Time Series Analysis

Nonstationary and Noninvertible Distribution Theory

Katsuto Tanaka

Chapter 1 Motivating Examples

We deal with linear time series models on which stationarity or invertibility is not imposed. Using simple examples arising from estimation and testing problems we indicate nonstandard aspects of the departure from stationarity or invertibility. In particular, asymptotic distributions of various statistics are derived by the eigenvalue approach under the normality assumption on the underlying processes. As a prelude to discussions in later chapters we also present equivalent expressions for limiting random variables based on the other approaches.

1.1. The test statistic for the parameter constancy

Let us consider the following model:

$$
(1.1)\qquad \begin{aligned} y_t &= \beta_t + \varepsilon_t\ , \\ \beta_t &= \beta_{t-1} + u_t\ , \qquad \beta_0 = 0,\ \ (t = 1, \cdots, T), \end{aligned}
$$

where

i) $\{y_t\}$ is an observable sequence, whereas $\{\beta_t\}$ is an unobservable sequence starting from $\beta_0 = 0$;

ii) $\{\varepsilon_t\}$ and $\{u_t\}$ are error sequences assumed to be independent of each other;

iii) $\{\varepsilon_t\}$ is normally independently distributed (NID) with common mean 0 and variance σ_ε^2 (> 0), which will be abbreviated as $\{\varepsilon_t\} \sim \mathrm{NID}(0, \sigma_\varepsilon^2)$; it is also assumed that $\{u_t\} \sim \mathrm{NID}(0, \sigma_u^2)$ $(\sigma_u^2 \geq 0)$.

The model (1.1) is the so-called *state space model* or the random walk plus noise model, and our concern is to test if β_t is constant, that is, $\beta_t = 0$ for all t. This is equivalent to testing

$$
H_0 : \rho = \frac{\sigma_u^2}{\sigma_\varepsilon^2} = 0 \qquad \text{against} \qquad H_1 : \rho > 0.
$$

Since $y_t = u_1 + \cdots + u_t + \varepsilon_t$, the observation vector $y = (y_1, \cdots, y_T)'$ has the distribution:

$$
(1.2)\qquad y = Cu + \varepsilon \sim \mathrm{N}(0, \sigma_\varepsilon^2(I_T + \rho CC')),
$$

where $u = (u_1, \cdots, u_T)'$, $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_T)'$, and I_T is the $T \times T$ identity matrix, while

$$
(1.3)\qquad C = \begin{pmatrix} 1 & & & & \\ \cdot & \cdot & & 0 & \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \\ 1 & \cdot & \cdot & \cdot & 1 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} 1 & & & & \\ -1 & \cdot & & 0 & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ 0 & & & -1 & 1 \end{pmatrix}.
$$

The matrix C necessarily appears from the random walk process $\beta_t = \beta_{t-1} + u_t$, and may be called the *random walk generating matrix*. Note that the (s,t)-th element of CC' is $\min(s,t)$ and the t-th largest eigenvalue λ_t of CC' is given by Rutherford (1946) (see also Problem 1.1) as

$$\lambda_t = \frac{1}{4}\left(\sin \frac{t-\frac{1}{2}}{2T+1}\pi\right)^{-2} . \tag{1.4}$$

For the present problem we consider the *Lagrange multiplier* (LM) or *score* test based on the derivative of the log-likelihood evaluated under H_0. The optimality of the LM test will be discussed in Chapter 9. Suppose, for simplicity, that σ_ε^2 is known and is assumed to be unity. Then, the log-likelihood $L(\rho)$ is given by

$$L(\rho) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\log|I_T + \rho CC'| - \frac{1}{2}y'(I_T+\rho CC')^{-1}y \tag{1.5}$$

so that

$$\begin{aligned} \left.\frac{dL(\rho)}{d\rho}\right|_{\rho=0} &= -\frac{1}{2}\mathrm{tr}(CC') + \frac{1}{2}y'CC'y \\ &= \frac{1}{2}\left(\varepsilon'CC'\varepsilon - \frac{T(T+1)}{2}\right) . \end{aligned} \tag{1.6}$$

The resulting statistic is a quadratic form in NID$(0,1)$ random variables plus a constant. Thus its exact distribution can be computed by Imhof's(1961) formula.

An asymptotic expansion for this distribution can also be obtained. In fact we have (Problem 1.2) that the c.f. $\phi_T(\theta)$ of $\varepsilon'CC'\varepsilon/(T+\frac{1}{2})^2$ can be expanded, up to $O(T^{-2})$, as

$$\begin{aligned} \phi_T(\theta) &= \prod_{t=1}^{T}\left(1 - 2i\theta\lambda_t/(T+1/2)^2\right)^{-\frac{1}{2}} \\ &\sim (\cos\sqrt{2i\theta})^{-\frac{1}{2}}\left[1 - \frac{i\theta}{8T^2}\left(1 - \frac{\sqrt{2i\theta}}{3}\tan\sqrt{2i\theta}\right)\right] . \end{aligned} \tag{1.7}$$

Note that the term of the order T^{-1} vanishes, while it can be verified that the c.f. of $\varepsilon'CC'\varepsilon/T^2$ contains the term of the order T^{-1}.

As for the limiting distribution it is known that, if regularity conditions hold, the first derivative of the log-likelihood divided by $\sqrt{T}$ tends to normality, but it is not the case with the present situation. One might argue that this is because the parameter

ρ to be tested is on the boundary of the parameter space $\rho \geq 0$ so that one of the regularity conditions does not hold. The LM statistic, however, tends to normality without this condition, in general, unlike the likelihood ratio and Wald statistics. In fact, if the testing problem is such that $H_0 : \rho = \rho_0 > 0$ against $H_1 : \rho > \rho_0$ for which the parameter space is $\rho_0 \leq \rho < \infty$, then $dL(\rho)/d\rho\,|_{\rho=\rho_0}$ tends to normality (Tanaka (1983a, b) and Problem 1.3).

Here we can make use of the knowledge of eigenvalues given in (1.4) to derive the asymptotic distribution of (1.6), whose approach may be referred to as the *eigenvalue approach*. Two other general approaches applicable to cases where eigenvalues are unknown and $\{\varepsilon_t\}$ is not necessarily normal and is dependent will be presented in later chapters. From (1.6) we have, by diagonalization,

$$\frac{2}{T^2}\left.\frac{dL(\rho)}{d\rho}\right|_{\rho=0} = \frac{1}{T^2}\sum_{t=1}^{T}\lambda_t\xi_t^2 - \frac{T+1}{2T}\,,$$

where $\{\xi_t\} \sim \mathrm{NID}(0,1)$. Denoting as V_T the first term on the right side, it can be shown (Problem 1.4) that

$$\text{(1.8)} \qquad \operatorname*{plim}_{T\to\infty}\left(V_T - \sum_{t=1}^{T}\frac{\xi_t^2}{\left(t-\frac{1}{2}\right)^2\pi^2}\right) = 0\,.$$

Thus it holds that, as $T \to \infty$,

$$\text{(1.9)} \qquad \mathcal{L}(V_T) \longrightarrow \mathcal{L}(V) = \mathcal{L}\left(\sum_{n=1}^{\infty}\frac{\xi_n^2}{\left(n-\frac{1}{2}\right)^2\pi^2}\right)\,,$$

where $\mathcal{L}(X)$ denotes the probability law of X so that $\mathcal{L}(2dL(\rho)/(T^2d\rho)\,|_{\rho=0}) \longrightarrow \mathcal{L}\left(V-\frac{1}{2}\right)$ as $T \to \infty$. Note that V is an infinite, weighted sum of independent $\chi^2(1)$ random variables with each weight being nonnegligible to the sum of weights $\sum_{n=1}^{\infty}\left(1\Big/\left(n-\frac{1}{2}\right)^2\pi^2\right) = 0.5$, although the first weight $1/(\pi/2)^2$=0.4053 is dominant.

