Journal of the Korean
Statistical Society
Vol. 23, No. 1, 1994

An Asymptotic Property of Multivariate Autoregressive Model with Multiple Unit Roots

Keyil Shin [1]

ABSTRACT

To estimate coefficient matrix in autoregressive model, usually ordinary least squares estimator or unconditional maximum likelihood estimator is used. It is unknown that for univariate AR(p) model, unconditional maximum likelihood estimator gives better power property than ordinary least squares estimator in testing for unit root with mean estimated. When autoregressive model contains multiple unit roots and unconditional likelihood function is used to estimate coefficient matrix, the separation of nonstationary part and stationary part of the eigenvalues in the estimated coefficient matrix in the limit is developed. This asymptotic property may give an idea to test for multiple unit roots.

KEYWORDS : Autoregressive model, Unconditional likelihood function, Nonstationarity, Multiple unit roots.

[1]Department of Statistics, Hankuk University of Foreign Studies, Yongin-Gun, Gyounggi-Do, 449-791, Korea

1. INTRODUCTION

Consider the k-dimensional multivariate first-order autoregressive AR(1) process defined by the rule

$$Y_t = HY_{t-1} + \eta_t, t = 1, 2, \cdots \tag{1.1}$$

where

$$Y_t = [Y_{1,t}, Y_{2,t}, \cdots, Y_{k,t}]',$$

$$\eta_t = [\eta_{1,t}, \eta_{2,t}, \cdots, \eta_{k,t}]',$$

$$Y_0 = \phi,$$

$\{\eta_t : t = 1, 2, \cdots\}$ is a sequence of independent and identically distributed multivariate normal variates with mean ϕ (a vector of 0's) and variance Ω and H is a k by k coefficient matrix.

Assume that there exists a real matrix R such that $R^{-1}HR = A$ is a diagonal matrix. Let

$$A^* = \begin{bmatrix} I_{rr} & \phi' \\ \phi & A^*_{22} \end{bmatrix} \tag{1.2}$$

where I_{rr} is an r by r identity matrix and A^*_{22} is a diagonal matrix with elements less than 1 in magnitude. When the matrix H with $A = A^*$ in the model defined in (1.1) we say that H has r multiple unit roots and the rest less than 1 in magnitude.

The transformed model by R^{-1} is

$$X_t = AX_{t-1} + \epsilon_t, t \geq 1 \tag{1.3}$$

where $X_t = R^{-1}Y_t$ and $\epsilon_t = R^{-1}\eta_t$. Since we are concerned with the eigenvalues of H the model transformed by R^{-1} gives the same eigenvalues. Therefore we will use the model (1.3) from now on.

Johansen (1988) investigated multivariate AR(p) model with the H matrix satifying rank $(H - I) = k - r$. Fountis and Dickey (1989) showed that the nonstationary part and the stationary part of the ordinary least squares estimators can be separated in the limit for an AR(p) model with one unit root and the rest less than 1 in magnitude. That is, the normalized information matrix converges to a block diagonal matrix. Shin (1994) extended this property to the multiple unit roots case. Shin (1992) showed that this separation also occurs when one uses unconditional maximum likelihood estimator for AR(p) model with one unit root and the rest less than one in magnitude. Here we call the likelihood function with initial values random "the unconditional likelihood function". See Cox (1991) and Cox and Llatas (1991). In section 2, we review the multiple unit roots case when ordinary least squares estimator is used to estimate the coefficient matrix in an AR(p) model. In section 3, we show that this separation also occurs when unconditional maximum likelihood estimator for AR(p) model with multiple unit roots is used.

2. MULTIVARIATE AR(1) MODEL : O.L.S

In this section we briefly review an asymptotic property developed by Shin (1994).

Consider the k-dimensional multivariate AR(1) model defined in (1.3) with $X_0 = 0$. Now we assume that data are generated by $X_t = A^* X_{t-1} + e_t, t \geq 1$, that is,

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} I_{rr} & \phi' \\ \phi & A^*_{22} \end{bmatrix} \begin{bmatrix} X_{1t-1} \\ X_{2t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}, t \geq 1 \tag{2.1}$$

where $X_{10} = 0, X_{20} = 0$ and e_t's are i.i.d $N(0, \Sigma)$.

