Statistica Sinica **25** (2015), 529-536
doi:http://dx.doi.org/10.5705/ss.2013.299

EXPECTATION OF THE LIMITING DISTRIBUTION OF THE LSE OF A UNIT ROOT PROCESS

Shi Jin and Wenbo V. Li

University of Delaware

Abstract: In this paper we prove a conjecture raised by Tanaka on the first moment of the limiting distribution of the least squares estimator (LSE) of the unit root $I(d)$ process. The limiting random variable is a ratio of quadratic functionals of the d-fold integrated Brownian motion. Its expectation can be found by using Karhunen-Loéve expansion and a property of the eigenfunctions of its covariance kernel.

Key words and phrases: Discrete Fourier matrix, integrated Brownian motion, Karhunen-Loéve expansion, nonstationary time series, unit root.

1. Introduction

Consider the integrated process

$$(1-L)^d y_j = \epsilon_j, \quad j = 1, \ldots, T, \tag{1.1}$$

where L is the lag operator such that $Ly_j = y_{j-1}$, d is a positive integer, $y_0 = 0$ and $\{\epsilon_j\}$ is an i.i.d. sequence with mean zero and variance σ^2. The process $\{y_j\}$ is known as the unit root process, a nonstationary time series. White (1958) and Dickey and Fuller (1979) showed that, when $d = 1$, the least squares estimator (LSE) of the autoregressive coefficient of the the process converges in distribution to a functional of stochastic integrals of Brownian motion. Chan and Wei (1988) and Tanaka (1996) showed by the Functional Central Limit Theorem that, for $d > 1$, the statistic

$$\hat{\rho} = \frac{\sum_{j=1}^{T} y_{j-1} y_j}{\sum_{j=1}^{T} y_{j-1}^2} \tag{1.2}$$

converges asymptotically to a functional of stochastic integrals of integrated Brownian motion. Specifically,

$$T(\hat{\rho} - 1) \Rightarrow \begin{cases} \frac{(W^2(1)-1)/2}{\int_0^1 W^2(t)dt}, & d = 1, \\ \frac{X_{d-1}^2(1)/2}{\int_0^1 X_{d-1}^2(t)dt}, & d > 1 \end{cases} \tag{1.3}$$

where $W(t)$ is the Brownian motion and $X_d(t)$ is the d-fold integrated Brownian motion, defined recursively as

$$X_d(t) = \int_0^t X_{d-1}(s)ds, \quad t \geq 0,\ d \geq 1,$$

for all positive integer d and $X_0(t) = W(t)$. As pointed out in Tanaka (1999), $\hat{\rho}$ in (1.2) can also be interpreted as the LSE of the coefficient ρ of the model

$$y_j = \rho y_{j-1} + v_j, \quad (1-L)^{d-1} v_j = \epsilon_j, \quad j = 1, \ldots, T. \tag{1.4}$$

The limiting distribution of the LSE is of interest for statistical inference. Chan and Wei (1988) considered this for a general nonstationary AR(p) model when the characteristic roots lie on or outside the unit circle. An AR(1) model with autoregressive coefficient converging to one has been studied in Chan and Wei (1987). To obtain the percentiles of the limiting distribution in the form of a stochastic integral of the Ornstein-Uhlenbeck process, Chan (1988) derived a corresponding expansion using the Karhunen-Loéve expansion. A review of inference on nonstationary time series models was given in Chan (2006).

For the limiting distribution (1.3) when $d = 1$, the LSE is the Dickey-Fuller statistic. The analytic form of the density function of its limiting distribution is known to be difficult, and earlier researches approximate the distribution by Monte Carlo simulations and by numerical inversion of its Laplace transform. For $d = 2$ and 3, Tanaka (1996) computed the Laplace transform of the limiting distribution using the Girsanov Theorem. The corresponding transforms for $d \geq 4$ have not been found. However, Tanaka noticed in the same paper that

$$E\left[\frac{X_d^2(1)/2}{\int_0^1 X_d^2(t)dt}\right] = d + 1 \tag{1.5}$$

for $d = 1, 2$. Combined with the fact that (1.5) is true for $d = 0$, he conjectured that (1.5) holds for any non-negative integer d.