The limiting distribution of $\varepsilon'CC'\varepsilon/T^2$ or $\varepsilon'C'C\varepsilon/T^2 = \sum_{t=1}^{T}S_t^2/T^2$ with $S_t = \sum_{j=1}^{t}\varepsilon_j$ was first dealt with by Erdös and Kac (1946) without assuming normality on $\{\varepsilon_t\}$.

Their assumption is that $\{\varepsilon_t\}$ is independent and identically distributed (i.i.d.) with common mean 0 and variance 1, which will be abbreviated as $\{\varepsilon_t\} \sim$ i.i.d. (0,1). Their work opened the way to the so-called *functional central limit theorem* or *invariance principle* to be discussed in Chapter 3.

The present testing problem was first discussed in Nyblom and Mäkeläinen (1983) and Tanaka (1983b), which was generalized by Nabeya and Tanaka (1988) and Nabeya (1989) to cases where eigenvalues cannot be obtained explicitly. A different testing problem leading to the same asymptotic result as in (1.9) was earlier discussed in Sen and Srivastava (1973) (see also Gardner (1969) and MacNeill (1974)). We shall return to this problem later, where general cases are treated together with the limiting distributions under local alternatives.

In passing we note that

$$\phi(\theta) = E(e^{i\theta V}) = \prod_{n=1}^{\infty} \left(1 - \frac{2i\theta}{\left(n - \frac{1}{2}\right)^2 \pi^2} \right)^{-\frac{1}{2}} = (\cos\sqrt{2i\theta})^{-\frac{1}{2}}$$

and it can be shown (Sen and Srivastava (1973) and Problem 1.5) that

$$P(V \le x) = 2\sqrt{2} \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} \Phi\left(-\frac{2n + \frac{1}{2}}{\sqrt{x}} \right) , \tag{1.10}$$

where Φ is the distribution function of N$(0,1)$.

As a prelude to subsequent discussions we give four equivalent expressions for V in the sense of distribution, that is,

$$\begin{aligned} \mathcal{L}(V) &= \mathcal{L}\left(\sum_{n=1}^{\infty} \frac{\xi_n^2}{\left(n - \frac{1}{2}\right)^2 \pi^2} \right) \\ &= \mathcal{L}\left(\int_0^1 w^2(t)dt \right) \\ &= \mathcal{L}\left(\int_0^1 \int_0^1 (1 - \max(s,t)) dw(s)dw(t) \right) \\ &= \mathcal{L}\left(\int_0^1 \int_0^1 \min(s,t) dw(s)dw(t) \right) , \end{aligned} \tag{1.11}$$

where $\{w(t)\}$ is the *standard Brownian motion* defined on [0,1], the definition and properties of which will be described in Chapter 2 together with the integrals involved in these expressions. The first of these expressions refers to the eigenvalue approach we have taken above and will take in this chapter, the second to the *stochastic process approach* to be discussed in Chapter 4, and the third and fourth to the *Fredholm approach* to be discussed in Chapter 5. The proof for the equivalence of the four expressions in (1.11) will be deferred until Chapter 5.

We have assumed the initial value β_0 to be zero. If β_0 is an unknown constant, the LM test statistic becomes different; so is the limiting distribution. In any case we have realized that the present model does produce nonstandard results. An intuitive reasoning may be that $\{y_t\}$ becomes an i.i.d. sequence under H_0, while it is nonstationary with an autoregressive (AR) unit root under H_1. Weakening the i.i.d. assumption on $\{\varepsilon_t\}$ to the extent of stationarity the present test may be interpreted as testing the null hypothesis of stationarity against nonstationarity. Moreover, by incorporating a deterministic trend into the model, the null hypothesis of stationarity around the trend can be tested against the AR unit root hypothesis. This idea is exploited in Kwiatkowski, Phillips, Schmidt and Shin (1992). More details will be discussed in Chapters 9 and 10.

Problems

1.1 Show that the eigenvalues of $(CC')^{-1}$ are $4\sin^2\left(\left(t-\frac{1}{2}\right)\pi\Big/(2T+1)\right)$ so that (1.4) results.

1.2 Derive the expansion in (1.7).

1.3 For the model (1.1) with $\sigma_\varepsilon^2 = 1$, derive the LM test for testing $H_0 : \rho = \rho_0\,(> 0)$ against $H_1 : \rho > \rho_0$ and show that the statistic tends to normality under H_0. More specifically, show that

$$\frac{1}{\sqrt{T}}\,\frac{dL(\rho)}{d\rho}\bigg|_{\rho=\rho_0} \longrightarrow \mathrm{N}(0,\sigma^2)\ ,$$

where $\sigma^2 = (\rho_0+2)/(4(\rho_0(\rho_0+4))^{\frac{3}{2}})\,.$

1.4 Prove that

$$\operatorname*{plim}_{T\to\infty}\left(\frac{4}{(2T+1)^2}\sum_{t=1}^{T}\lambda_t\xi_t^2-\sum_{t=1}^{T}\frac{\xi_t^2}{\left(t-\frac{1}{2}\right)^2\pi^2}\right)=0$$

so that (1.8) holds.

1.5 Establish the formula (1.10) using the fact that the inverse Laplace transform of $e^{-c\sqrt{\theta}}$ is $c\exp(-c^2/(4x))/(2\sqrt{\pi x^3})$ for $c>0$.

1.2. The test statistic for a moving average unit root

Let us next consider the first-order moving average (MA(1)) model :

$$y_t=\varepsilon_t-\alpha\varepsilon_{t-1}\quad (t=1,\cdots,T), \tag{1.12}$$

where $\varepsilon_0,\varepsilon_1,\cdots,$ are NID$(0,\sigma^2)$ random variables. The parameter α is restricted to be $|\alpha|\le 1$ because of the identifiability condition. The MA(1) model (1.12) is said to be *noninvertible* when $|\alpha|=1$. The nonstandard nature of the noninvertible MA(1) model was first recognized by Kang (1975), which was followed by theoretical work of Cryer and Ledolter (1981), Sargan and Bhargava (1983), Anderson and Takemura (1986), Tanaka and Satchell (1989) and Davis and Dunsmuir (1993).

Our purpose here is to test if the MA(1) model is noninvertible, that is, to test

$$H_0:\alpha=1\qquad\text{against}\qquad H_1:\alpha<1\,.$$

Following Tanaka (1990b) we consider an LM type test, the optimality of which will be demonstrated in Chapter 10. The log-likelihood $L(\alpha,\sigma^2)$ for $y=(y_1,\cdots,y_T)'$ is given by

$$L(\alpha,\sigma^2)=-\frac{T}{2}\log(2\pi\sigma^2)-\frac{1}{2}\log|\Omega(\alpha)|-\frac{1}{2\sigma^2}y'\Omega^{-1}(\alpha)y\,,$$

where

$$\Omega(\alpha)=\begin{pmatrix}1+\alpha^2 & -\alpha & & & 0\\ -\alpha & 1+\alpha^2 & \cdot & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & -\alpha\\ 0 & & & -\alpha & 1+\alpha^2\end{pmatrix}.$$