The usual ordinary least squares estimator for the matrix A in (1.3) is

$$\begin{aligned} \hat{A}_{ols} &= \sum_{t=1}^{n}(X_t X'_{t-1})(\sum_{t=1}^{n} X_{t-1}X'_{t-1})^{-1} \\ &= \begin{bmatrix} \sum_{t=1}^{n} X_{1t}X'_{1t-1} & \sum_{t=1}^{n} X_{1t}X'_{2t-1} \\ \sum_{t=1}^{n} X_{2t}X'_{1t-1} & \sum_{t=1}^{n} X_{2t}X'_{2t-1} \end{bmatrix} \begin{bmatrix} \sum_{t=1}^{n} X_{1t-1}X'_{1t-1} & \sum_{t=1}^{n} X_{1t-1}X'_{2t-1} \\ \sum_{t=1}^{n} X_{2t-1}X'_{1t-1} & \sum_{t=1}^{n} X_{2t-1}X'_{2t-1} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \hat{A}_{ols11} & \hat{A}_{ols12} \\ \hat{A}_{ols21} & \hat{A}_{ols22} \end{bmatrix}. \end{aligned} \tag{2.2}$$

Fountis and Dickey (1989) showed that for a single unit root, the properly normalized $n(\hat{A}_{ols11}-1)$ converges weakly to ξ/Γ where (ξ,Γ) is the weak limit of $(\sum_{t=1}^{n} e_{1t}X_{1t-1}/n, \sum_{t=1}^{n} X_{1t-1}^2/n^2)$ where $e_{1t}=X_{1t}-X_{1t-1}$. This is the same limiting distribution for the univariate model studied by Dickey and Fuller (1979). Thus we say that the nonstationary part and the stationary part can be separated for the AR(1) model with one unit root. Shin (1994) showed

$$n(\hat{A}_{ols11}-I)=[\sum_{t=1}^{n} e_{1t}X'_{1t-1}/n][\sum_{t=1}^{n} X_{1t-1}X'_{1t-1}/n^2]^{-1}+o_p(1) \quad (2.3)$$

where X_{1t} is an r-dimensional nonstationary vector process and $e_{1t}=X_{1t}-X_{1t-1}$.

Let

$$n(\hat{B}_{ols11}-I)=[\sum_{t=1}^{n} e_{1t}X'_{1t-1}/n][\sum_{t=1}^{n} X_{1t-1}X'_{1t-1}/n^2]^{-1}. \quad (2.4)$$

Then $\hat{B}_{ols11}$ is the ordinary least squares estimator of the coefficient matrix in the r-dimensional AR(1) model with all unit roots. Shin (1994) showed that (2.3) and (2.4) have the same limiting distribution of the estimated eigenvalues. Therefore when one uses the ordindary least squares estimator, regardless of the stationary part, we have the same limiting distribution for the nonstationary part.

3. MULTIVARIATE AR(1) MODEL : M.L.E

Consider the k-dimensional multivariate AR(1) model

$$X_t = AX_{t-1}+\epsilon_t, t\geq 1$$

where ϵ_t's are i.i.d $N(0,\Sigma)$. By the Yule-Walker equations we have

$$V = AVA'+\Sigma$$

where V is the variance matrix of $X_t, t\geq 1$. Now we assume that the distribution of X_0 is $N(0,V)$. With this setup the unconditional likelihood function for a k-dimensional stationary multivariate process X_t is L where, for data $X_0, X_1, \ldots, X_n$,

$$Ln(L) = \frac{-(n+1)k}{2} ln(2\pi) - \frac{1}{2} ln(|V|) - \frac{n}{2} ln(|\Sigma|) - \frac{1}{2} X_0{}' V^{-1} X_0$$
$$-\frac{1}{2} \sum_{t=1}^{n} (X_t - AX_{t-1})' \Sigma^{-1} (X_t - AX_{t-1}). \tag{3.1}$$

If A has unit roots, the determinant of V matrix is infinite and the unconditional maximum likelihood estimator can not be found. So we use the unconditional stationary likelihood function which can be maximized for any set of data.

