THE INSTITUTE OF STATISTICS

Testing a multivariate process for a unit root using unconditional likelihood

by

Key-Il Shin and David A. Dickey

Institute of Statistics

Mimeo Series # ~~2410~~ 2246

North Carolina State University

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina

Testing a multivariate process for a unit root using unconditional likelihood

Key-Il Shin
North Carolina State University
Department of Statistics
Raleigh

David A. Dickey
North Carolina State University
Department of Statistics
Raleigh

1993

Abstract

Multivariate time series with unit roots are important in the study of cointegration. Most statistical methods deal with maximum likelihood estimators using a likelihood conditioned on an initial observation. We investigate the effect of using the unconditional likelihood function and find that this produces tests with better power properties.

1. Introduction

Vector time series are of interest in many areas of research. Autoregressive models, which express the current vector as a linear function of its predecessors, often suffice for observed data. Least squares estimators for the parameter matrices are easy to obtain.

As with univariate time series models, the concept of stationarity plays on important role. Stationarity requires a constant mean and variance as well as a covariance that depends only on the time separation between vectors. In univariate autoregressions, stationarity is characterized by the location of the roots of a "Characteristic polynomial" whose coefficients are the autoregressive coefficients. In vector series, autoregressive coefficient matrices are used to construct the characteristic polynomial and the roots are solutions to a determinantal equation.

Often data appears to be nonstationary, a phenomenon that suggests roots of unit modulus. Of course, since the true autoregressive matrices are not available, we can only test for unit roots using an estimated model. A vector process Y of dimension k might have r unit roots with $0<r<k$ of such a nature that a linear transformation Z=TY with T a square matrix produces Z with r nonstationary and k-r stationary components. In this case the vector Y is said to be share r common trends and have k-r cointegrating vectors, these being k-r rows of T. The study of cointegration, common trends, and the related topic of error correction models has enjoyed high visibility in recent econometric literature.

Test for unit roots, based on maximizing the likelihood conditional on the first observations, have been suggested. Computations use standard least squares techniques - univariate or multivariate. The distributions of the resulting estimates, however, are nonstandard even in the limit when unit roots are present.

Motivated by the work of Gonzales-Farias (1992) in the univariate case, we consider unconditional likelihood maximization as a vector estimation technique. Power advantages, reported by Gonzales-Farias for univariate series with estimated means, seem to hold in the vector case as well. We study the case of a single unit root, obtaining the same limit distribution as Gonzales-Farias got for the univariate case. Unlike the stationary case where least squares and exact maximum likelihood estimates converge to the same normal distribution, the estimates and test statistics converge to different distributions when unit roots are present.

2. Least squares and maximum likelihood

In order to test for unit roots, one usually assumes that the initial values are fixed. This assumption makes calculations easier especially if one wants to use a maximum likelihood estimator of the parameters in the model. Using a likelihood in which the initial values are assumed random may give better power to test for unit roots even though one needs more complicated calculations to get the estimators. In this paper we call the likelihood function with initial values random "the unconditional likelihood function" following Cox (1991) and Cox and Llatas (1991) who study the nearly nonstationary case.

Dickey (1976) and Dickey and Fuller (1979) studied univariate AR(1) models with given initial values to test for a unit root using ordinary least squares estimators. Gonzales-Farias (1992) used the unconditional likelihood function to estimate the parameters in the univariate AR(p) models.

Fountis (1983), and Fountis and Dickey (1989) investigated the multivariate AR(p) model with one unit root and others less than one. They assumed that the initial values are fixed and used the ordinary least squares estimators that arise from the conditional likelihood. They obtained the same limiting distribution for the normalized estimate of the unit root as that in Fuller (1976). Johansen (1988, 1991) derived the conditional maximum likelihood estimator of the cointegration vectors for an autoregressive process with independent Gaussian errors and initial values fixed. He derived a likelihood ratio test for the hypothesis that there are a given number of these.

In this paper we study vector processes with one unit root and the rest less than one in modulus as was investigated by Fountis (1989). Instead of using Fountis' ordinary least squares estimates, we use the estimates which maximize the stationary unconditional Gaussian likelihood function.

3. Development of likelihood

3.1 Multivariate AR(1) model : no mean case

The unconditional likelihood function for a stationary model is developed here. It is an appropriate likelihood under the alternative hypothesis of stationary, but can be maximized for any set of observed data.