The maximum likelihood estimators (MLE's) of α and σ^2 under H_0 are $\hat{\alpha} = 1$ and $\hat{\sigma}^2 = y'\Omega^{-1}y/T$ with $\Omega^{-1} = \Omega^{-1}(1)$. It can be checked easily that

$$\left.\frac{d\Omega(\alpha)}{d\alpha}\right|_{\alpha=1} = \Omega\,, \qquad \left.\frac{d^2\Omega(\alpha)}{d\alpha^2}\right|_{\alpha=1} = 2I_T\,,$$

$$\frac{\partial L(\alpha,\sigma^2)}{\partial\alpha} = -\frac{1}{2}\mathrm{tr}\left(\Omega^{-1}(\alpha)\frac{d\Omega(\alpha)}{d\alpha}\right) - \frac{1}{2\sigma^2}\,y'\,\frac{d\Omega^{-1}(\alpha)}{d\alpha}y\;,$$

$$\frac{\partial^2 L(\alpha,\sigma^2)}{\partial\alpha^2} = -\frac{1}{2}\mathrm{tr}\left(\frac{d\Omega^{-1}(\alpha)}{d\alpha}\frac{d\Omega(\alpha)}{d\alpha} + \Omega^{-1}(\alpha)\frac{d^2\Omega(\alpha)}{d\alpha^2}\right) - \frac{1}{2\sigma^2}y'\frac{d^2\Omega^{-1}(\alpha)}{d\alpha^2}y\;.$$

These yield $\partial L(\alpha,\sigma^2)/\partial\alpha\,|_{\alpha=1,\,\sigma^2=\hat{\sigma}^2} = 0$ and

$$\begin{aligned}\left.\frac{\partial^2 L(\alpha,\sigma^2)}{\partial\alpha^2}\right|_{\alpha=1,\,\sigma^2=\hat{\sigma}^2} &= -\frac{1}{2}\mathrm{tr}(-I_T + 2\Omega^{-1}) - T\,\frac{y'(\Omega^{-1}-\Omega^{-2})y}{y'\Omega^{-1}y}\\ &= -\frac{T(T+5)}{6} + T\,\frac{y'\Omega^{-2}y}{y'\Omega^{-1}y}\;,\end{aligned}$$

where we have used the fact that

$$\begin{aligned}\Omega^{-1} &= [(CC')^{-1} + e_Te_T']^{-1} = CC' - \frac{1}{T+1}Cee'C'\\ &= [(C'C)^{-1} + e_1e_1']^{-1} = C'C - \frac{1}{T+1}C'ee'C\end{aligned}$$

with $e = (1,\cdots,1)'$, $e_1 = (1,0,\cdots,0)' : T\times 1$ and $e_T = (0,\cdots,0,1)' : T\times 1$.

The LM test considered here rejects H_0 if

$$\begin{aligned}S_T &= \frac{1}{T^2}\left(\left.\frac{\partial^2 L(\alpha,\sigma^2)}{\partial\alpha^2}\right|_{\alpha=1,\,\sigma^2=\hat{\sigma}^2} + \frac{T(T+5)}{6}\right)\\ &= \frac{1}{T}\,\frac{y'\Omega^{-2}y}{y'\Omega^{-1}y}\end{aligned}\tag{1.13}$$

takes large values. The limiting distribution under H_0 can be derived by the eigenvalue approach as follows. Put $\xi = \Omega^{-1/2}y/\sigma$ so that $\xi \sim \mathrm{N}(0, I_T)$ and

$$\begin{aligned}\frac{1}{T\sigma^2}y'\Omega^{-1}y &= \frac{1}{T}\xi'\xi \longrightarrow 1 \qquad \text{in probability,}\\ \frac{1}{T^2\sigma^2}y'\Omega^{-2}y &= \frac{1}{T^2}\xi'\Omega^{-1}\xi = \frac{1}{T^2}\sum_{t=1}^{T}\delta_t\xi_t^2\;,\end{aligned}$$

where δ_t is the t-th largest eigenvalue of Ω^{-1} given by Anderson (1971) (see also Problem 2.1) as

$$(1.14) \qquad \delta_t = \frac{1}{4}\left(\sin\frac{t\pi}{2(T+1)}\right)^{-2} .$$

Here the c.f. $\phi_T(\theta)$ of $\xi'\Omega^{-1}\xi/(T+1)^2$ can be expanded (Problem 2.2), up to $O(T^{-2})$, as

$$\begin{aligned}(1.15) \qquad \phi_T(\theta) &= \prod_{t=1}^{T}\left(1-2i\theta\delta_t/(T+1)^2\right)^{-\frac{1}{2}} \\ &\sim \left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-\frac{1}{2}}\left[1-\frac{i\theta}{8T^2}\left(1+\frac{\sqrt{2i\theta}}{3}\cot\sqrt{2i\theta}\right)\right] .\end{aligned}$$

Note that, as in (1.7), the term of the order T^{-1} vanishes, while it can be verified that the c.f. of $\xi'\Omega^{-1}\xi/T^2$ contains the term of the order T^{-1}.

As for the limiting distribution we first have (Nyblom and Mäkeläinen (1983) and Problem 2.3)

$$(1.16) \qquad \operatorname*{plim}_{T\to\infty}\left(\frac{1}{T^2}\sum_{t=1}^{T}\delta_t\xi_t^2-\sum_{t=1}^{T}\frac{\xi_t^2}{t^2\pi^2}\right)=0$$

so that, as $T\to\infty$ under H_0,

$$(1.17) \qquad \mathcal{L}(S_T)\longrightarrow\mathcal{L}(W)=\mathcal{L}\left(\sum_{n=1}^{\infty}\frac{\xi_n^2}{n^2\pi^2}\right) .$$

Thus we have

$$\phi(\theta)=E(e^{i\theta W})=\prod_{n=1}^{\infty}\left(1-\frac{2i\theta}{n^2\pi^2}\right)^{-\frac{1}{2}}=\left(\frac{\sin\sqrt{2i\theta}}{\sqrt{2i\theta}}\right)^{-\frac{1}{2}} .$$

The limiting distribution in (1.17) was first dealt with by Anderson and Darling (1952) in connection with goodness of fit tests. They showed that

$$P(W\le x)=\frac{1}{\pi\sqrt{x}}\sum_{n=0}^{\infty}(-1)^n\begin{pmatrix}-\frac{1}{2}\\ n\end{pmatrix}\sqrt{4n+1}\,e^{-b_n/x}K_{\frac{1}{4}}\left(\frac{b_n}{x}\right) ,$$

where $b_n=(4n+1)^2/16$ and $K_\nu(z)$ is the modified Bessel function defined by

$$K_\nu(z)=\frac{\sqrt{\pi}(z/2)^\nu}{\Gamma\left(\nu+\dfrac{1}{2}\right)}\int_0^\infty e^{-z\cosh x}\sinh^{2\nu}x\,dx ,$$

and tabulated percent points. It will be recognized in Chapter 5 that

$$
\begin{aligned}
(1.18)\qquad \mathcal{L}(W) &= \mathcal{L}\left(\sum_{n=1}^{\infty}\frac{\xi_n^2}{n^2\pi^2}\right) \\
&= \mathcal{L}\left(\int_0^1 (w(t)-tw(1))^2 dt\right) \\
&= \mathcal{L}\left(\int_0^1\int_0^1 (\min(s,t)-st)dw(s)dw(t)\right) \\
&= \mathcal{L}\left(\int_0^1\int_0^1 \left(\frac{1}{3}-\max(s,t)+\frac{s^2+t^2}{2}\right)dw(s)dw(t)\right),
\end{aligned}
$$

where $\{w(t)-tw(1)\}$ is the *Brownian bridge process* to be introduced in Chapter 2. The four expressions in (1.18) are comparable with those in (1.11).