We now consider this unconditional likelihood under the null hypothesis that the data have r unit roots. For the moment assume data generated by $X_t = A^* X_{t-1} + e_t$ where A^* is defined in (1.2).

To get a maximum likelihood estimator, take the derivatives of (3.1) with respect to A. Using the results of McDonald and Swaminathan (1973) we have

i) $\partial\{-1/2 \sum_{t=1}^{n} (X_t - AX_{t-1})' \Sigma^{-1} (X_t - AX_{t-1})\} / \partial A =$

$$\Sigma^{-1} \sum_{t=1}^{n} (X_t - AX_{t-1}) X'_{t-1} \tag{3.2}$$

ii) $\partial ln(|V|) / \partial vec(A')$

$$= \{(I \otimes VA') + E(VA' \otimes I)\}\{I - A' \otimes A'\}^{-1} vec(2V^{-1} - diag V^{-1}) \tag{3.3}$$

iii) $\partial\{-1/2 X_0{}' V^{-1} X_0\} / \partial vec(A')$

$$= 1/2[(I \otimes VA') + E(VA' \otimes I)\}\{I - A' \otimes A'\}^{-1}][V^{-1} X_0 \otimes V^{-1} X_0] \tag{3.4}$$

where $E = \partial vec'(A) / \partial vec(A')$.

This complicated equations does not give an explicit form of the maximum likelihood estimator of A. To make it easier to solve this equation we assume that the derivative matrix of (3.1) with respect to A has unique solution which maximizes (3.1).

Motivated by the least squares results of Shin (1994) we anticipate the $\hat{A}_{MLE}$ which maximizes (3.1) has probability arbitrarily close to 1 of being in the set S_M of matrices

$$\hat{A}_{MLE} \in S_M = diag(1, \cdots, 1, \alpha_{r+1}, \cdots, \alpha_k) + \begin{bmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1k} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \delta_{k1} & \delta_{k2} & \cdots & \delta_{kk} \end{bmatrix} \quad (3.5)$$

where for $0 \le M_1 < M_2, M_1/n < |\delta_{ii}| < M_2/n, 1 \le i \le r, |\delta_{ij}| < M_2/n, i \ne j$, $1 \le i \le r, j \le k$ and $|\delta_{ij}| < M_2/n^{1/2}, r+1 \le j \le k$. These are the convergence orders obtained by Shin (1994).

By Lemmas 4 and 5 in appendices, we have for $A \in S_M$

i) $[-1/2\partial ln(|V|)/\partial A]D_n^{-1} =$

$$\begin{bmatrix} -1/2n\partial ln(|V_{11}|)/\partial A_{11} & \phi' \\ \phi & O \end{bmatrix} + O(n^{-1/2}) = O(1) \quad (3.6)$$

where ϕ and O are proper zero matrices.

ii) $[\partial\{-1/2X_0{}'V^{-1}X_0\}/\partial A]D_n^{-1} = O_p(n^{-1/2})$. (3.7)

Here $D_n = diag(n, \ldots, n, n^{1/2}, \ldots, n^{1/2})$.

Now multiplying D_n^{-1} times the derivatives of $Ln(L)$ with respect to A and setting these equal to 0 we have

$$\begin{bmatrix} -1/2n\partial ln(|V_{11}|)/\partial A_{11} & \phi' \\ \phi & O \end{bmatrix} + \Sigma^{-1}\sum_{t=1}^{n}(X_t - AX_{t-1}\}X'_{t-1}D_n^{-1}$$
$$+O_p(n^{-1/2}) = 0.$$

This equation can be written as

$$\Sigma\begin{bmatrix} -1/2n\partial ln(|V_{11}|)/\partial A_{11} & \phi' \\ \phi & O \end{bmatrix} + \sum_{t=1}^{n}\{e_t - (A - A^*)X_{t-1})X'_{t-1}D_n^{-1}$$

$$+O_p(n^{-1/2}) = 0. \tag{3.8}$$

where $e_t = X_t - A^* X_{t-1}$. Now since for $A \in S_M, -1/2n\partial ln(|V_{11}|)/\partial A_{11} = O(1), \sum_{t=1}^n e_t X'_{t-1} D_n^{-1} = O_p(1)$ and $D_n^{-1}\sum_{t=1}^n X_{t-1}X'_{t-1}D_n^{-1} = O_p(1)$ we have the maximum likelihood estimator in S_M.

