Journal of the Korean
Statistical Society
Vol. 26, No. 3, 1997

A Unit Root Test for Multivariate Autoregressive Model with Multiple Unit Roots

Key Il Shin[1]

Abstract

Recently maximum likelihood estimators using unconditional likelihood function are used for testing unit roots. When one wants to use this method the determinant term of initial values in the multivariate unconditional likelihood function produces a complicated function of the elements in the coefficient matrix and variance matrix. In this paper an approximation of the determinant term is calculated and based on this approximation an approximated unconditional likelihood function is calculated. The approximated unconditional maximum likelihood estimators can be used to test for unit roots. When multivariate process has one unit root the limiting distribution obtained by this method and the limiting distribution using exact unconditional likelihood function are the same.

Key Words : Autoregressive model; Unconditional likelihood function; Cointegration; Unit root.

[1]Department of Statistics, Hankuk university of Foreign Studies, Yong in-Si, Gyunggi-Do, 449-791, Korea.

1. INTRODUCTION

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

$$X_t = AX_{t-1} + \epsilon_t, \ t \geq 1 \tag{1.1}$$

where $X_0 = 0$ and $\{\epsilon_t : t = 1, 2, \cdots\}$ is a sequence of independent and identically distributed multivariate normal variates with mean 0 and variance matrix Σ. When all eigenvalues of A are less than 1 we say that this process is stationary. If not we say that this process is nonstationary. Especially if one or more than one eigenvalues of A are equal to 1 we say that process has unit root(s). When $A = I_k$, a k by k identity matrix, differencing makes process stationary. Hence the testing for null hypothesis $A = I_k$ is of interest. When A has r, $1 \leq r \leq k-1$, unit roots, one may be interested in finding some vectors which are the coefficients of the combination of the elements of X_t being stationary. These vectors are known as cointegrating vectors. See Engle and Granger(1987) and Murry(1994). Since the number of cointegrating vectors is the same as the number of the eigenvalues of A less than 1 in magnitude, the unit roots test is also of intest.

Phillips and Durlauf(1986) suggested several estimators for testing $A = I_k$. Fountis and Dickey(1989) showed that the nonstationary part(unit root 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 limiting distribution is the same as that of the univariate process studied by Dickey and Fuller(1979) even though there exist stationary processes. Also Johansen(1988) studied multiple unit roots problem.

In section 2 we develop the approximation of the determinant of the variance matrix and the approximated unconditional likelihood function based on this approximation. In section 3 we study maximum likelihood estimators based on the approximated unconditional likelihood function when process has multiple unit roots and others less than one in magnitude, In section 4 we make an empirical distribution of the smallest eigenvalue of maximum likelihood estimators of A based on the approximated unconditional likelihood function. In section 5 we have some concluding remarks.

2. APPROXIMATION OF UNCONDITIONAL LIKELIHOOD FUNCTION

Consider the k-dimensional multivariate AR(1) model

$$Y_t = BY_{t-1} + \eta_t,\ t \geq 1 \tag{2.1}$$

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

$$\Gamma = B\Gamma B' + \Omega \tag{2.2}$$

where Γ is the variance of $Y_t, t \geq 1$. Now we assume that the distribution of Y_0 is $N(0, \Gamma)$. With this setup the unconditional likelihood function for the k-dimensional stationary multivariate process Y_t is L where, for data, $Y_0, Y_1, \cdots, Y_n$,

$$\begin{aligned} Ln(L) &= -(n+1)k/2ln(2\pi) - 1/2ln(|\Gamma|) - n/2ln(|\Omega|) - 1/2Y_0'\Gamma^{-1}Y_0 \\ &\quad -1/2\sum_{t=1}^{n}(Y_t - BY_{t-1})'\Omega^{-1}(Y_t - BY_{t-1}). \end{aligned} \tag{2.3}$$

Note that when B has unit roots, the Γ matrix in (2.2) can not be defined. So (2.3) is no more loglikelihood function. However one can maximize (2.3) for any set of data including nonstationary process. Also the Γ matrix in (2.3) produces difficult calculations to obtain maximum likelihood estimators of B. For the stationary process when n is large one can ignore the effect of initial condition. But when the process is nonstationary one can not do that.