Consider the k-dimensional multivariate AR(1) model

$$X_t = A\, X_{t-1} + \epsilon_t, \quad t \geq 1$$

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 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 $X_0, X_1, \ldots, X_n$,

$$\begin{aligned} \mathrm{Ln}(L) = & -(n+1)k/2 \ln(2\pi) - 1/2\ln(|V|) - n/2 \ln(|\Sigma|) - 1/2 X_0' V^{-1} X_0 \\ & -1/2 \sum_{t=1}^{n} (X_t - AX_{t-1})' \Sigma^{-1} (X_t - AX_{t-1}) \end{aligned} \qquad (1)$$

We now consider this unconditional likelihood under the null hypothesis that the data have a unit root. For the moment we assume data generated by $X_t = A^* X_{t-1} + \epsilon_t$ where $A^* = \mathrm{diag}(1,\alpha_2, \ldots, \alpha_k)$ is a diagonal matrix with $|\alpha_i|<1$, $i = 2, \ldots, k$.

Motivated by the least squares results of Fountis (1989) we anticipate the A which maximizes (1) has probability arbitrarily close to 1 of being in the set S_M of matrices

$$A \in S_M = \begin{bmatrix} 1+\epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1k} \\ \epsilon_{21} & \alpha_2+\epsilon_{22} & \cdots & \epsilon_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \epsilon_{k1} & \epsilon_{k2} & \cdots & \alpha_k+\epsilon_{kk} \end{bmatrix}$$

where $M_1/n<|\epsilon_{11}|<M_2/n$, $|\epsilon_{i1}|<M_2/n$, $i = 2, \ldots k$ and $|\epsilon_{jk}|<M_2/n^{1/2}$, $j = 1, \ldots p$, $k \neq 1$. These are the convergence orders Fountis obtained.

Note that this does not mean we restrict parameter space. Our strategy is that, since the derivative matrices are complicated, we first look in S_M to see if a zero derivative occurs there. If the MLE does not occur in this subspace, we should try again in another subspace. Fortunately we have good information on the estimation space from the work of Fountis (1989).

To get a maximum likelihood estimator, we need the derivatives of (1) with respect to A. Shin (1992) shows

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

For $A \in S_M$

$$\text{ii)}\quad [-1/2\ \partial \ln(|V|)/\partial A]\ D_n^{-1} = \begin{bmatrix} 1/\{2n(a_{11}-1)\} & \phi' \\ \phi & O \end{bmatrix} + O(n^{-1/2}) \qquad (3)$$

where ϕ and O are a proper zero vector and a zero matrix.

$$\text{iii)}\quad [\partial\{-1/2X_0V^{-1}X_0\}/\partial A]\ D_n^{-1} = O_p(n^{-1/2}) \text{ where } D_n^{-1}= \text{diag}(n, n^{1/2}, \ldots, n^{1/2}). \qquad (4)$$

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

$$\begin{bmatrix} 1/\{2n(a_{11}-1)\} & \phi' \\ \phi & O \end{bmatrix} + \Sigma^{-1}\sum_{t=1}^{n}(X_t - AX_{t-1})X_{t-1}' D_n^{-1} + O_p(n^{-1/2}) = 0.$$

This equation can be written as

$$\Sigma\begin{bmatrix} 1/\{2n(a_{11}-1)\} & \phi' \\ \phi & O \end{bmatrix} + \sum_{t=1}^{n}\{e_t - (A^* - A)X_{t-1}\}X_{t-1}' D_n^{-1} + O_p(n^{-1/2}) = 0. \quad (5)$$

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

Fountis (1989) shows that $\sum_{t=1}^{n} e_t X_{t-1}' D_n^{-1}$ has upper left element converging to $\xi\sigma_1^2 = \int W(t)dW(t)\sigma_1^2$ where $\sigma_1^2 = \text{Var}(\epsilon_{1t})$ and W(t) is a Wiener process. He also shows that $D_n^{-1}\sum_{t=1}^{n} X_{t-1}X_{t-1}' D_n^{-1}$ converges to a block diagonal matrix with $\Gamma\sigma_1^2 = \int_0^1 W^2(t)dt\sigma_1^2$ in the upper left corner. Also Shin (1992) shows the MLE of Σ is consistent. Therefore the upper left corner element converges to

$$\sigma_1^2/\{2n(a_{11}-1)\} + \xi\,\sigma_1^2 - \{n(a_{11}-1)\}\Gamma\sigma_1^2 = 0. \quad (6)$$