We provide a method to compute the expectation of this type of functional using the Karhunen-Loéve expansion of $X_d(t)$, and prove the conjecture.

2. Karhunen-Loéve Expansion of Integrated Brownian Motion

Suppose that $\{X(t) : 0 \leq t \leq 1\}$ is a mean zero Gaussian process with covariance function $K(t, s) = \mathbb{E}(X(t)X(s))$. The Karhunen-Loéve Theorem states that $X(t)$ can be decomposed as

$$X(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} e_k(t) \xi_k,$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ are the eigenvalues of the kernel $K(t,s)$ corresponding to the eigenfunctions $e_1(t), e_2(t), \ldots$ such that

$$\lambda e(t) = \int_0^1 K(t,s)e(s)ds \quad \text{ for } 0 \leq t \leq 1,$$

with $\{\xi_k : k \geq 1\}$ independent $N(0,1)$ random variables. The eigenfunctions $\{e_k(t) : k \geq 1\}$ are orthonormal in $L^2[0,1]$, and

$$K(t,s) = \sum_{k=1}^{\infty} \lambda_k e_k(t) e_k(s).$$

A natural consequence of the KL expansion is the distributional identity

$$\mathcal{J} = \int_0^1 X^2(t)dt \overset{\text{law}}{=} \sum_{k=1}^{\infty} \lambda_k \xi_k^2.$$

The characteristic function of $\mathcal{J}$ can be derived as

$$\mathbb{E}(\exp(iu\mathcal{J})) = \prod_{k=1}^{\infty} (1 - 2iu\lambda_k)^{-1/2}, \quad \text{ for } u \in \mathbb{R}.$$

The significance of the KL expansion is that it minimizes the total mean squared error compared to other expansions of stochastic processes. Very few Gaussian processes have their KL expansion explicitly computed. Freedman (1999) computed the KL expansion for the integrated Brownian motion $X_d(t)$ when $d = 1$, and showed that the eigenvalues are roots of the equation $1 + \cos(1/x^{1/4})\cosh(1/x^{1/4}) = 0$. Gao, Hannig, and Torcaso (2003) studied the KL expansion of $X_d(t)$ for any positive integer d. They showed the eigenfunctions of $\mathrm{Cov}(X_d(t), X_d(s))$ satisfy the following Sturm-Liouville problem on the interval [0,1]:

$$\lambda f^{(2d+2)}(t) = (-1)^{d+1} f(t) = (i)^{2d+2} f(t)$$

with boundary conditions $f^{(k)}(0) = f^{(d+1+k)}(1) = 0$ for $k = 0, 1, \ldots, d$, i the imaginary unit. Thus, the eigenfunctions are the nontrivial functions of the form

$$f(t) = \sum_{j=0}^{2d+1} c_j e^{\alpha_j t} \tag{2.1}$$

with the c_j's undetermined constants, $\alpha_j = \lambda^{-1/(2d+2)} i\omega_j$ and $\omega_j = \exp((j\pi/(d+1))i)$ satisfying the boundary conditions. The eigenvalues λ's are determined by setting the determinant of

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \omega_0 & \omega_1 & \cdots & \omega_{2d+1} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_0^d & \omega_1^d & \cdots & \omega_{2d+1}^d \\ \omega_0^{d+1}e^{\alpha_0} & \omega_1^{d+1}e^{\alpha_1} & \cdots & \omega_{2d+1}^{d+1}e^{\alpha_{2d+1}} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_0^{2d+1}e^{\alpha_0} & \omega_1^{2d+1}e^{\alpha_1} & \cdots & \omega_{2d+1}^{2d+1}e^{\alpha_{2d+1}} \end{pmatrix}$$

to be zero.

We study the behavior of the orthonormal eigenfunctions of $X_d(t)$ for any positive integer d at $t = 1$, and show that this is the key for evaluating the expectation. First, we have a lemma regarding the upper half of a discrete Fourier matrix. Its proof is given in the Supplementary Material.

