Journal of the Korean Statistical Society (1999), **28**: *1, pp 45–55*

A Multiple Unit Roots Test Based on Least Squares Estimator

Key-Il Shin[1]

ABSTRACT

Knowing the number of unit roots is important in the analysis of k-dimensional multivariate autoregressive process. In this paper we suggest simple multiple unit roots test statistics based on least squares estimator for the multivariate AR(1) process in which some eigenvalues are one and the rest are less than one in magnitude. The empirical distributions are tabulated for suggested test statistics. We have small Monte-Calro studies to compare the powers of the test statistics suggested by Johansen(1988) and in this paper.

Keywords: Autoregressive process; Least squares estimator; Cointegration; Unit root test

1. INTRODUCTION

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

$$Y_t - \mu = \Phi(Y_{t-1} - \mu) + \eta_t, \ t = 1, 2, \cdots, \tag{1.1}$$

where $Y_0 = 0$ and $\{\eta_t : t = 1, 2, \cdots\}$ is a sequence of independent and identically distributed multivariate normal variates with mean 0 and variance Ω. When some eigenvalues of Φ are one, we say that the process is nonstationary. Since the number of cointegrating vectors is the same as that of the eigenvalues of Φ less than one in magnitude, multiple unit roots test is closely related to the cointegration problem. See Engle and Granger(1987) and Murray(1994) for cointegration. Phillips and Durlauf(1986) suggested several statistics for testing $H_0 : \Phi = I_k$. Fountis and Dickey(1989) studied nonstationary AR(p) process where only one eigenvalue is one and the rest are less than one in magnitude. They showed

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

that the nonstationary part(unit root part) and the stationary part of the ordinary least squares estimators can be separated in the limit. Johansen(1988) studied multiple unit roots problem using likelihood ratio test statistic. Among the many fine discussions of these issues are those provided by Hamilton(1994) and Fuller(1996). In section 2 we generalize the method suggested by Fountis and Dickey(1989) and suggest simple test statistics for multiple unit roots test. In secton 3 we tabulate the empirical distributions of test statistics developed in section 2 and compare the empirical powers of the test statistics in this paper and suggested by Johansen(1988). We have some concluding remarks in section 4.

2. DERIVATION OF TEST STATISTICS

Consider the k-dimensional multivariate AR(1) model defined in (1.1)

$$Y_t - \mu = \Phi(Y_{t-1} - \mu) + \eta_t,\ t = 1, 2, \cdots .$$

Assume that there exists a real matrix R such that

$$A = R^{-1}\Phi R = \begin{bmatrix} I_r & O' \\ O & A_{22}^* \end{bmatrix}, \tag{2.1}$$

where O is a proper zero matrix and A_{22}^* has eigenvalues less than one in magnitude. Then the transformed process by R^{-1} becomes

$$Z_t - \nu = A(Z_{t-1} - \nu) + \epsilon_t,\ t = 1, 2, \cdots , \tag{2.2}$$

where $Z_t = R^{-1}Y_t, \nu = R^{-1}\mu$ and $\epsilon_t = R^{-1}\eta_t$. Note that since A and Φ are similar matrices, the eigenvalues of A and Φ are the same. So testing that Φ has r unit roots is the same as testing A has r unit roots.

For a given data set Y_t 's , one can obtain the ordinary least squares estimator of Φ and get eigenvalues from that estimator. These eigenvalues are the same as those of the ordinary least squares estimator of A. See Fountis and Dickey (1989, p 421).

In this paper we suggest test statistics based on estimated eigenvalues of Φ which are the same as those of A for unit roots test. Since the eigenvalues do not depend on the matrix R, the test statistics based on eigenvalues do not depend on the matrix R. Using the eigenvalues for unit roots test, those models (1.1) and (2.2) can be treated as same. So in the following sections we use the transformed process by R^{-1} for brevity.

2.1. Testing for $A = I_k$

Consider the k-dimensional multivariate AR(1) process with mean zero defined by

$$Z_t = AZ_{t-1} + \epsilon_t,\ t = 1, 2, \cdots, \tag{2.3}$$

where $\epsilon_t{}'s$ are indepenent and identically distributed with mean zero and variance Σ.

Assume that $1 \geq \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_k > -1$ where α_i , $i = 1, \cdots, k$ are the eigenvalues of A and we want to test $H_0 : A = I_k$.