The rest of the likelihood derivation involves equations studied by Fountis (1983). Therefore the set S_M contains the MLE with arbitrary high probability as the sample size increase. Now let $\hat{\rho}_m$ be the largest eigenvalue of the estimated matrix A_M. Then we have $n(a_{11}-1) = n(\hat{\rho}_m - 1) + O_p(n^{-1/2})$ from the work of Fountis (1989). Hence we have

$$1/\{2n(\hat{\rho}_m - 1)\} + \xi - \{n(\hat{\rho}_m - 1)\}\,\Gamma = 0. \quad (7)$$

3.2 Multivariate AR(p) model : no mean case

In this section we extend our AR(1) results to the general AR(p) model and we show in general that the assumption of a diagonal A matrix is not really needed. The key step is casting the previous results in the so called statespace representation.

Consider the k-dimensional AR(p) model

$$X_t = A_1 X_{t-1} + A_2 X_{t-2} + \ldots + A_p X_{t-p} + \epsilon_t, \quad t \geq p+1 \tag{8}$$

where ϵ_t's are i.i.d $N(0,\Sigma)$. For this model we assume initial conditions

$$X_t = A_{t-1,1} X_{t-1} + A_{t-1,2} X_{t-2} + \ldots + A_{t-1,t-1} X_1 + \epsilon_t, \quad 1 \leq t < p+1 \tag{9}$$

where e_t's are independent. Let $E(X_t X_{t+h}') = \Gamma(h)$, $\Gamma(-h) = \Gamma(h)'$ for $1 \leq t$, $t+h \leq p$. Then A_{ti}'s are defined by $\Gamma(-1)=A_{11}\Gamma(0)$, $\Gamma(-2)=A_{21}\Gamma(-1)+A_{22}\Gamma(0)$, $\ldots$, $\Gamma(-p+1)=A_{p-1,1}\Gamma(-p+2)+A_{p-1,2}\Gamma(-p+3)+\ldots+A_{p-1,p-1}\Gamma(0)$ and A_{ti} is a 0 matrix if $i>t$. These define variances of ϵ_t's for $t \leq p$. Denote $Var(\epsilon_i) = V_i$ for $i \leq p$.

With this setup the unconditional likelihood function for a k-dimensional stationary multivariate process X_t is L where, for data $X_0, X_1, \ldots X_n$,

$$
\begin{aligned}
Ln(L) &= -nk/2 \ln(2\pi) - 1/2 \sum_{t=1}^{p} \ln(|V_t|) - (n-p)/2 \ln(|\Sigma|) \\
&- 1/2 \sum_{t=1}^{p} (X_t - A_{t-1,1}X_{t-1} - A_{t-1,2}X_{t-2} - \ldots - A_{t-1,t-1}X_1) V_t^{-1} (X_t - A_{t-1,1}X_{t-1} - A_{t-1,2}X_{t-2} - \ldots - A_{t-1,t-1}X_1)' \\
&- 1/2 \sum_{t=p+1}^{n} (X_t - A_1 X_{t-1} - A_2 X_{t-2} - \ldots - A_p X_{t-p}) \Sigma^{-1} (X_t - A_1 X_{t-1} - A_2 X_{t-2} - \ldots - A_p X_{t-p})'
\end{aligned}
\tag{10}
$$

Using the statespace representation we can write the model like this.

Let $Y_t = (X_t, X_{t-1}, \ldots, X_{t-p+1})'$, $\eta_t = (\epsilon_t, 0, \ldots 0)'$,

$$B = \begin{bmatrix} A_1 & A_2 \ldots & A_p \\ I & \phi \quad \ldots & \phi \\ \phi & I \quad \ldots & \phi \\ \vdots & \ddots & \vdots \\ \phi & \phi \ldots I & \phi \end{bmatrix}.$$

Then the new model is

$$Y_t = B\, Y_{t-1} + \eta_t, \quad t \geq p+1. \tag{11}$$

We assume data generated by

$$X_t = A_1^* X_{t-1} + A_2^* X_{t-2} + \ldots + A_p^* X_{t-p} + \epsilon_t, \quad t \geq 1 \tag{12}$$

where the characteristic equation,

$$q(m) = |m^p I - m^{p-1} A_1^* - \ldots - A_p^*| = 0 \tag{13}$$

has one root 1 and others α_i, $i = 2, \ldots, kp$ less than 1 in magnitude.