Since $D_n^{-1}\sum_{t=1}^n X_{t-1}X'_{t-1}D_n^{-1}$ converges to a block diagonal matrix, by section 2, the upper left corner block matrix of (3.8) is

$$-1/2n\Sigma_{11}\partial ln(|V_{11}|)/\partial A_{11}$$

$$+\sum_{t=1}^{n}\{e_{1t} - (A_{11} - I)X_{1t-1}\}X'_{1t-1}/n + o_p(1) = 0. \tag{3.9}$$

Therefore the unconditional maximum likelihood estimator can be separated into two parts, the nonstationary part and the stationary part, in the limit.

4. APPENDICES

Lemma 1. For $A \in S_M$ defined in (3.5),

$$vec(V_{11}) = (I - A_{11} \otimes A_{11})^{-1} vec(\Sigma_{11}) + O(n^{1/2}) = O(n),$$

$$vec(V_{12}) = (I - A_{22} \otimes A_{11})^{-1}[(A_{21} \otimes A_{11})vec(V_{11}) + vec(\Sigma_{12})]$$

$$+O(n^{-1/2}) = O(1)$$

and

$$vec(V_{22}) = (I - A_{22} \otimes A_{22})^{-1} vec(\Sigma_{22}) + O(n^{-1/2}) = O(1). \tag{4.1}$$

Proof. By Yule-Walker equations we have $V = AVA' + \Sigma$. Now using vec operations we have $vec(V) = (I - A \otimes A)^{-1} vec(\Sigma)$.
Let F be a permutation matrix such that for a k^2 by k^2 B matrix, $F\ vec(B) = vec(C)$ where

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \qquad vec(C) = \begin{bmatrix} vec(B_{11}) \\ vec(B_{21}) \\ vec(B_{12}) \\ vec(B_{22}) \end{bmatrix}.$$

Then we have

$$F(B \otimes B)F' =$$

$$\begin{bmatrix} B_{11} \otimes B_{11} & B_{11} \otimes B_{12} & B_{12} \otimes B_{11} & B_{12} \otimes B_{12} \\ B_{11} \otimes B_{21} & B_{11} \otimes B_{22} & B_{12} \otimes B_{21} & B_{12} \otimes B_{22} \\ B_{21} \otimes B_{11} & B_{21} \otimes B_{12} & B_{22} \otimes B_{11} & B_{22} \otimes B_{12} \\ B_{21} \otimes B_{21} & B_{21} \otimes B_{22} & B_{22} \otimes B_{21} & B_{22} \otimes B_{22} \end{bmatrix}. \tag{4.2}$$

Now since $F^{-1} = F'$ we have

$$Fvec(V) = \begin{bmatrix} vec(V_{11}) \\ vec(V_{21}) \\ vec(V_{12}) \\ vec(V_{22}) \end{bmatrix} = \{F(I - A \otimes A)F'\}^{-1} Fvec(\Sigma) \tag{4.3}$$

$$= \begin{bmatrix} I - A_{11} \otimes A_{11} & A_{11} \otimes A_{12} & A_{12} \otimes A_{11} & A_{12} \otimes A_{12} \\ A_{11} \otimes A_{21} & I - A_{11} \otimes A_{22} & A_{12} \otimes A_{21} & A_{12} \otimes A_{22} \\ A_{21} \otimes A_{11} & A_{21} \otimes A_{12} & I - A_{22} \otimes A_{11} & A_{22} \otimes A_{12} \\ A_{21} \otimes A_{21} & A_{21} \otimes A_{22} & A_{22} \otimes A_{21} & I - A_{22} \otimes A_{22} \end{bmatrix}^{-1} Fvec(\Sigma).$$

Using the inverse formula of the partition matrix we have the results.