The ordinary least squares estimator of A is defined by $A_{ols} = \sum_{t=2}^{n} Z_t Z'_{t-1} \left[\sum_{t=2}^{n} Z_{t-1} Z'_{t-1}\right]^{-1}$. Phillips and Durlauf (1986) studied A_{ols} and obtained the limiting distribution of $n(A_{ols} - I_k)$. The result is stated in Theorem 2.1.

Theorem 2.1. *(Phillips and Durlauf (1986)) When characteristic equation for k -dimensional multivariate AR(1) process has all roots one, the limiting distribution of $n(A_{ols} - I_k)$ is as follow.*

$$n(A_{ols} - I_k) \to^L \Sigma^{1/2} \left[\int_0^1 W(t)dW(t)' + I_k\right] \left[\int_0^1 W(t)W(t)'dt\right]^{-1} \Sigma^{-1/2} \tag{2.4}$$

Here " $\to^L$ " denotes convergence in distribution and W(t) is a Wiener process.

Note that when all eigenvalues of A are one then determinant of A must be one and the reverse is true with the assumption after (2.3). Also trace of A with all roots one must be k and the reverse is true.

So we suggest new test statistics, $n(det(A_{ols}) - 1)$ and $n(trace(A_{ols}) - k)$, for testing $H_0 : A = I_k$ *.vs.* H_A : some eigenvalues are less than one.

Theorem 2.2. *The limiting distribution of $n(det(A_{ols}) - 1)$ is as follow.*

$$n(det(A_{ols}) - 1) \to^L tr(\left[\int_0^1 W(t)dW(t)' + I_k\right] \left[\int_0^1 W(t)W(t)'dt\right]^{-1}) \tag{2.5}$$

Proof: Using diagonal expansion of determinant, see Searl(1982), $det(A_{ols}) = det(A_{ols} - I_k + I_k) = 1 + tr_1(A_{ols} - I_k) + tr_2(A_{ols} - I_k) + \cdots + tr_k(A_{ols} - I_k)$. Here $tr_i(B)$ is the sum of principal minors of order i of $det(B)$. Note that $tr_1(A_{ols} - I_k) = trace(A_{ols} - I_k)$ and $tr_k(A_{ols} - I_k) = det(A_{ols} - I_k)$. Now using Theorem 2.1 we have $tr_i(A_{ols} - I_k) = O_p(n^{-i})$. So $n(det(A_{ols}) - 1) = ntrace(A_{ols} - I_k) + O_p(n^{-1})$. □

Theorem 2.2 shows that suggested two statistics share the common limiting distribution. This approach can be extended to the mean estimated case.

Consider AR(1) process with nonzero mean vector ν defined in (2.2). Let $A_{\nu,ols}$ be the ordinary least squares estimator of A for the model (2.2). Then the limiting distribution of $n(A_{\nu,ols} - I_k)$ for AR(1) model with mean vector ν is found in Phillips and Durlauf (1986). Using this result we have Theorem 2.3.

Theorem 2.3. *With the assumptions defined after (2.1), the limiting distribution of $n(det(A_{\nu,ols}) - 1)$ is as follow.*

$$n(det(A_{\nu,ols}) - 1) \to^{L}$$

$$tr\left(\left[\int_0^1 dW(t)W(t)' + I_k - W(1)\int_0^1 W(t)'dt\right]\left[\int_0^1 W(t)W(t)'dt - \int_0^1 W(t)dt\int_0^1 W(t)'dt\right]^{-1}\right) \tag{2.6}$$

The proof of this theorem is similar to that of Theorem 2.2. Again $n(det(A_{\nu,ols}) - 1)$ and $n(trace(A_{\nu,ols}) - k)$ have the same limiting distribution.

2.2. Testing unit roots

Consider the k-dimensional multivariate AR(1) model defined in (2.3) again. This time we assume that A has r unit roots and the rest are less than one in magnitude. So we want to test H_0 : A has r unit roots , $r \leq k$. Then one can use the proposed statistics, $n(det(A_{ols}) - 1)$ and $n(trace(A_{ols}) - r)$. First consider A_{ols}, least squares estimator of A.