It is important to note that the assumption on the initial value ε_0 is very crucial. If we assume $\varepsilon_0 = 0$, which may be referred to as the conditional case, so that $\{y_t\}$ is not stationary, the LM test becomes different; so is the limiting distribution of the LM statistic (Problem 2.5).

An MA unit root is often caused by overdifferencing of the original time series. From this point of view Saikkonen and Luukkonen (1993a) suggested the following model:

$$
(1.19)\qquad
\begin{aligned}
y_1 &= \mu + \varepsilon_1\,, \\
\Delta y_t &= y_t - y_{t-1} = \varepsilon_t - \alpha\varepsilon_{t-1}\,, \quad (t = 2,\cdots,T)\,,
\end{aligned}
$$

where μ is a constant and $\varepsilon_1,\cdots,\varepsilon_T \sim \mathrm{NID}(0,\sigma^2)$. Then the null hypothesis H_0 : $\alpha = 1$ implies overdifferencing. Note that, if μ is known and is assumed to be zero, $(y_1, y_2-y_1,\cdots,y_T-y_{T-1})'$ follows the conditional MA(1) model. The LM test in the present case rejects H_0 when $\left(\sum_{t=1}^{T} y_t\right)^2 \Big/ \sum_{t=1}^{T} y_t^2$ takes large values (Problem 2.6).

Suppose that the constant μ in (1.19) is unknown. Then the LM test rejects H_0 for large values of

$$
(1.20)\qquad SL_T = \frac{1}{T-1}\frac{y'MCC'My}{y'My}\,,
$$

where $M = I_T - ee'/T$ with $e = (1, \cdots, 1)'$ and C is defined in (1.3) (Saikkonen and Luukkonen (1993a) and Problem 2.7). It can be shown (Problem 2.8) that SL_T in (1.20) is rewritten as

$$SL_T = \frac{1}{T-1} \frac{(\Delta y)' \Omega_*^{-2} (\Delta y)}{(\Delta y)' \Omega_*^{-1} (\Delta y)}, \tag{1.21}$$

where $\Delta y = (y_2 - y_1, \cdots, y_T - y_{T-1})' : (T-1) \times 1$ and Ω_* is the first $(T-1) \times (T-1)$ submatrix of Ω. Comparing SL_T with S_T in (1.13) we can conclude that the LM statistic for the model (1.19) is derived completely in the same way as in (1.13) just by disregarding the first equation in (1.19) and replacing y_t by Δy_t $(t = 2, \cdots, T)$. Nonetheless the formulation (1.19) is meaningful in connection with the determination of the order of integration of $\{y_t\}$, that is, the order of the AR unit root. If $\{\Delta^{d+1} y_t\}$ is found to have an MA unit root, while $\{\Delta^d y_t\}$ is not, then the order of integration of $\{y_t\}$ is supposed to be d. The MA unit root test may be useful for that purpose.

Problems

2.1 Derive (1.14) by computing first the eigenvalues of Ω.

2.2 Derive the expansion in (1.15).

2.3 Establish (1.16).

2.4 Obtain the limiting c.f. of $\varepsilon' \Omega^{-2} \varepsilon / T^4$ for $\varepsilon \sim \mathrm{N}(0, I_T)$.

2.5 Derive the LM test for testing $H_0 : \alpha = 1$ in the model (1.12) with $\varepsilon_0 = 0$. Obtain also the asymptotic distribution of the LM statistic under H_0.

2.6 Show that the LM test for testing $H_0 : \alpha = 1$ in the model (1.19) with $\mu = 0$ rejects H_0 for large values of $\left(\sum_{t=1}^{T} y_t \right)^2 \Big/ \sum_{t=1}^{T} y_t^2$.

2.7 Derive the LM statistic in (1.20) for testing $\alpha = 1$ in the model (1.19) with μ being unknown.

2.8 Show that the statistic SL_T in (1.20) is rewritten as in (1.21).

1.3. Statistics from the one-dimensional random walk

So far we have discussed the limiting distributions of $\varepsilon' CC'\varepsilon/T^2$ in (1.9) and $\varepsilon'\Omega^{-1}\varepsilon/T^2$ in (1.17) assuming $\varepsilon \sim \mathrm{N}(0, I_T)$. Here we deal with these statistics from a different point of view. For this purpose consider the one-dimensional random walk:

$$(1.22) \qquad y_t = y_{t-1} + \varepsilon_t\,, \quad y_0 = 0\,, \qquad (t = 1, \cdots, T)\,,$$

where $\varepsilon_1\,, \varepsilon_2\,, \cdots$ are NID$(0,1)$ random variables and define

$$S_{1T} = \frac{1}{T^2}\sum_{t=1}^{T} y_t^2 = \frac{1}{T^2}\varepsilon' C' C\varepsilon\,,$$

$$S_{2T} = \frac{1}{T^2}\sum_{t=1}^{T} (y_t - \bar{y})^2 = \frac{1}{T^2}\varepsilon' C' MC\varepsilon\,,$$

where $M = I_T - ee'/T$ with $e = (1, \cdots, 1)'$. The statistic S_{1T} is the one discussed in Erdös and Kac (1946), as was described in Section 1, and it has the same distribution as that of $\varepsilon' CC'\varepsilon/T^2$. As for S_{2T} it can be shown (Problem 3.1) that

$$(1.23) \qquad \begin{aligned} \mathcal{L}(S_{2T}) &= \mathcal{L}\left(\frac{1}{T^2}\varepsilon' MCC'M\varepsilon\right) \\ &= \mathcal{L}\left(\frac{1}{T^2}\sum_{t=1}^{T-1}\gamma_t\xi_t^2\right)\,, \end{aligned}$$

where $\{\xi_t\} \sim$ NID$(0,1)$ and

$$(1.24) \qquad \gamma_t = \frac{1}{4}\left(\sin\frac{t\pi}{2T}\right)^{-2}\,, \qquad (t = 1, \cdots, T-1).$$

Therefore S_{2T} has the same limiting distribution as that of $\varepsilon'\Omega^{-1}\varepsilon/T^2$.

To summarize we have

$$(1.25) \qquad \mathcal{L}\left(\frac{1}{T^2}\sum_{t=1}^{T} y_t^2\right) \longrightarrow \mathcal{L}\left(\sum_{n=1}^{\infty}\frac{\xi_n^2}{\left(n-\frac{1}{2}\right)^2\pi^2}\right)\,,$$

$$(1.26) \qquad \mathcal{L}\left(\frac{1}{T^2}\sum_{t=1}^{T}(y_t - \bar{y})^2\right) \longrightarrow \mathcal{L}\left(\sum_{n=1}^{\infty}\frac{\xi_n^2}{n^2\pi^2}\right)\,.$$

These expressions tell us clearly that mean correction does affect the asymptotic distribution, unlike in the stationary case.

Figure 1.1 draws the densities of distributions in (1.25) for T=10, 20, 50 and ∞. The densities $f_T(x) = dP(V_T \le x)/dx$, where $V_T = \sum_{t=1}^{T} y_t^2/T^2$, were computed numerically following the inversion formula (Problem 3.2):

$$f_T(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\theta x}\phi_T(\theta)d\theta = \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}(e^{-i\theta x}\phi_T(\theta))d\theta\,, \tag{1.27}$$

where $\mathrm{Re}(z)$ is the real part of z; $\phi_T(\theta) = (\cos\sqrt{2i\theta})^{-\frac{1}{2}}$ for $T = \infty$, while, for T finite,

$$\phi_T(\theta) = \prod_{t=1}^{T}(1 - 2i\theta\lambda_t)^{-\frac{1}{2}}$$

with λ_t defined in (1.4). The numerical computation involves the square root of complex variables and how to compute this together with numerical integration as in (1.27) will be discussed in Chapter 6. It is seen from Figure 1.1 that the finite sample densities converge rapidly to the limiting density, although the former have a heavier right-hand tail.