Lemma 2. For $A \in S_M$ defined in (3.5),

$$[VA']_{11} = V_{11}A'_{11} + O(n^{1/2}) = O(n),$$

$$[VA']_{12} = O(1), [VA']_{21} = O(1) \quad \text{and} [VA']_{22} = O(1). \tag{4.4}$$

Also $(2V^{-1} - diagV^{-1})_{11} = 2V_{11}^{-1} - diagV_{11}^{-1} + O(n^{-3/2}) = O(n^{-1})$,

$$(2V^{-1} - diagV^{-1})_{12} = O(n^{-1}) \quad \text{and} (2V^{-1} - diagV^{-1})_{22} = O(1). \tag{4.5}$$

Proof. From Lemma 1 we have the orders of the V matrix elements.

Therefore for $A \in S_M$ we have the results in (4.4). Also using the inverse formula for a partitioned matrix we have (4.5).

Lemma 3.

$$\{F(I \otimes VA')F' + FEF'F(VA' \otimes I)F'\} =$$

$$\begin{bmatrix} \{(I \otimes VA'_{11}) + G_{11}(VA'_{11} \otimes I)\} + O(n^{1/2}) & O(1) & O(1) & \phi \\ O(1) & O(1) & O(1) & \phi \\ \phi & O(n) & O(n) & O(1) \\ \phi & O(1) & O(1) & O(1) \end{bmatrix} \tag{4.6}$$

and

$$\{F(I - A' \otimes A')F'\}^{-1} =$$

$$\begin{bmatrix} (I - A'_{11} \otimes A'_{11})^{-1} + O(n^{1/2}) & O(1) & O(1) & O(n^{-1}) \\ O(n^{1/2}) & O(1) & O(n^{-1/2}) & O(n^{-1/2}) \\ O(n^{1/2}) & O(n^{-1/2}) & O(1) & O(n^{-1/2}) \\ O(1) & O(n^{-1/2}) & O(n^{-1/2}) & O(1) \end{bmatrix} \tag{4.7}$$

where $E = \partial vec'(B)/\partial vec(B'), G_{ij} = \partial vec'(B_{ij})/\partial vec(B'_{ij}), i, j = 1, 2$.

Proof. Using F defined in Lemma 1 we have

$$F(I \otimes VA')F' =$$

$$\begin{bmatrix} I \otimes (VA')_{11} & I \otimes (VA')_{12} & \phi & \phi \\ I \otimes (VA')_{21} & I \otimes (VA')_{22} & \phi & \phi \\ \phi & \phi & I \otimes (VA')_{11} & I \otimes (VA')_{12} \\ \phi & \phi & I \otimes (VA')_{21} & I \otimes (VA')_{22} \end{bmatrix}$$

and

$$F(VA' \otimes I)F' =$$

$$\begin{bmatrix} (VA')_{11}\otimes I & \phi & (VA')_{12}\otimes I & \phi \\ \phi & (VA')_{11}\otimes I & \phi & (VA')_{12}\otimes I \\ (VA')_{21}\otimes I & \phi & (VA')_{22}\otimes I & \phi \\ \phi & (VA')_{21}\otimes I & \phi & (VA')_{22}\otimes I \end{bmatrix}.$$

Now for $E = \partial vec'(B)/\partial vec(B')$ we have

$$G = FEF' = \begin{bmatrix} G_{11} & \phi & \phi & \phi \\ \phi & \phi & G_{12} & \phi \\ \phi & G_{21} & \phi & \phi \\ \phi & \phi & \phi & G_{22} \end{bmatrix}$$

where $G_{ij} = \partial vec'(B_{ij})/\partial vec(B'_{ij}), i,j = 1,2$ and B is defined in Lemma 1.

$$FEF'F(VA'\otimes I)F' = GF(VA'\otimes I)F'$$

$$\begin{bmatrix} G_{11}(VA')_{11}\otimes I & \phi & G_{11}(VA')_{12}\otimes I & \phi \\ G_{12}(VA')_{21}\otimes I & \phi & G_{12}(VA')_{22}\otimes I & \phi \\ \phi & G_{21}(VA')_{11}\otimes I & \phi & G_{21}(VA')_{12}\otimes I \\ \phi & G_{22}(VA')_{21}\otimes I & \phi & G_{22}(VA')_{22}\otimes I \end{bmatrix}.$$

Now using Lemma 2 we have (4.6). Also By Lemma 1 we have (4.7).