$$\begin{aligned} A_{ols} &= \begin{bmatrix} A_{ols11} & A_{ols12} \\ A_{ols21} & A_{ols22} \end{bmatrix} = \sum_{t=2}^{n} Z_t Z_{t-1}' \left[\sum_{t=2}^{n} Z_{t-1}Z_{t-1}'\right]^{-1} \\ &= \begin{bmatrix} \sum_{t=2}^n Z_{1t}Z_{1t-1}' & \sum_{t=2}^n Z_{1t}Z_{2t-1}' \\ \sum_{t=2}^n Z_{2t}Z_{1t-1}' & \sum_{t=2}^n Z_{2t}Z_{2t-1}' \end{bmatrix} \cdot \\ &\quad \begin{bmatrix} \sum_{t=2}^n Z_{1t-1}Z_{1t-1}' & \sum_{t=2}^n Z_{1t-1}Z_{2t-1}' \\ \sum_{t=2}^n Z_{2t-1}Z_{1t-1}' & \sum_{t=2}^n Z_{2t-1}Z_{2t-1}' \end{bmatrix}^{-1} \end{aligned}$$

where Z_{1t} is a r-dimensional vector corresponding to nonstationary process.

Let $D_n = diag(n, \cdots, n, n^{1/2}, \cdots, n^{1/2})$. Then it is known that $(A_{ols} - A)D_n = \sum_{t=1}^n e_t Z_{t-1}' D_n^{-1}(D_n^{-1}\sum_{t=1}^n Z_{t-1}Z_{t-1}'D_n^{-1})^{-1} = O_p(1)$ where $e_t = Z_t -$

AZ_{t-1}. With convergence rates in Fountis and Dickey(1989) we have

$$n(A_{ols11} - I_r) = \sum_{t=1}^{n} e_{1t} Z'_{1t-1}/n \left[\sum_{t=1}^{n} Z_{1t-1} Z'_{1t-1}/n^2\right]^{-1} + o_p(1) \qquad (2.7)$$

where $e_{1t} = Z_{1t} - Z_{1t-1}$.

So existence of the stationary part does not affect the limiting distribution of estimated eigenvalues in the nonstationary part. Hence by checking how the determinant or trace of A_{ols} corresponding to the largest r eigenvalues of A_{ols11} are close to one or r, one can test for the number of unit roots. Note that Fountis and Dickey (1989) studied the case of r=1 which is a special case of these suggested statistics. This method can be applied to the mean estimated case.

Consider AR(1) process with nonzero mean vector ν defined in (2.2). Again we assume that A has r unit roots and the rest are less than one in magnitude and we want to test H_0 : A has r unit roots , $r \leq k$. Let $A_{\nu,ols}$ be the ordinary least squares estimator of A in the model (2.2) and $A_{\nu,ols11}$ be the upper left corner matrix of $A_{\nu,ols}$ corresponding to the largest r eigenvalues. Then using the similar arguments developed above, the limiting distribution of $n(det(A_{\nu,ols11}) - r)$ can be obtained by Theorem 2.3.

3. SIMULATION STUDIES

In this section we tabulate the empirical distributions of suggested test statistics. With the 5 % critical values, we compare the empirical powers of three test statistics; $n(det(\Phi_{\mu,ols}) - 1)$, $n(trace(\Phi_{\mu,ols}) - k)$ and statistic by Johansen.

3.1. Percentiles

For simplicity we do not tabulate the percentiles for zero mean case because a time series with zero mean is seldom encountered in practice. We just consider k-dimensional AR(1) process with mean vector μ and the variance matrix Ω. So the process used for simulation is $Y_t - \mu = \Phi(Y_{t-1} - \mu) + \eta_t$. To construct the percentiles, we use $Y_0 = 0, \mu = 0, \Phi = I_k$, and generate η_t's as independent standard normal random variates with variance I_k. The RANNOR function in FORTRAN is used to generate η_t's. For a given sample size n, 50, 100, 250, we generate 50,000 replications of each sample size n and compute the test statistics. As one can see, the percentiles based on trace and determinant of $\Phi_{\mu,ols}$ become closer as n increases. These results confirm Theorem 2.3. The percentiles are reported in Table 3.1 to Table 3.3.