Figure 1.1

Figure 1.2 draws the densities of distributions in (1.26) for T=10, 20 and ∞. The densities $g_T(x) = dP(W_T \le x)/dx$ with $W_T = \sum_{t=1}^{T}(y_t - \bar{y})^2/T^2$ were computed following (1.27) with $\phi_T(\theta) = (\sin\sqrt{2i\theta}/\sqrt{2i\theta})^{-\frac{1}{2}}$ for $T = \infty$ and, for T finite,

$$\phi_T(\theta) = \prod_{t=1}^{T-1}(1 - 2i\theta\gamma_t)^{-\frac{1}{2}}\,,$$

where γ_t is defined in (1.24). Note that Figure 1.2 does not contain the density for T=50 because it was found to be very close to that for $T = \infty$, while it is not much the case with Figure 1.1. This is related with theoretical findings about asymptotic expansions given in (1.7) and (1.15). The normalizer $\left(T + \frac{1}{2}\right)^2$, instead of T^2 in (1.25), could make finite sample densities closer to the limiting density, which will be exemplified shortly by presenting tables for percent points.

Figure 1.2

Table 1.1 reports percent points and means for distributions of $\sum_{t=1}^{T} y_t^2 \Big/ \left(T+\frac{1}{2}\right)^2$ for T=10, 20, 50 and ∞, where "E" stands for exact distributions, while "A" for distributions based on the asymptotic expansion given in (1.7). Table 1.2 are those for distributions of $\sum_{t=1}^{T}(y_t-\bar{y})^2/T^2$, where the asymptotic expansion "A" is based on (1.15). It is seen from these tables that the finite sample distributions are really close to the limiting distribution. Especially, percent points for T=50 are identical with those for $T=\infty$ within the deviation of 3/10000. Asymptotic expansions also give a fairly good approximation to finite sample distributions. In most cases they give a correct value up to the fourth decimal point.

Table 1.1 Table 1.2

It is an easy matter to compute moments of these distributions. Let $\kappa_{1T}^{(j)}$ be the j-th order cumulant for the distribution of $\sum_{t=1}^{T} y_t^2 \Big/ \left(T+\frac{1}{2}\right)^2$ based on the asymptotic expansion in (1.7). Define $\kappa_{2T}^{(j)}$ similarly for the distribution of $\sum_{t=1}^{T}(y_t-\bar{y})^2/T^2$ based on the asymptotic expansion in (1.15). Then we have (Problem 3.3), up to $O(T^{-2})$,

$$\begin{aligned}
&\kappa_{1T}^{(1)} \sim \frac{1}{2}-\frac{1}{8T^2}, && \kappa_{2T}^{(1)} \sim \frac{1}{6}-\frac{1}{6T^2}, \\
&\kappa_{1T}^{(2)} \sim \frac{1}{3}+\frac{1}{6T^2}, && \kappa_{2T}^{(2)} \sim \frac{1}{45}+\frac{1}{18T^2}, \\
&\kappa_{1T}^{(3)} \sim \frac{8}{15}+\frac{1}{3T^2}, && \kappa_{2T}^{(3)} \sim \frac{8}{945}+\frac{1}{45T^2}, \\
&\kappa_{1T}^{(4)} \sim \frac{136}{105}+\frac{16}{15T^2}, && \kappa_{2T}^{(4)} \sim \frac{8}{1575}+\frac{16}{945T^2}.
\end{aligned} \tag{1.28}$$

As for the limiting distributions we have (Problem 3.4)

$$\kappa_1^{(j)} = \lim_{T\to\infty}\kappa_{1T}^{(j)} = \frac{(j-1)!\,2^{3j-2}(2^{2j}-1)}{(2j)!}B_j, \tag{1.29}$$

$$\kappa_2^{(j)} = \lim_{T\to\infty}\kappa_{2T}^{(j)} = \frac{(j-1)!\,2^{3j-2}}{(2j)!}B_j, \tag{1.30}$$

where B_j's are the Bernoulli numbers : $B_1=\frac{1}{6}$, $B_2=\frac{1}{30}$, $B_3=\frac{1}{42}$, $B_4=\frac{1}{30}$, and so on. The skewness $\kappa_1^{(3)}/\left(\kappa_1^{(2)}\right)^{\frac{3}{2}}$ and kurtosis $\kappa_1^{(4)}/\left(\kappa_1^{(2)}\right)^2-3$ are 2.771 and 8.657, respectively, while $\kappa_2^{(3)}/\left(\kappa_2^{(2)}\right)^{\frac{3}{2}}$=2.556 and $\kappa_2^{(4)}/\left(\kappa_2^{(2)}\right)^2-3$=7.286.

As another example let us consider the following statistic:

$$\hat{\rho}_\delta = \sum_{t=2}^{T} y_{t-1}y_t \bigg/ \left(\sum_{t=2}^{T} y_{t-1}^2 + \delta y_T^2 \right) ,$$

where y_t follows (1.22) and δ is a given constant. The statistic $\hat{\rho}_\delta$ may be regarded as an estimator for the model

$$(1.31) \qquad y_t = \rho y_{t-1} + \varepsilon_t\,, \quad y_0 = 0, \quad (t = 1, \cdots, T),$$

where $\{\varepsilon_t\} \sim \mathrm{NID}(0,1)$. In particular $\hat{\rho}_0$ becomes both the least squares estimator (LSE) and the MLE of ρ for the model (1.31), while $\hat{\rho}_1$ is called the Yule-Walker estimator. We shall show that the asymptotic distribution of $\hat{\rho}_\delta$ does depend on the value of δ when $\rho = 1$, unlike in the stationary case, by deriving the limiting distribution of a suitably normalized quantity of $\hat{\rho}_\delta$. White (1958) first obtained the limiting c.f. associated with $T(\hat{\rho}_0 - 1)$ as $T \to \infty$ under $|\rho| \geq 1$. Here we maintain to assume that $\rho = 1$ and follow his approach.

Consider now $T(\hat{\rho}_\delta - 1) = U_T/V_T$, where

$$\begin{aligned} U_T &= \frac{1}{T}\sum_{t=2}^{T} y_{t-1}\varepsilon_t - \frac{\delta}{T} y_T^2\,, \\ V_T &= \frac{1}{T^2}\sum_{t=2}^{T} y_{t-1}^2 + \frac{\delta}{T^2} y_T^2\,. \end{aligned}$$

Then we have (Problem 3.5), for any real x,

$$(1.32) \qquad \begin{aligned} X_T &= xV_T - U_T \\ &= \varepsilon' \left[\frac{x}{T^2} C' \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & \delta \end{pmatrix} C - \frac{1-2\delta}{2T} ee' + \frac{1}{2T} I_T \right] \varepsilon\,, \end{aligned}$$

where $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_T)'$, $e = (1, \cdots, 1)' : T \times 1$ and C is the random walk generating matrix defined in (1.3). Note that $P(T(\hat{\rho}_\delta - 1) \leq x) = P(xV_T - U_T \geq 0) = P(X_T \geq 0)$. We can show (White (1958) and Problem 3.6) that the moment generating function (m.g.f.) $m_T(\theta)$ of X_T is given by

$$(1.33) \qquad m_T(\theta) = \left[r^T \left\{ \cos T\omega - \frac{r\cos\omega - d}{r\sin\omega} \sin T\omega \right\} \right]^{-\frac{1}{2}} ,$$

where

$$r = 1 - \frac{\theta}{T}, \quad d = 1 - \frac{2\delta\theta}{T} - \frac{2\delta\theta x}{T^2},$$

$$\cos\omega = 1 - \frac{\theta x}{rT^2}, \quad \sin\omega = \frac{1}{rT}\sqrt{2r\theta x - \frac{\theta^2 x^2}{T^2}}.$$