Lemma 4. For $A \in S_M$ defined in (3.5),

$$[-1/2\partial ln(|V|)/\partial A']_{11} = -1/2\partial ln(|V_{11}|)/\partial A_{11} + O(n^{1/2}) = O(n),$$

$$[-1/2\partial ln(|V|)/\partial A']_{12} = O(1),$$

$$[-1/2\partial ln(|V|)/\partial A']_{21} = O(n^{1/2}) \quad \text{and}$$

$$[-1/2\partial ln(|V|)/\partial A']_{22} = O(1).$$

Proof. McDonald and Swaminathan (1973) showed that

$$vec(\partial ln(|V|)/\partial A')$$

$$= \{(I\otimes VA') + E(VA'\otimes I)\}\{I - A'\otimes A'\}^{-1}vec(2V^{-1} - diagV^{-1}) \quad (4.7)$$

where $E = \partial vec'(B)/\partial vec(B')$.
Now let F be the same matrix defined in Lemma 1. Then premultiplying (4.8) by F we have

$$Fvec(\partial ln(|V|)/\partial A') =$$

$$\{F(I \otimes VA')F' + FEF'F(VA' \otimes I)F'\}\{F(I - A' \otimes A')F'\}^{-1}$$

$$Fvec(2V^{-1} - diagV^{-1}).$$

Then using Lemmas 2 and 3 we have the results.

Lemma 5. For $A \in S_M$ defined in (3.5), $[\partial\{-1/2X_0'V^{-1}X_0\}/\partial A]D_n^{-1} = O_p(n^{-1/2})$ where $D_n = diag(n, \cdots, n, n^{1/2}, \cdots, n^{1/2})$.

Proof. McDonald and Swaminathan (1973) showed that

$$\partial\{X_0'V^{-1}X_0\}/\partial vec(A') = [-\partial vec(V)/\partial vec(A')][V^{-1}X_0 \otimes V^{-1}X_0].$$

Since $[\partial vec(V)/\partial vec(A')] = \{(I \otimes VA') + E(VA' \otimes I)\}\{I - A' \otimes A'\}^{-1}$, by Lemmas 1 and 3 we have the results.

5. REFERENCES

(1) Cox, D. D. (1991). Gaussian Likelihood Estimation for Nearly Nonstationary AR(1) processes. *The Annals of Statistics*, 19, 1129 - 1142.

(2) Cox, D. D., and Llatas, I. (1991). The Maximum Likelihood Type Estimation for Nearly Nonstationary Autoregressive Time Series. *The Annals of Statistics*, 19, 1109 - 1128.

(3) Dickey, D. A., and Fuller, W. A. (1979). Distribution of the Estimators for Autoregerssive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, 427 - 431.

(4) Fountis, N. G., and Dickey, D. A. (1989). Testing for a unit Root Nonstationarity in Multivariate Autoregressive Time Series. *The Annals of Statistics*, 17, 419 - 428.

(5) McDonald, R. P., and Swaminathan, M. (1973). A Simple Matrix Calculus with Applications to Multivariate Analysis. *General Systems*, 18, 37 -54.

(6) Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, 12, 231 - 254.

(7) Shin, k. (1992). A Unit Root Test for Multivariate Autoregressive Time Series. Ph.D. Thesis. North Carolina State University, Raleigh, North Carolina.

(8) Shin, K. (1994). Testing a Multivariate Process for Multiple Unit Roots, *The Korean Journal of Applied Statistics*, 7, 103 - 112.