Lemma 1. *Let $\omega = \exp(i\pi/(d+1))$ be the $(2d+2)$th root of unity. Let $\tilde{M}$ be the $(d+1)\times(2d+2)$ matrix with entries $\tilde{M}_{jk} = \omega^{(j-1)(k-1)}$, for $j = 1, \ldots, d+1, k = 1, \ldots, 2d+2$. Let $\boldsymbol{c} = [c_0, c_1, \ldots, c_{2d+1}]'$ satisfy $\tilde{M}\boldsymbol{c} = 0$. The vector $\boldsymbol{c}$ satisfies*

(1) $$\sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} \frac{c_j c_k}{\omega^j + \omega^k} = 0.$$

(2) $$\sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} (-1)^{j+k} \frac{c_j c_k}{\omega^j + \omega^k} = 0.$$

(3) $$\sum_{j,k=0}^{2d+1} (-1)^{j+k} c_j c_k - (2d+2) \sum_{|j-k|=d+1} (-1)^{j+k} c_j c_k = 0.$$

Theorem 1. *For the d-fold integrated Brownian motion $X_d(t)$, its orthonormal eigenfunctions satisfy $e_k^2(1) = 2d+2$ for every positive integer k.*

Proof. Let $\bar{\boldsymbol{c}} = [\bar{c}_0, \bar{c}_1, \ldots, \bar{c}_{2d+1}]'$ with $\bar{c}_i = (-1)^i e^{\alpha_i} c_i$. The boundary conditions on the eigenfunction at (2.1) show that $M\boldsymbol{c} = 0$. We split this into $\tilde{M}\boldsymbol{c} = 0$ and $\tilde{M}\bar{\boldsymbol{c}} = 0$. Since $f(t)$ is an eigenfunction, for any k we may write

$$e_k^2(1) = \frac{f^2(1)}{\int_0^1 f^2(t)dt}.$$

Thus we only have to show that

$$f^2(1) = (2d+2)\int_0^1 f^2(t)dt. \tag{2.2}$$

Plugging (2.1) into (2.2) yields

$$\sum_{j,k=0}^{2d+1} c_j c_k e^{\alpha_j+\alpha_k} = (2d+2) \sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} \frac{c_j c_k}{\alpha_j+\alpha_k}(e^{\alpha_j+\alpha_k}-1) + (2d+2) \sum_{|j-k|=d+1} c_j c_k. \tag{2.3}$$

Rearranging the terms in (2.3) and substituting $c_j e^{\alpha_j}$ with $(-1)^j \bar{c}_j$, we have

$$\begin{aligned} &(2d+2) \sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} \frac{c_j c_k}{\alpha_j+\alpha_k} - (2d+2) \sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} (-1)^{j+k} \frac{\bar{c}_j \bar{c}_k}{\alpha_j+\alpha_k} \\ &\quad + \sum_{j,k=0}^{2d+1} c_j c_k e^{\alpha_j+\alpha_k} - (2d+2) \sum_{|j-k|=d+1} c_j c_k = 0. \end{aligned} \tag{2.4}$$

As both $\boldsymbol{c}$ and $\bar{\boldsymbol{c}}$ are in the null space of $\tilde{M}$, by Lemma 1, we obtain

$$\sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} \frac{c_j c_k}{\alpha_j+\alpha_k} = \frac{1}{\lambda^{-1/(2d+2)} i} \sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} \frac{c_j c_k}{\omega_j+\omega_k} = 0,$$

$$\sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} (-1)^{j+k} \frac{\bar{c}_j \bar{c}_k}{\alpha_j+\alpha_k} = \frac{1}{\lambda^{-1/(2d+2)} i} \sum_{\substack{j,k=0 \\ |j-k|\neq d+1}}^{2d+1} (-1)^{j+k} \frac{\bar{c}_j \bar{c}_k}{\omega_j+\omega_k} = 0,$$

$$\sum_{j,k=0}^{2d+1} c_j c_k e^{\alpha_j+\alpha_k} - (2d+2) \sum_{|j-k|=d+1} c_j c_k = 0.$$

Hence, we have proven (2.4) and the result follows.