Consider eigenvalue-eigenvector decomposition of B. Let S be the eigenvector matrix such that $S^{-1}BS = A$ and $|S| = 1$ where $A = diag(\alpha_1, \alpha_2, \cdots, \alpha_k)$, $|\alpha_i| < 1, i = 1, \cdots, k$, is the eigenvalue matrix which is diagonal. We only consider B matrix which can be decomposed by $S^{-1}BS = A$. Now pre-multiplying S^{-1} and post-multiplying $S^{-1'}$ to (2.2) we have

$$V = AVA' + \Sigma \tag{2.4}$$

where $V = S^{-1}\Gamma S^{-1'}$ and $\Sigma = S^{-1}\Omega S^{-1'}$.

Let σ_{ij} be the ij-th element of Σ. Then from (2.4) we have

$$V = \begin{bmatrix} \sigma_{11}/(1-\alpha_1^2) & \sigma_{12}/(1-\alpha_1\alpha_2) & \cdots & \sigma_{1k}/(1-\alpha_1\alpha_k) \\ \sigma_{21}/(1-\alpha_1\alpha_2) & \sigma_{22}/(1-\alpha_2^2) & \cdots & \sigma_{2k}/(1-\alpha_2\alpha_k) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1}/(1-\alpha_1\alpha_2) & \sigma_{k2}/(1-\alpha_2\alpha_k) & \cdots & \sigma_{kk}/(1-\alpha_k^2) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^*_{11} & \sigma^*_{12} & \cdots & \sigma^*_{1k} \\ \sigma^*_{21} & \sigma^*_{22} & \cdots & \sigma^*_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^*_{k1} & \sigma^*_{k2} & \cdots & \sigma^*_{kk} \end{bmatrix} (I - A^2)^{-1}$$

$$= \Sigma^*(I - A^2)^{-1} \tag{2.5}$$

where $\sigma^*_{ij} = \sigma_{ij}(1-\alpha_j^2)/(1-\alpha_i\alpha_j)$ and σ^*_{ij} is the ij-th element of Σ^*

Therefore since $|S| = 1$ we have $|\Gamma| = |V| = |\Sigma^*||I - A^2|^{-1} = |\Sigma^*||I + A|^{-1}|I - A|^{-1}$. Hence

$$|\Gamma| = |\Sigma^*||I + B|^{-1}|I - B|^{-1} \tag{2.6}$$

So (2.3) becomes

$$\begin{aligned} Ln(L) &= -(n+1)k/2ln(2\pi) - 1/2ln(|\Sigma^*|) + 1/2ln(|I+B|) \\ &\quad +1/2ln(I-B) - n/2ln(|\Omega|) - 1/2Y_0'\Gamma^{-1}Y_0 \\ &\quad -1/2\sum_{t=1}^{n}(Y_t - BY_{t-1})'\Omega^{-1}(Y_t - BY_{t-1}). \end{aligned} \tag{2.7}$$

Let us consider the model transformed by S^{-1} ,

$$X_t = AX_{t-1} + \epsilon_t,\ t \geq 1 \tag{2.8}$$

where $X_t = S^{-1}Y_t$, $A = S^{-1}BS$ and $\epsilon_t = S^{-1}\eta_t$.

Note that because A and B are similar matrices the eigenvalues of A are also the eigenvalues of B.

Let m_{im} be the eigenvalues of A_m which is the maximum likelihood estimators of A. Then we have $|m_{im}| < 1$ because A_m is obtained subject to (2.4). Therefore for fixed Σ, we have the maximum likelihood estimators of Σ^*, say $\Sigma^*_m = \left\{\sigma^*_{ijm}\right\}$,

$$\begin{aligned} \sigma^*_{ijm} &= \sigma_{ij}(1 - m^2_{jm})/(1 - m_{im}m_{jm}) \\ &= \sigma_{ij}(1 + m_{jm})\left\{m_{im} + (1 - m_{im})/(1 - m_{jm})\right\}. \end{aligned} \tag{2.9}$$