Table 3.1: Percentiles of $n(det(\Phi_{\mu,ols}) - 1)$ and $n(trace(\Phi_{\mu,ols}) - k)$ for two unit roots

Length of data		percentiles				
		0.01	0.025	0.05	0.1	0.25
n=50	trace	-32.05	-27.35	-25.59	-22.38	-12.83
	det	-29.10	-26.03	-23.57	-20.81	-12.28
n=100	trace	-34.21	-29.37	-26.64	-23.00	-13.04
	det	-32.50	-28.70	-25.56	-22.21	-12.77
n=250	trace	-34.47	-30.99	-27.21	-23.58	-13.24
	det	-32.12	-30.30	-26.81	-23.24	-13.13

Table 3.2: Percentiles of $n(det(\Phi_{\mu,ols}) - 1)$ and $n(trace(\Phi_{\mu,ols}) - k)$ for three unit roots

Length of data		percentiles				
		0.01	0.025	0.05	0.1	0.25
n=50	trace	-47.36	-43.85	-40.02	-36.29	-24.48
	det	-37.47	-34.89	-32.65	-30.12	-21.52
n=100	trace	-51.26	-46.14	-42.57	-38.39	-25.46
	det	-45.49	-41.63	-38.35	-34.94	-23.86
n=250	trace	-53.85	-48.10	-44.04	-39.48	-25.88
	det	-51.21	-46.34	-42.22	-38.03	-25.22

Table 3.3: Percentiles of $n(det(\Phi_{\mu,ols}) - 1)$ and $n(trace(\Phi_{\mu,ols}) - k)$ for four unit roots

Length of data		percentiles				
		0.01	0.025	0.05	0.1	0.25
n=50	trace	-63.05	-58.29	-54.54	-50.48	-37.11
	det	-42.59	-40.54	-38.13	-36.72	-29.30
n=100	trace	-69.33	-63.08	-59.61	-54.83	-39.97
	det	-56.29	-52.81	-49.83	-46.44	-35.28
n=250	trace	-73.91	-68.14	-63.30	-58.09	-41.87
	det	-67.90	-62.92	-58.74	-54.23	-39.79

3.2. Empirical powers

In this section we study the empirical powers of the test statistics suggested in this paper and by Johansen(1988). Here we only consider $H_0 : A = I_2$. Critical values for $\alpha = 0.05$ level tests are given in Table 3.1 to Table 3.3. For the power of statistic suggested by Johansen, we use 18.59 for n=50 and 18.42 for n=100 as critical values. We consider two cases, (1) $\Phi = A$, $\eta_t \sim N(0, I_2)$, (2) $\Phi = A$, $\eta_t \sim N(0, \Omega)$, where $A = diag(\alpha_1, \alpha_2)$ and $\Omega = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. For each case, we generate data with sample size n=50,100 and $\alpha_1, \alpha_2 = 0.8$, 0.85, 0.9, 0.95 and 1. From the results of Table 3.4 to Table 3.7 we observe the followings.

1. The power of $n(trace(\Phi_{\mu,ols}) - 2)$ is better than that of $n(det(\Phi_{\mu,ols}) - 1)$ when α_1 and α_2 are less than one.

2. When $\alpha_1 = 1$, $\alpha_2 < 1$, the power of $n(trace(\Phi_{\mu,ols}) - 2)$ is almost the same as that of $n(det(\Phi_{\mu,ols}) - 1)$.

3. The powers of test statistics developed in this paper become lower when correlation exists.

4. When Var(η_t)= I, the power of $n(trace(\Phi_{\mu,ols}) - 2)$ is better than that of Johansen's method.

5. When Var(ϵ_t)= Ω and the difference of α_1 and α_2 is large, actually depending on the difference and the magnitude of α_i's, the power of Johansen's method is better than that of $n(trace(\Phi_{\mu,ols}) - 2)$.

Note that for simulation we use diagonal matrix for A since the effect of non-identity variance marix of errors on the multiple unit roots test is equivalent to that of non-diagonal A matrix. To see the effect of Corr(η_{1t} , η_{2t})= ρ, correlation of errors, we calculate the powers of test statistics for (α_1 , α_2) = (0.9 ,0.85), n=100 with various correlations. Result is reported in Table 3.8.

Table 3.4: Empirical power of size 0.05 with n=50, Var(η_t) $= I$

			α_1				
			0.8	0.85	0.9	0.95	1.0
α_2	0.8	trace	66.08	55.53	44.38	34.97	24.04
		det	59.45	49.50	39.61	31.88	23.26
		Jo	37.09	29.15	22.24	17.57	14.37
	0.85	trace		44.38	33.92	25.33	17.49
		det		39.61	29.72	22.79	16.69
		Jo		22.24	16.41	13.06	10.53
	0.9	trace			24.89	18.06	11.81
		det			21.93	16.15	11.07
		Jo			12.48	9.52	7.66
	0.95	trace				12.35	7.90
		det				11.15	7.59
		Jo				7.39	6.05
	1.0	trace					5.12
		det					5.01
		Jo					4.83