The m.g.f. $m_T(\theta)$ may be expanded (Knight and Satchell (1993) and Problem 3.7), up to $O(T^{-1})$, as

$$(1.34) \qquad m_T(\theta) \sim e^{\frac{\theta}{2}}\left[\cos A + \theta(1-2\delta)\frac{\sin A}{A}\right]^{-\frac{1}{2}} \times \left[1 + \frac{2\delta\theta^2\cos A + \theta\{(\theta-1)(\theta(1-2\delta)+2x)+4\delta x\}\dfrac{\sin A}{A}}{4T\left\{\cos A + \theta(1-2\delta)\dfrac{\sin A}{A}\right\}}\right],$$

where $A = \sqrt{2\theta x}$, and thus the limiting c.f. $\phi_\delta(\theta; x)$ of X_T is given by

$$(1.35) \qquad \phi_\delta(\theta; x) = e^{\frac{i\theta}{2}}\left[\cos\sqrt{2i\theta x} + i\theta(1-2\delta)\frac{\sin\sqrt{2i\theta x}}{\sqrt{2i\theta x}}\right]^{-\frac{1}{2}}.$$

Unlike the asymptotic expansions obtained before, it seems impossible to find a normalizer with which the term of the order T^{-1} is eliminated.

In any case we have

$$(1.36) \qquad \begin{aligned} F(x;\delta) &= \lim_{T\to\infty} P(T(\hat{\rho}_\delta - 1) \le x) \\ &= \lim_{T\to\infty} P(X_T \ge 0) \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{1}{\theta}\,\mathrm{Im}(\phi_\delta(\theta; x))d\theta, \end{aligned}$$

where $\mathrm{Im}(z)$ is the imaginary part of z. The limiting probability density $f(x;\delta)$ of $T(\hat{\rho}_\delta - 1)$ is computed as $f(x;\delta) = \partial F(x;\delta)/\partial x$. The following equivalent expressions will be obtained in later chapters for the weak convergence of X_T in (1.32):

$$(1.37) \qquad \begin{aligned} \mathcal{L}(X_T) &\longrightarrow \mathcal{L}\left(\sum_{n=1}^\infty \frac{\xi_n^2}{\lambda_n} + \frac{1}{2}\right) \\ &= \mathcal{L}\left(x\int_0^1 w^2(t)dt - (1-2\delta)\int_0^1 w(t)dw(t) + \delta\right) \\ &= \mathcal{L}\left(x\int_0^1\int_0^1 (1-\max(s,t))dw(s)dw(t)\right. \\ &\qquad \left. - \frac{1-\delta}{2}\int_0^1\int_0^1 dw(s)dw(t) + \frac{1}{2}\right), \end{aligned}$$

where $\{\lambda_n\}$ is a sequence of solutions to

$$\cos\sqrt{\lambda x} + \frac{\lambda(1-2\delta)}{2}\frac{\sin\sqrt{\lambda x}}{\sqrt{\lambda x}} = 0\,,$$

while the integral $\int_0^1 w(t)dw(t)$ is called the Ito integral, whose definition will be given in Chapter 2. Alternatively we shall have

$$\begin{aligned}
(1.38)\qquad \mathcal{L}(T(\hat\rho_\delta - 1)) &= \mathcal{L}(U_T/V_T) \\
&\longrightarrow \mathcal{L}\left(\frac{(1-2\delta)\int_0^1 w(t)dw(t) - \delta}{\int_0^1 w^2(t)dt}\right) \\
&= \mathcal{L}\left(\frac{\frac{1-2\delta}{2}\int_0^1\int_0^1 dw(s)dw(t) - \frac{1}{2}}{\int_0^1\int_0^1 (1-\max(s,t))dw(s)dw(t)}\right).
\end{aligned}$$

We note in passing that the first expression in (1.37) cannot be converted explicitly into the ratio form as in (1.38).

Moments of the limiting distribution of $T(\hat\rho_\delta - 1)$ can be derived following Evans and Savin (1981b). Suppose that $\mathcal{L}(T(\hat\rho_\delta - 1)) \longrightarrow \mathcal{L}(U/V)$ with $P(V>0)=1$ and put $\psi(\theta_1,\theta_2) = E[\exp(\theta_1 U + \theta_2 V)]$. Then the k-th order raw moment $\mu_\delta(k)$ of $F(x;\delta)$ in (1.36) is given by

$$\begin{aligned}
(1.39)\qquad \mu_\delta(k) &= \frac{1}{(k-1)!}\int_0^\infty \theta_2^{k-1} E(U^k e^{-\theta_2 V})d\theta_2 \\
&= \frac{1}{(k-1)!}\int_0^\infty \theta_2^{k-1} \left.\frac{\partial^k \psi(\theta_1,-\theta_2)}{\partial\theta_1^k}\right|_{\theta_1=0} d\theta_2\,,
\end{aligned}$$

where

$$\begin{aligned}
\psi(\theta_1,-\theta_2) &= E\left[\exp\left\{(-\theta_1)\left(\frac{\theta_2}{\theta_1}V - U\right)\right\}\right] \\
&= \phi_\delta(i\theta_1;\theta_2/\theta_1) \\
&= e^{-\frac{\theta_1}{2}}\left[\cosh\sqrt{2\theta_2} - \theta_1(1-2\delta)\frac{\sinh\sqrt{2\theta_2}}{\sqrt{2\theta_2}}\right]^{-\frac{1}{2}}.
\end{aligned}$$

Figure 1.3 draws the limiting probability densities $f(x;\delta)$ of $T(\hat\rho_\delta - 1)$ for $\delta = 0, 0.5$ and 1. It is seen that $f(x;1)$ is located to the left of $f(x;0)$, as is expected from the definition of $T(\hat{\rho}_\delta - 1)$, though these are not congruent. Table 1.3 reports percent points, means and standard deviations (SD's) of the limiting distributions of $T(\hat{\rho}_\delta - 1)$ for various values of δ. The limiting distribution of $T(\hat{\rho}_0 - 1)$ was earlier tabulated in Fuller (1976) by simulations, while tables based on numerical integration were provided by Evans and Savin (1981a), Bobkoski (1983), Perron (1989) and Nabeya and Tanaka (1990a). A closer look at the values of means and SD's in Table 1.3 leads us to conjecture the following. Let the limit in distribution of $T(\hat{\rho}_\delta - 1)$ be U_δ/V, where

$$\begin{aligned} U_\delta &= (1-2\delta)\int_0^1 w(t)dw(t) - \delta = U_{1-\delta} + 2(1-2\delta)\left(U_0 + \frac{1}{2}\right), \\ V &= \int_0^1 w^2(t)dt\,. \end{aligned}$$

Then we have

$$\begin{aligned} (1.40) \qquad E\left(\frac{U_0 + \frac{1}{2}}{V}\right) &= E\left(\frac{\int_0^1 w(t)dw(t) + \frac{1}{2}}{\int_0^1 w^2(t)dt}\right) \\ &= E\left(\frac{1}{2(1-2\delta)}\,\frac{U_\delta - U_{1-\delta}}{V}\right) = 1\,, \\ (1.41) \qquad \mathrm{Var}\left(\frac{U_\delta}{V}\right) &= \mathrm{Var}\left(\frac{U_{1-\delta}}{V}\right). \end{aligned}$$