3. Expectation of the Limiting Distribution

Theorem 2. *Suppose $X(t)$ is the mean zero Gaussian process $X(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, e_k(t)\xi_k$. If $e_k^2(1) = c$, then*

$$E\left[\frac{X^2(1)/2}{\int_0^1 X^2(t)dt}\right] = \frac{c}{2}.$$

Proof. Denote the Laplace transform of $\mathcal{J} = \int_0^1 X^2(t)dt$ by $\phi(u)$. Then for $u > 0$,

$$\phi(u) = E\left[\exp(-u\mathcal{J})\right] = E\Big[\exp\Big\{-u\sum_{m=1}^{\infty} \lambda_m \xi_m^2\Big\}\Big]$$

$$= \prod_{m=1}^{\infty} E\left[\exp\left\{-u\lambda_m \xi_m^2\right\}\right]$$
$$= \prod_{m=1}^{\infty} (1+2u\lambda_m)^{-1/2}.$$

Since $a^{-1} = \int_0^\infty e^{-au} du, a > 0$, we have

$$\begin{aligned}
&E\left[\frac{X^2(1)/2}{\int_0^1 X^2(t)dt}\right] \\
&= \frac{1}{2} E\left[X^2(1)\int_0^\infty \exp\left\{-u\int_0^1 X^2(t)dt\right\} du\right] \\
&= \frac{1}{2} E\left[\left(\sum_{k=1}^{\infty}\sqrt{\lambda_k}e_k(1)\xi_k\right)^2 \int_0^\infty \exp\left\{-u\sum_{m=1}^{\infty}\lambda_m\xi_m^2\right\} du\right] \\
&= \frac{1}{2}\int_0^\infty E\left[\left(\sum_{k,j=1}^{\infty}\sqrt{\lambda_k\lambda_j}e_k(1)e_j(1)\xi_k\xi_j\right)\prod_{m=1}^{\infty}\exp\left\{-u\lambda_m\xi_m^2\right\}\right] du \\
&= \frac{1}{2}\int_0^\infty \sum_{k,j=1}^{\infty}\sqrt{\lambda_k\lambda_j}e_k(1)e_j(1)\cdot E\left[\xi_k\xi_j\prod_{m=1}^{\infty}\exp\left\{-u\lambda_m\xi_m^2\right\}\right] du \\
&= \frac{1}{2}\int_0^\infty \sum_{k=1}^{\infty}\lambda_k e_k^2(1)\cdot E\left[\xi_k^2\prod_{m=1}^{\infty}\exp\left\{-u\lambda_m\xi_m^2\right\}\right] du \\
&= \frac{1}{2}\int_0^\infty \sum_{k=1}^{\infty}\lambda_k e_k^2(1)\cdot \prod_{m\neq k}^{\infty} E\left[\exp\left\{-u\lambda_m\xi_m^2\right\}\right]\cdot E\left[\xi_k^2\exp\left\{-u\lambda_k\xi_k^2\right\}\right] du.
\end{aligned}$$

Here ξ_k^2 has the Chi-squared distribution with moment generating function

$$E\left[\exp\left\{-t\xi_m^2\right\}\right] = (1+2t)^{-1/2}. \tag{3.1}$$

Differentiating (3.1) with respect to t, we obtain $E\left[\xi_k^2\exp\{-t\xi_k^2\}\right] = (1+2t)^{-3/2}$. It follows that,

$$\begin{aligned}
E\left[\frac{X^2(1)/2}{\int_0^1 X^2(t)dt}\right] &= \frac{1}{2}\int_0^\infty \sum_{k=1}^{\infty}\lambda_k e_k^2(1)\cdot\prod_{m\neq k}^{\infty}(1+2u\lambda_m)^{-1/2}\cdot(1+2u\lambda_k)^{-3/2} du \\
&= \frac{1}{2}\int_0^\infty \prod_{m=1}^{\infty}(1+2u\lambda_m)^{-1/2}\cdot\sum_{k=1}^{\infty}\lambda_k e_k^2(1)\cdot(1+2u\lambda_k)^{-1} du \\
&= \frac{1}{2}\int_0^\infty \prod_{m=1}^{\infty}(1+2u\lambda_m)^{-1/2}\cdot\sum_{k=1}^{\infty}e_k^2(1)\cdot\frac{d}{du}\log\left[(1+2u\lambda_k)^{1/2}\right] du
\end{aligned}$$