Table 3.5: Empirical power of size 0.05 with n=50, Var(η_t)$= \Omega$

			α_1				
			0.8	0.85	0.9	0.95	1.0
α_2	0.8	trace	66.29	54.54	41.16	27.80	16.99
		det	59.92	48.50	36.66	25.35	16.79
		Jo	37.33	30.34	27.88	30.12	35.29
	0.85	trace		43.94	32.73	21.82	11.68
		det		38.63	28.77	19.21	11.36
		Jo		21.95	17.82	18.39	21.77
	0.9	trace			25.06	16.82	8.61
		det			21.94	15.04	8.35
		Jo			12.24	10.77	12.24
	0.95	trace				12.38	6.68
		det				11.26	6.45
		Jo				7.31	6.86
	1.0	trace					4.95
		det					5.02
		Jo					5.10

Table 3.6: Empirical power of size 0.05 with n=100, Var(η_t)= I

			α_1				
			0.8	0.85	0.9	0.95	1.0
α_2	0.8	trace	99.72	98.23	93.91	83.64	67.08
		det	99.60	97.67	92.75	82.38	67.68
		Jo	94.99	86.32	73.19	57.28	43.42
	0.85	trace		93.59	82.56	64.99	44.21
		det		92.06	80.04	62.73	44.12
		Jo		72.71	54.32	37.35	26.10
	0.9	trace			64.01	43.08	24.41
		det			60.83	40.65	23.88
		Jo			35.98	21.88	14.07
	0.95	trace				24.07	11.32
		det				22.57	11.14
		Jo				11.90	7.47
	1.0	trace					4.88
		det					4.91
		Jo					4.70

Table 3.7: Empirical power of size 0.05 with n=100, Var(ϵ_t)= Ω

			α_1				
			0.8	0.85	0.9	0.95	1.0
α_2	0.8	trace	99.67	98.22	92.22	75.88	53.47
		det	99.52	97.66	90.74	73.98	53.74
		Jo	94.68	88.26	82.64	83.07	88.48
	0.85	trace		93.55	81.16	57.66	32.02
		det		91.96	78.61	55.39	31.95
		Jo		72.56	58.15	55.65	65.74
	0.9	trace			63.78	39.77	16.58
		det			60.71	37.65	16.45
		Jo			35.44	26.59	35.02
	0.95	trace				24.38	8.39
		det				22.79	8.20
		Jo				12.00	12.42
	1.0	trace					4.85
		det					4.86
		Jo					4.89

Table 3.8: Empirical power of size 0.05 with various ρ and $(\alpha_1, \alpha_2) = (0.9, 0.85)$

ρ	0.0	0.45	0.71	0.83	0.89	0.93	0.95	0.96
Trace	82.46	82.38	81.27	79.71	77.82	75.70	73.73	71.18
Det	80.00	79.94	78.65	77.08	75.21	72.96	70.90	68.20
Jo	54.36	55.84	58.35	62.94	68.73	75.54	81.80	87.88

Table 3.8 shows that powers of suggested test statistics in this paper decrease as correlation increases but the result of Johansen's method is reversed. Notice that the R matrix defined in (2.1) is closely related to the correlation of errors and so the form of R matrix affects the powers of test statistics.

4. CONCLUDING REMARKS

Many researches are devoted to finding the number of unit roots in multivariate autoregressive model. In this paper we generalize the method developed by Fountis and Dickey (1989). So when the number of unit roots in a model is one, the suggested methods become the same as that developed by Fountis and Dickey (1989). The suggested test statistics can be easily calculated even though some of roots are complex values. In simulation study we use the absolute value of roots. Comparing two test statistics, test statistic based on trace of least squares estimator is better than that based on determinant. When the difference between α_i's or the correlation of errors are not large, the test statistic based on the trace of estimator of Φ gives better power properties than that of Johansen(1988).

REFERENCES

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*, Vol. 74, 427-431.

Engle, R. F. and Granger, C. W. I (1987). "Co-Integration and Error Correction : Representation, Estimation, and Testing," *Econometrica*, Vol 55, No.2,251-276.

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.

Fuller, W. A. (1996). *Introduction to Statistical Time Series*, John Wiley and Sons, INC.

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

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

Murray, M. P. (1994). "A Drunk and Her Dog : An Illustrations of Cointegration and Error Correction," *The American Statistician*, Vol.48, No. 1, 37-39.

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

Seal, S. R. (1982). *Matrix Algebra Useful for Statistics*, John Wiley and Sons, INC.