The relation in (1.40) can be proved (Problem 3.8) by showing that the mean $E((U_0 + \frac{1}{2})/V)$ of the limiting distribution of

$$(1.42) \qquad T(\hat{\rho}_0 - 1) + \frac{\frac{1}{2}}{\frac{1}{T^2}\sum_{t=2}^{T} y_{t-1}^2} = \frac{\frac{1}{T}\sum_{t=2}^{T} y_{t-1}\varepsilon_t + \frac{1}{2}}{\frac{1}{T^2}\sum_{t=2}^{T} y_{t-1}^2}$$

is equal to 1. On the other hand the relation in (1.41) can be proved (Problem 3.9) by showing that

$$(1.43) \qquad \mu_\delta(2) - (\mu_\delta(1))^2 = \mu_{1-\delta}(2) - (\mu_{1-\delta}(1))^2\,.$$

Figure 1.3 Table 1.3

The estimator $\hat{\rho}_\delta$ may be used to test the unit root hypothesis $H_0 : \rho = 1$ against $H_1 : \rho < 1$ for the model (1.31) with $|\rho| \leq 1$. The limiting local powers will be computed in Chapter 9 by numerical integration for various values of δ and it will be found that the test based on $\delta = 0$ is the best of all the tests based on $\hat{\rho}_\delta$.

Problems

3.1 Prove the distributional equivalence in (1.23).

3.2 Derive the last equality in (1.27) from the second.

3.3 Compute cumulants in (1.28).

3.4 Obtain the expressions for cumulants in (1.29) and (1.30).

3.5 Obtain the last expression in (1.32).

3.6 Show that the m.g.f. $m_T(\theta)$ of X_T in (1.32) is given by (1.33).

3.7 Derive the asymptotic expansion of $m_T(\theta)$ given in (1.34).

3.8 Prove $E\left(\left(U_0 + \frac{1}{2}\right)/V\right) = 1$ in (1.40) by showing that the mean of the limiting distribution of (1.42) is equal to 1.

3.9 Prove (1.41) by showing that (1.43) holds.

1.4. Statistics from the two-dimensional random walk

As a sequel to the previous section we consider the two-dimensional random walk:

$$(1.44) \qquad y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \quad (t = 1, \cdots, T),$$

where $y_0 = 0$ and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \sim \mathrm{NID}(0, I_2)$. Under this last assumption $\{y_{1t}\}$ and $\{y_{2t}\}$ are independent of each other so that $\mathrm{Cov}(y_{1s}, y_{2t}) = 0$ for any s and t.

The nonstandard nature of statistics arising from the model in (1.44) can be best seen from the following example. Consider

$$(1.45) \qquad S_T = \frac{1}{T^2} \sum_{t=1}^{T} y_{1t} y_{2t} = \frac{1}{T^2} \underset{\sim}{\varepsilon}_1' C' C \underset{\sim}{\varepsilon}_2$$

$$= \frac{1}{2T^2} \underset{\sim}{\varepsilon}' \begin{pmatrix} 0 & C'C \\ C'C & 0 \end{pmatrix} \underset{\sim}{\varepsilon},$$

where $\underset{\sim}{\varepsilon}_j = (\varepsilon_{j1}, \cdots, \varepsilon_{jT})'$ $(j = 1, 2)$, $\underset{\sim}{\varepsilon} = (\underset{\sim}{\varepsilon}_1', \underset{\sim}{\varepsilon}_2')'$ and C is defined in (1.3). Then the c.f. $\phi_T(\theta)$ of S_T is given (Problem 4.1) by

$$\phi_T(\theta) = \prod_{t=1}^{T} \left(1 + \frac{\theta^2}{T^4}\lambda_t^2\right)^{-\frac{1}{2}} = \prod_{t=1}^{T} \left[\left(1 - \frac{2i\theta}{2T^2}\lambda_t\right)\left(1 + \frac{2i\theta}{2T^2}\lambda_t\right)\right]^{-\frac{1}{2}}, \tag{1.46}$$

where λ_t is the t-th largest eigenvalue of $C'C$ or CC' given in (1.4). It is noted that the distribution of S_T is symmetric about the origin since $\phi_T(\theta)$ is real. From the expression in (1.46) we have

$$\mathcal{L}(S_T) = \mathcal{L}\left(\frac{1}{2T^2}\sum_{t=1}^{T} \lambda_t(\xi_{1t}^2 - \xi_{2t}^2)\right),$$

where $\xi_t = (\xi_{1t}, \xi_{2t})' \sim \mathrm{NID}(0, I_2)$ and thus

$$\mathcal{L}(S_T) \longrightarrow \mathcal{L}(S) = \mathcal{L}\left(\frac{1}{2}\sum_{n=1}^{\infty} \frac{\xi_{1n}^2 - \xi_{2n}^2}{\left(n - \frac{1}{2}\right)^2 \pi^2}\right),$$

$$\begin{aligned} \phi(\theta) &= E(e^{i\theta S}) = \prod_{n=1}^{\infty}\left(1 + \frac{\theta^2}{\left(n - \frac{1}{2}\right)^4 \pi^4}\right)^{-\frac{1}{2}} \\ &= (\cos\sqrt{i\theta})^{-\frac{1}{2}}(\cosh\sqrt{i\theta})^{-\frac{1}{2}}. \end{aligned}$$

Therefore $\sum_{t=1}^{T} y_{1t}y_{2t}/T^2$ has a nondegenerate limiting distribution even if $\{y_{1t}\}$ and $\{y_{2t}\}$ are independent of each other with $E(y_{1t}) = E(y_{2t}) = 0$. Note that the limiting distribution is also symmetric about the origin. Three equivalent expressions for the limiting random variable S are

$$\begin{aligned} \mathcal{L}(S) &= \mathcal{L}\left(\frac{1}{2}\sum_{n=1}^{\infty} \frac{\xi_{1n}^2 - \xi_{2n}^2}{\left(n - \frac{1}{2}\right)^2 \pi^2}\right) \\ &= \mathcal{L}\left(\int_0^1 w'(t)Hw(t)dt\right) \\ &= \mathcal{L}\left(\int_0^1\int_0^1 (1 - \max(s,t))dw'(s)Hdw(t)\right), \end{aligned} \tag{1.47}$$

where $w(t) = (w_1(t), w_2(t))'$ is the two-dimensional standard Brownian motion to be introduced in Chapter 2, while

$$H = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} .$$

As the next example consider

$$(1.48) \qquad U_T = \frac{1}{T}\sum_{t=1}^{T} y_{1t}\varepsilon_{2t} = \frac{1}{2T}\underset{\sim}{\varepsilon}' \begin{pmatrix} 0 & C' \\ C & 0 \end{pmatrix} \underset{\sim}{\varepsilon},$$

which has the c.f.

$$\phi_T(\theta) = \left| I_T + \frac{\theta^2}{T^2}C'C \right|^{-\frac{1}{2}} = \prod_{t=1}^{T}\left(1 + \frac{\theta^2}{T^2}\lambda_t\right)^{-\frac{1}{2}} .$$