Since for any $i, j \leq k$, $m_{im} + (1 - m_{im})/(1 - m_{jm})$ does not go to 0 in probability we have $\sigma^*_{ijm} = O_p(1)$. Let V_m and Γ_m be the maximum likelihood estimators of V and Γ respectively. Then since V_m^{-1} is $O_p(1)$, see Shin(1994b), we have $\Gamma_m^{-1} = O_p(1)$. So we can ignore these three terms $-1/2Y_0'\Gamma^{-1}Y_0$,$1/2ln(|I+B|)$ and -1/2 $ln(|\Sigma^*|)$ which do not affect the maximum likelihood estimators of B for large n.

Therefore the approximated loglikelihood function is

$$\begin{aligned} Ln(L) &= -(n+1)k/2ln(2\pi) + 1/2(|I-B|) - n/2ln(|\Omega|) \\ &\quad -1/2\sum_{t=1}^{n}(Y_t - BY_{t-1})'\Omega^{-1}(Y_t - BY_{t-1}). \end{aligned} \tag{2.10}$$

Note that (2.2) makes the eigenvalues of B_m, maximum likelihood estimators of B, less than 1 and (2.3) contains this property for existence of maximum likelihood estimators of Γ matrix. But since $|I - B_m|$ can be positive with some of the eigenvalues of B_m which are greater than 1, (2.10) may not give the eigenvalues of B_m less than 1. So we restrict B_m to have eigenvalues of B_m less than one.

3. MAXIMUM LIKELIHOOD ESTIMATORS

In this section we develop the maximum likelihood estimators which maximizes (2.10). Consider the k-dimensional multivariate AR(1) model

$$X_t = AX_{t-1} + \epsilon_t, \ t \geq 1 \tag{3.1}$$

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

$$V = AVA' + \Sigma \tag{3.2}$$

where V is the variance of X_t, $t \geq 1$ and $\Sigma = \{\sigma_{ij}\}$. Now we assume that the distribution of X_0 is $N(0, V)$. Assume that data are generated by

$$X_t = A^* X_{t-1} + \epsilon_t, \ t \geq 1 \tag{3.3}$$

where $A^* = diag(1, 1, \cdots, 1, \alpha_{r+1}, \cdots \alpha_k)$ is a diagonal matrix with $|\alpha_i| < 1, i = r+1, \cdots, k$. We use the approximated likelihood function instead of unconditional likelihood function for maximum likelihood estimators. Now the approximated likelihood function developed in section 2 is

$$\begin{aligned} Ln(L) &= -(n+1)k/2ln(2\pi) + 1/2ln(|I-A|) - n/2ln(|\Sigma|) \\ &\quad -1/2\sum_{t=1}^{n}(X_t - AX_{t-1})'\Sigma^{-1}(X_t - AX_{t-1}). \end{aligned} \tag{3.4}$$

Let us take derivatives of (3.4) with respect to A.

$$i) \quad \partial\{-1/2\sum_{t=1}^{n}(X_t - AX_{t-1})'\Sigma^{-1}(X_t - AX_{t-1})\}/\partial A$$
$$= \Sigma^{-1}\sum_{t=1}^{n}(X_t - AX_{t-1})X'_{t-1} \quad (3.5)$$
$$ii) \quad [1/2\partial ln(|I - A|)/\partial A] = -1/2\{(I - A')\}^{-1} \quad (3.6)$$

Note that comparing (3.6) with the exact derivatives of $-1/2ln(|V|)$ with respect to A in Shin(1994b) we have much simpler expression.

Let $D_n = diag(n, \cdots, n, n^{1/2}, \cdots, n^{1/2})$. Then $\{\partial Ln(L)/\partial A\}D_n^{-1} = 0$ becomes

$$-1/2\{D_n(I - A')\}^{-1} + \Sigma^{-1}\sum_{t=1}^{n}(X_t - AX_{t-1})X'_{t-1}D_n^{-1} = 0 \quad (3.7)$$

and so

$$-1/2\{D_n(I - A')\}^{-1} + \Sigma^{-1}\sum_{t=1}^{n}\{e_t + (A^* - A)X_{t-1}\}X'_{t-1}D_n^{-1} = 0 \quad (3.8)$$

where $e_t = X_t - A^*X_{t-1}$.