Let S be an eigenmatrix such that for the true parameter matrix B_o of B, $S^{-1}B_oS = \mathrm{diag}(1, \alpha_2, \alpha_3, \ldots, \alpha_{kp})$ where $|\alpha_i|<1$, $\alpha_i \neq \alpha_j$ and $|S|=|S^{-1}|=1$. Let $C = S^{-1}BS$. Then (11) becomes

$$Z_t = C\, Z_{t-1} + \delta_t, \quad t \geq p+1. \tag{14}$$

where $Z_t = S^{-1}Y_t$, $\delta_t = S^{-1}\eta_t$. Then the unconditional likelihood function is L where

$$\mathrm{Ln}(L) = -nk/2 \ln(2\pi) - 1/2 \ln(|V|) - (n-p)/2 \ln(|\Omega|) - 1/2\, Z_p V^{-1} Z_p'$$

$$-1/2 \sum_{t=p+1}^{n} (Z_t - C\, Z_{t-1})\, \Omega^{-1} (Z_t - C\, Z_{t-1})' \tag{15}$$

where $V=Var(Z_p)$, and $\Omega=Var(\delta_t)$.

Again we anticipate that the C which maximizes (15) has probability arbitrarily close to 1 of being in the set S_M of matrices

$$C \in S_M = \begin{bmatrix} 1+\epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1kp} \\ \epsilon_{21} & \alpha_2+\epsilon_{22} & \cdots & \epsilon_{2kp} \\ \vdots & \vdots & \vdots & \vdots \\ \epsilon_{k1} & \epsilon_{k2} & \cdots & \alpha_{kp}+\epsilon_{kp\,kp} \end{bmatrix}$$

where $M_1/n<|\epsilon_{11}|<M_2/n$, $|\epsilon_{i1}|<M_2/n$, $i = 2, \ldots kp$ and $|\epsilon_{jl}|<M_2/n^{1/2}$, $j = 1, \ldots kp$, $l \neq 1$.

Unfortunately the elements c_{ij} in the C matrix are not functionally independent so we can not directly apply the results for the AR(1) model here. Nevertheless, Shin (1992) shows that $\Sigma\ \partial Ln(L)/\partial(A_1, A_2, \ldots, A_p) = 0$ can be written as

$$\Omega \begin{bmatrix} 1/\{2n(c_{11}-1)\} & \phi' \\ \phi & O \end{bmatrix} + \sum_{t=1}^{n} \{e_t-(C^*-C)Z_{t-1}\}Z_{t-1}'\, D_n^{-1} + O_p(n^{-1/2}) = 0. \tag{16}$$

where $C^* = diag(1, \alpha_2, \alpha_3, \ldots, \alpha_{kp})$, $e_t = Z_t - C^*Z_{t-1}$. Therefore from the previous results we have the same limit distribution for the AR(p) model as in the AR(1) case, that is

$$1/\{2n(\hat{\rho}_m-1)\} + \xi - \{n(\hat{\rho}_m-1)\}\,\Gamma = 0. \tag{17}$$

4. Effect of mean

4.1 Preliminary statement

In the following sections we consider the k-dimensional AR(1) and AR(p) models with the mean estimated. The only difference from the models in section 3 is that the models now include the unknown parameter μ representing the mean vector. Hence some assumptions in the previous sections are directly used in these models.

For the AR(1) model with mean

i) $X_0 - \mu$ is normally distributed and $\mathrm{Var}(X_0-\mu) = V = \Gamma(0)$.

ii) Data are generated by $X_t - \mu = A^* (X_{t-1} - \mu) + \epsilon_t$ where $A^* = \mathrm{diag}(1, \alpha_2, \alpha_3, \ldots, \alpha_k)$, $|\alpha_i| < 1$ for $i = 2, \ldots k$.

iii) We anticipate that the MLE of A is in the set S_M defined in section 3.1 with high probability.

For the AR(p) model with mean

iv) $X_t - \mu = A_{t1}(X_{t-1} - \mu) + A_{t2}(X_{t-2} - \mu) + \ldots + A_{ti}(X_{t-i} - \mu) + \epsilon_t$, $1 \le t < p+1$ where e_t's are independent, A_{ti}'s and $\Gamma(h)$ are as defined in (9).

v) Data are generated by $X_t - \mu = A_1^* (X_{t-1} - \mu) + A_2^* (X_{t-2} - \mu) + \ldots + A_p^* (X_{t-p} - \mu) + \epsilon_t$, $t \ge 1$ where the characteristic equation has one root 1 and others α_i, $i = 2, \ldots, kp$ less than 1 in magnitude.

vi) We also anticipate that the MLE of C, which is obtained by using the statespace representation and transformed by the inverse of the true parameter matrix, is in the set S_M defined in section 3.2 with high probability.