$$
\begin{aligned}
&=\frac{1}{2}\int_0^\infty \phi(u)\cdot\sum_{k=1}^\infty e_k^2(1)\cdot\frac{d}{du}\log\left[(1+2u\lambda_k)^{1/2}\right]du\\
&=\frac{c}{2}\int_0^\infty \phi(u)\cdot\sum_{k=1}^\infty \frac{d}{du}\log\left[(1+2u\lambda_k)^{1/2}\right]du\\
&=\frac{c}{2}\int_0^\infty \phi(u)\cdot\frac{d}{du}\log\left[\prod_{k=1}^\infty(1+2u\lambda_k)^{1/2}\right]du\\
&=-\frac{c}{2}\int_0^\infty \phi(u)\cdot\frac{d}{du}\log\left[\phi(u)\right]du=-\frac{c}{2}\int_0^\infty d\phi(u)\\
&=-\frac{c}{2}\left(\lim_{u\to\infty}\phi(u)-\phi(0)\right)=\frac{c}{2}.
\end{aligned}
$$

Corollary 1. *For any positive integer d,*

$$
E\left[\frac{X_d^2(1)/2}{\int_0^1 X_d^2(t)dt}\right]=d+1. \tag{3.2}
$$

Remark. If we let $X(t)$ be the Brownian bridge $B(t)=W(t)-tW(1)$, then $B(1)=0$, and

$$
E\left[\frac{B^2(1)/2}{\int_0^1 B^2(t)dt}\right]=0,
$$

which can also be verified by its eigenfunction $f_k(t)=\sin(k\pi t)$ as above.

As an application of Corollary 1, we obtain the asymptotic expression for the bias of the least square estimator $\hat{\rho}$ defined in (1.1).

Proposition 1. *As $T\to\infty$, $E(T(\hat{\rho}-1))\to d+1$, and $E(\hat{\rho}-1)=(d+1)T+o(T^{-1})$.*

Proof. By Lemma 1 and Remark 3 in Ing, Sin, and Yu (2010), we have $E(|T(\hat{\rho}-1)|)=O(1)$. Thus, $T(\hat{\rho}-1)$ is uniformly integrable and with (3.2), give the result.

Acknowledgement

The authors would like to thank the editor and the referees for their valuable comments and suggestions. The first author would like to thank Yuk J. Leung for his advice and guidance in carrying out the work. This research was supported by National Science Foundation grants DMS-1106938.

References

Chan, N. H. (1988). The parameter inference for nearly non stationary time series. *J. Amer. Statist. Assoc.* **83**, 857-862.

Chan, N. H. (2006). Inference for time series and stochastic process. *Statist. Sinica* **16**, 683-696.

Chan, N. H. and Wei, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15**, 1050-1063.

Chan, N. H. and Wei, C. Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statist.* **16**, 367-401.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* **74**, 427-431.

Freedman, D. (1999). On the Bernstein-von Mises theorem with infinite-dimensional parameters. *Ann. Statist.* **27**, 1119-1141.

Gao, F., Hannig, J. and Torcaso, F. (2003). Integrated Brownian motions and exact L_2-small balls. *Ann. Probab.* **31**, 1320-1337.

Ing, C. K., Sin, C. Y. and Yu, S. H. (2010). Prediction errors in nonstationary autoregressions of infinite order. *Econometric Theory* **26**, 774-803.

Tanaka, K. (1996). *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory.* Wiley, New York.

Tanaka, K. (1999). The nonstationary fractional unit root. *Econom. Theory* **15**, 549-582.

White, J. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.* **29**, 1188-1197.

Department of Mathematical Sciences, University of Delaware, 15 Orchard Road, Newark, DE 19716, USA.

E-mail: jin@udel.edu

Department of Mathematical Sciences, University of Delaware, 15 Orchard Road, Newark, DE 19716, USA.

(Received October 2013; accepted February 2014)