From Shin(1994b) and the upper left block matrix of (3.8) we have

$$-1/2\Sigma_{11}\{n(I - A'_{11})\}^{-1} + \sum_{t=1}^{n}\{e_{1t} + (I - A_{11})X_{1t}\}X'_{1t-1}/n + O_p(n^{-1/2}) = 0 \quad (3.9)$$

where Σ_{11} and A_{11} are r by r upper left block matrix of Σ and A_{11} corresponding to X_{1t} respectively. Let $P^* = n(I - A'_{11})$, $Q^* = \sum_{t=1}^{n}(e_{t1}X'_{1t-1})/n$, $R^* = \sum_{t=1}^{n}(X_{1t-1}X'_{1t-1})/n^2$.

Then for large n (3.9) becomes

$$-1/2\Sigma_{11}P^{*-1} + Q^* + P^{*'}R^* = 0 \quad (3.10)$$

Since Σ_{11} is a positive definite matrix, there exists a symmetric matrix $\Sigma_{11}^{1/2}$ such that $\Sigma_{11} = \Sigma_{11}^{1/2}\Sigma_{11}^{1/2}$.

Now pre-multiplying $\Sigma_{11}^{-1/2}$ and post-multiplying $\Sigma_{11}^{-1/2}$ we have

$$-1/2P_m^{-1} + Q + P'_m R = 0. \quad (3.11)$$

where $P_m = \Sigma_{11}^{1/2}P^*\Sigma_{11}^{-1/2}$, $Q = \Sigma_{11}^{-1/2}Q^*\Sigma_{11}^{-1/2}$, $R = \Sigma_{11}^{-1/2}R^*\Sigma_{11}^{-1/2}$.

Phillips and Durlauf(1986) showed that Q converged weakly to $\int_0^1 W(t)$ $dW(t)$ and R converges weakly to $\int_0^1 W(t)W(t)'dt$ where $W(t)$ is a Brownian motion in k dimensions. Unfortunately (3.11) does not give a closed form of P_m which satisfies (3.11). Also this equation does not have a unique solution. Hence one can not use these equations to obtain an maximum likelihood estimators. However we have simpler expression than (3.9) in Shin (1994b). Also (2.10) is a simple form comparing to (2.3) one may obtain maximum likelihood estimators A from (2.10) numerically.

When the process is univariate, studied by Gonzalez-Farias(1992) and Pantula *at el.*(1994), we have $P_m = 1/2\{-R^{-1}Q' - [(QR)^2 + 2R^{-1}]^{1/2}\}$. Also as n goes infinity, P_m converges weakly to $1/2\{-\xi/\Gamma - (\xi^2/\Gamma^2 + 2/\Gamma)^{1/2}\}$ where (ξ, Γ) is the weak limit of $(\sum_{t=1}^n e_t X_{t-1}/n, \sum_{t=1}^n X_{t-1}^2/n^2)$ which is obtained by Gonzalez-Farias(1992).

4. EMPIRICAL DISTRIBUTION

For k-dimensional multivariate AR(1) model we have $k^2 + k(k+1)/2$ unknown parameters. These many parameters give difficulties to maximize (2.10). In this section we calculate the empirical distribution only for bivariate AR(1) model.

Taking derivatives of (3.4) with respect to Σ we have $\Sigma_m = \sum_{t=1}^n (X_t - A_m X_{t-1})(X_t - A_m X_{t-1})'/n$ which is the maximum likelihood estimators of Σ and A_m is the maximum likelihood estimators of A matrix. In this paper we use Σ_{ols}, ordinary least squares estimators obtained by substituting A_{ols} to A_m in Σ_m , instead of Σ_m for simplicity even though one can find A_m and Σ_m iteratively or simultaneously.