4.2 Multivariate AR(1) model : with mean estimated

Consider the k-dimensional multivariate AR(1) model

$$X_t-\mu = A\,(X_{t-1}-\mu) + \epsilon_t, \quad t \geq 1$$

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

$$V = AVA' + \Sigma.$$

where $V=\Gamma(0)$ is the variance of X_t, $t \geq 1$. Now using the assumptions in section 4.1 the unconditional likelihood function is L where, for data $X_0, X_1, \ldots, X_n$,

$$\begin{aligned} \mathrm{Ln}(L) = & -(n+1)k/2\,\ln(2\pi) - 1/2\ln(|V|) - n/2\,\ln(|\Sigma|) - 1/2(X_0-\mu)'\,V^{-1}(X_0-\mu) \\ & -1/2\sum_{t=1}^{n}(X_t-\mu-A(X_{t-1}-\mu))'\,\Sigma^{-1}(X_t-\mu-A(X_{t-1}-\mu)). \end{aligned} \tag{18}$$

Shin (1992) computes the derivatives of (18) with respect to μ and A. Setting these equal to 0 we have

$$\Sigma\,((A-I)\,V)'^{-1}(X_0-\mu) - \sum_{t=1}^{n}(X_t-\mu-A(X_{t-1}-\mu)) = 0 \quad \text{and} \tag{19}$$

for $A \in S_M$, defined in section 3.1,

$$\Sigma\begin{bmatrix} 1/\{2n(a_{11}-1)\} & \phi' \\ \phi & O \end{bmatrix} + \begin{bmatrix} \mu_1^2/n & \phi' \\ \phi & O \end{bmatrix}$$

$$+\sum_{t=1}^{n}(X_t-\mu-A(X_{t-1}-\mu))(X_{t-1}-\mu)'\,D_n^{-1} + O_p(n^{-1/2}) = 0. \tag{20}$$

Solving these equations simultaneously we have

$$\sigma_1^2/\{2n(a_{11}-1)\} + \sum_{t=1}^{n} e_{1t}X_{1t-1}/n - n(a_{11}-1)\sum_{t=1}^{n} X_{1t-1}^2/n^2$$

$$+ n(a_{11}-1)\{\sum_{t=1}^{n} X_{1t-1}/n^2\}\{\sum_{t=1}^{n} e_{1t} + n(1-a_{11})\sum_{t=1}^{n} X_{1t-1}/n\}/\{2+n(1-a_{11})\}$$

$$- 1/n\{\sum_{t=1}^{n} e_{1t} + n(1-a_{11})\sum_{t=1}^{n} X_{1t-1}/n\}^2/\{2+n(1-a_{11})\} + O_p(n^{-1/2}) = 0. \quad (21)$$

Define
$$\begin{bmatrix} \Gamma \\ W \\ T \end{bmatrix} = \lim_{n\to\infty} \begin{bmatrix} n^{-2}\sum_{t=1}^{n} X_{1t-1}^2 \\ n^{-3/2}\sum_{t=1}^{n} X_{1t-1} \\ n^{-1/2}X_{1n} \end{bmatrix}$$

as in the Dickey and Fuller (1979) and let N be the limit of $n(a_{11}-1)$. We have

$$(-2\Gamma+2W^2)N^4 + (8\Gamma-6W^2+T^2-2WT-1)N^3 + (5-4T^2-8\Gamma+8WT)N^2 - (2T^2-8)N + 4 = 0. \quad (22)$$

where we have multiplied (21) through by $2n(a_{11}-1)$ and $\{2+n(a_{11}-1)\}$ resulting in a polynomial expression which converges uniformly on any closed interval to the expression (22). Note that $\Gamma - W^2$ is a positive random variable. Since $n(a_{11}-1) = n(\hat{\rho}_m-1) + O_p(n^{-1/2})$ we have the result.