Table 1. Empirical distribution of $n(1 - \lambda_{small})$ for bivariate AR(1) model

Length of data	percentile				
	0.01	0.025	0.05	0.1	0.25
n=50	-22.21	-19.49	-16.70	-13.30	-8.80
n=100	-24.22	-20.21	-16.91	-13.68	-8.83
n=250	-25.15	-20.53	-17.10	-14.04	-9.73

We generate data with length $n = 50, 100, 250$ using SAS/RANNOR following the rule $X_t = X_{t-1} + \epsilon_t$,$t \geq 1$ where $X_0 = (0,0)'$,ϵ_t 's are i.i.d $N(0, I_2)$

and I_2 is a 2 by 2 identity matrix. First we find Σ_{ols} and then given Σ_{ols} we maximizes (2.10) using NLIN procedure in SAS. We repeat these steps 5000 times. Table 1 shows empirical distribution of $n(1-\lambda_{small})$ where λ_{small} is the smallest eigenvalues of A_m.

For univariate case, when the length of data, n, is greater than 100, Gonzalez-Farias(1992) showed that the power of approximated unconditional maximum likelihood estimator is better than that of ordinary least squares estimator. Also see Pantula *at el.*(1994) for power comparison among the several unit root test statistics.

For bivariate case with one unit root, Shin(1993a) showed that the unconditional maximum likelihood estimators have better power than the ordinary least squares estimators. So one may expect that this power advantage will carry over to the multivariate AR(1) model with multiple unit roots.

5. CONCLUDING REMARKS

Cointegration problem which has been rigorously studied is directly related with unit roots test. The existence of r unit roots of k-dimensional multivariate AR(1) process means that $k-r$ cointegrating vectors exist in this process. Consider S matrix defined in (2.4). Let $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$. Then $k-r$ by k lower block matrix $[S_{21}S_{22}]$ produces cointegrating vectors. For example, let $c_{k-r+i} = (s_{k-r+i,1}s_{k-r+i,2},\cdots,s_{k-r+i,k})$ be the i^{th} row of $[S_{21}S_{22}]$. Then c_{k-r+i} is a cointegrating vector.

Once we know the number of cointegrating vectors, or the number of unit roots, we may find the cointegrating vectors by simply using regression method. That is first using Phillips' Triangular Representation of cointegrated system we have $S_{22} = I_{k-r}$. Then regress X_{2t} on X_{1t}. The estimated coefficient matrix of X_{1t} gives us the estimated cointegrating matrix S_{21}. See Hamilton(1994) for more details. So the multiple unit roots test plays an important role to the cointegration problem.

In this paper we do not study cointegration problem. But more powerful unit roots test gives better idea to find cointegrating vectors. Even though the P_m matrix defined in (3.11) does not have a closed form and obtaining the eigenvalues of the approximated unconditional maximum likelihood estimator P_m needs complicated calculations, this approach can be used for cointegration problem.

REFERENCES

(1) Engle, R.F., and Granger, C.W.I(1987). Co-Integration and Error correction : Representation, Estimation, and Testing. *Econometrica*, **55**, 2, 251-276.

(2) 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.

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

(4) Gonzalez-Farias, G.(1992). *A New Unit Root Test For Autoregressive Time Series*, Ph.D. Thesis, North Carolina State University, Raleigh, North Carolina.

(5) Hamilton, J.D.(1994). *Time Series Analysis*, Princeton University Press.

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

(7) Murray, M.P.(1994). A Drunk and Her Dog : An Illustrations of Cointegration and Error Correction. *The American Statistician*, **48**, 1, 37-39.

(8) Pantula, S.G., Gonzalez-Farias, G., and Fuller, W.A.(1994). A Comparison of Unit Root Test Criteria. *Journal of Business and Economic Statistics*, **12**, 4, 449-450.

(9) Phillips, P. C. B., and Durlauf, S. N.(1986). Multiple Time Series Regression with Integrated processes. *Review of Economic Studies*, **53**, 473-495.

(10) Shin, Key-Il.(1993a). *A Unit Root Test for Multivariate Autoregressive Time Series*, Ph.D. Thesis, North Carolina State University, Raleigh, North Carolina.

(11) Shin, Key-Il.(1994b). An Asymptotic Property of Multivariate Autoregressive Model with Multiple Unit Roots. *Journal of the Korean Statistical Society*, **23**, 1, 167-178.