4.3 Multivariate AR(p) model : with mean estimates

Now consider the k-dimensional AR(p) model

$$X_t-\mu = A_1(X_{t-1}-\mu) + A_2(X_{t-2}-\mu) + \ldots + A_p(X_{t-p}-\mu) + \epsilon_t, \quad t \geq p+1 \quad (23)$$

where ϵ_t's are i.i.d $N(0,\Sigma)$. Using the statespace representation and transformation through the inverse matrix of an eigenmatrix of the true parameter B_o we have a new model

$$Z_t-\gamma = C\,(Z_{t-1}-\gamma) + \delta_t\,,\ t \geq p+1. \tag{24}$$

Using the assumptions in section 4.1 the unconditional likelihood function is L where

$$\begin{aligned} Ln(L) = &-nk/2\,\ln(2\pi) -1/2\,\ln(|V|) -(n-p)/2\,\ln(|\Omega|) -1/2\,(Z_p-\gamma)V^{-1}(Z_p-\gamma)' \\ &-1/2 \sum_{t=p+1}^{n} (Z_t-\gamma- C\,(Z_{t-1}-\gamma))\,\Omega^{-1}(Z_t-\gamma- C\,(Z_{t-1}-\gamma))', \end{aligned} \tag{25}$$

$V=Var(Z_p)$, and $\Omega=Var(\delta_t)$. Again we anticipate that C is in the set S_M defined in section 3.2 with high probability. Then by similar methods developed in sections 3.2 and 4.2 we have two matrix equations

$$\Omega\,((C-I)V)'^{\,-1}(Z_0-\mu) - \sum_{t=p+1}^{n}(Z_t-\gamma -C(Z_{t-1}-\gamma)) = 0 \ \text{ and} \tag{26}$$

for $C \in S_M$,

$$\Omega\begin{bmatrix} 1/\{2n(c_{11}-1)\} & \phi' \\ \phi & O \end{bmatrix} + \begin{bmatrix} \gamma_1^2/n & \phi' \\ \phi & O \end{bmatrix}$$

$$+\sum_{t=p+1}^{n}(Z_t-\gamma -C(Z_{t-1}-\gamma))(Z_{t-1}-\mu)'\,D_n^{-1}+ O_p(n^{-1/2}) = 0. \tag{27}$$

Solving these equations simultaneously we have the same limit distribution as in the AR(1) model with mean case.

5. Comparison and Monte-Carlo results

Through the sections 3.1 to 4.3 we develop the limit distributions of $n(\hat{\rho}_m\text{-}1)$ and it turns out that all cases have the same limit distribution. When one uses the ordinary least squares estimators, the limit distributions are also same regardless of the dimensions and orders of the model. Gonzales-Farias (1992) tabulates the limit distribution of $n(\hat{\rho}_m\text{-}1)$ in the univariate model and compares the power of the MLE with that of the OLS. In the mean estimated case the power of the MLE is better than that of the OLS. We expect that this power advantage will carry over to the multivariate model. A small Monte-Carlo experiment was run to confirm that MLE power is better than that of the OLS. The study used samples of length n=50 and the model $Y_t = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} Y_{t\text{-}1} + \epsilon_t,\ \epsilon_t \sim N(0, I)$. Now the powers are following.

Power for bivariate AR(1) model n=50 : with mean estimated

(In absolute value)

α_1	0.5	0.8	0.9	0.8	0.9	0.9
α_2	0.5	0.5	0.5	0.8	0.8	0.9
MLE	98.74	52.44	21.46	41.34	20.60	17.86
OLS	97.44	45.82	18.36	35.88	17.82	14.46

The powers of both statistics decrease as the second root α_2 increases. For all parameter combinations the MLE is superior to the OLS and for some parameter combinations, the MLE superiority is pretty good. Compared to the powers of the univariate case which Gonzales-Farias (1992) tabulated, the powers of both statistics are lower. This makes sense since additional nuisance parameters are estimated here.

6. References

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

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

Dickey, D. A. (1976). Estimation and Hypothesis Testing in Nonstationary Time Series. Ph.D. Thesis. Iowa State University, Ames.

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.

Fountis, N. G. (1983). Testing for Unit Roots in Multivariate Autoregressive. Ph.D. Thesis. North Carolina State University, Raleigh, North Carolina.

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. (1976), Introduction to Statistical Time Series, Wiley, New York.

Gonzalez-Farias, G. (1992). A new unit root test for autoregressive time series. Ph.D. Thesis. North Carolina State University, Raleigh, North Carolina.

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

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica,* 59, 1551 - 1580.

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