As $T \to \infty$ it holds that

$$(1.49) \qquad \begin{aligned} \mathcal{L}(U_T) \longrightarrow \mathcal{L}(U) &= \mathcal{L}\left(\frac{1}{2}\sum_{n=1}^{\infty}\frac{\xi_{1n}^2 - \xi_{2n}^2}{\left(n - \frac{1}{2}\right)\pi}\right) \\ &= \mathcal{L}\left(\int_0^1 w_1(t)dw_2(t)\right) , \end{aligned}$$

where the integral is the Ito integral to be introduced in Chapter 2. In the present case we cannot express the limit in distribution in (1.49) using a double integral with a continuous integrand as in the last expression in (1.47). The c.f. $\phi(\theta)$ of U in (1.49) is $\phi(\theta) = (\cosh\theta)^{-\frac{1}{2}}$. If we consider, instead of U_T in (1.48),

$$(1.50) \qquad V_T = \frac{1}{T}\sum_{t=1}^{T} y_{1t}\varepsilon_{1t} ,$$

it holds (Problem 4.2) that

$$(1.51) \qquad \begin{aligned} \mathcal{L}(V_T) &\longrightarrow \mathcal{L}\left(\frac{1}{2}(\xi^2 + 1)\right) \\ &= \mathcal{L}\left(\int_0^1 w_1(t)dw_1(t) + 1\right) \\ &= \mathcal{L}\left(\frac{1}{2}\int_0^1\int_0^1 dw_1(s)dw_1(t) + \frac{1}{2}\right) , \end{aligned}$$

where $\xi \sim \mathrm{N}(0,1)$ and the single integral is again the Ito integral. The double integral expression is possible in the present case, although an additive constant term emerges.

As the third example let us consider

$$(1.52) \qquad W_{1T} = \frac{1}{2T}\sum_{t=1}^{T}(y_{1t}\varepsilon_{2t} + y_{2t}\varepsilon_{1t}) = \frac{1}{4T}\underset{\sim}{\varepsilon}'\begin{pmatrix} 0 & C'+C \\ C'+C & 0 \end{pmatrix}\underset{\sim}{\varepsilon},$$

$$(1.53) \qquad W_{2T} = \frac{1}{2T}\sum_{t=1}^{T}(y_{1t}\varepsilon_{2t} - y_{2t}\varepsilon_{1t}) = \frac{1}{4T}\underset{\sim}{\varepsilon}'\begin{pmatrix} 0 & C'-C \\ C-C' & 0 \end{pmatrix}\underset{\sim}{\varepsilon},$$

which are mixed versions of U_T in (1.48). The c.f.'s ϕ_{jT}'s of W_{jT}'s $(j = 1, 2)$ are given (Problem 4.3) by

$$(1.54) \qquad \phi_{1T}(\theta) = \left[\left(1 + \frac{\theta^2}{4}\left(1 + \frac{1}{T}\right)^2\right)\left(1 + \frac{\theta^2}{4T^2}\right)^{T-1}\right]^{-\frac{1}{2}},$$

$$(1.55) \qquad \phi_{2T}(\theta) = \left[\frac{1}{2}\left\{\left(1 + \frac{\theta}{2T}\right)^T + \left(1 - \frac{\theta}{2T}\right)^T\right\}\right]^{-1}.$$

Therefore the distributions of W_{1T} and W_{2T} are symmetric about the origin. It evidently holds that

$$(1.56) \qquad \lim_{T\to\infty}\phi_{1T}(\theta) = \phi_1(\theta) = \left(1 + \frac{\theta^2}{4}\right)^{-\frac{1}{2}},$$

$$(1.57) \qquad \lim_{T\to\infty}\phi_{2T}(\theta) = \phi_2(\theta) = \left(\cosh\frac{\theta}{2}\right)^{-1} = \prod_{n=1}^{\infty}\left(1 + \frac{\theta^2}{((2n-1)\pi)^2}\right)^{-1}.$$

We shall have the following equivalent expressions :

$$(1.58) \qquad \begin{aligned} \mathcal{L}(W_{1T}) &\longrightarrow \mathcal{L}\left(\frac{1}{4}(\xi_1^2 - \xi_2^2)\right) \\ &= \mathcal{L}\left(\frac{1}{2}\int_0^1 w'(t)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}dw(t)\right) \\ &= \mathcal{L}\left(\frac{1}{4}\int_0^1\int_0^1 dw'(s)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}dw(t)\right), \end{aligned}$$

$$(1.59) \qquad \begin{aligned} \mathcal{L}(W_{2T}) &\longrightarrow \mathcal{L}\left(\frac{1}{4}\sum_{n=1}^{\infty}\frac{\xi_{1n}^2 + \xi_{2n}^2 - \xi_{3n}^2 - \xi_{4n}^2}{\left(n - \frac{1}{2}\right)\pi}\right) \\ &= \mathcal{L}\left(\frac{1}{2}\int_0^1 w'(t)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}dw(t)\right), \end{aligned}$$

where $(\xi_1, \xi_2)' \sim \mathrm{N}(0, I_2)$, $(\xi_{1n}, \xi_{2n}, \xi_{3n}, \xi_{4n})' \sim \mathrm{NID}(0, I_4)$ and $w(t)$ is the two-dimensional standard Brownian motion.

The limiting distributions of W_{1T} and W_{2T} are also symmetric about the origin. The former can be interpreted from the first expression in (1.58) as the distribution of the difference of two independent $\chi^2(1)/4$ random variables, while the latter is known as the distribution of Lévy's stochastic area (Hida (1980)). In the latter case the double integral expression is not possible, unlike in (1.58). This is closely related with the fact that the matrix appearing in (1.59) is not symmetric. Detailed discussions will be given in Chapter 3. Comparing (1.59) with (1.49) the following relation is seen to hold:

$$\mathcal{L}\left(\int_0^1 (w_1(t)dw_2(t) - w_2(t)dw_1(t))\right) = \mathcal{L}\left(\int_0^1 (w_1(t)dw_2(t) + w_3(t)dw_4(t))\right),$$

where $(w_1(t), w_2(t), w_3(t), w_4(t))'$ is the four-dimensional standard Brownian motion.

It is easy to obtain cumulants $\kappa_1^{(j)}$ and $\kappa_2^{(j)}$ for the limiting distributions in (1.58) and (1.59), respectively. We have (Problem 4.4)

$$\kappa_1^{(j)} = \begin{cases} 0 & j : \text{odd} \\ \dfrac{(2l)!}{l\, 2^{2l+1}} & j = 2l\,, \end{cases} \tag{1.60}$$

$$\kappa_2^{(j)} = \begin{cases} 0 & j : \text{odd} \\ \dfrac{2^{2l} - 1}{2l} B_l & j = 2l\,, \end{cases} \tag{1.61}$$

where B_l is the Bernoulli number.

Figure 1.4 draws the limiting probability densities $f_1(x)$ and $f_2(x)$ of $2 \times W_{1T}$ in (1.58) and $2 \times W_{2T}$ in (1.59), respectively, together with the density of $\mathrm{N}(0,1)$. The three distributions have means 0 and variances 1. We computed $f_1(x)$ and $f_2(x)$ following

$$f_1(x) = \frac{1}{\pi}\int_0^\infty \frac{\cos\theta x}{\sqrt{1+\theta^2}}\, d\theta\,, \tag{1.62}$$

$$f_2(x) = \frac{1}{\pi}\int_0^\infty \frac{\cos\theta x}{\cosh\theta}\, d\theta = \frac{1}{2}\frac{1}{\cosh(\pi x/2)}\,. \tag{1.63}$$

From the computational point of view it is quite easy to deal with (1.63), but we have difficulty in computing (1.62) since the integrand is oscillating and $1/\sqrt{1+\theta^2}$ approaches 0 rather slowly. Chapter 6 will suggest a method for overcoming this difficulty, from which Figure 1.4 has been produced. Percent points for the three distributions are tabulated in Table 1.4.

Figure 1.4 Table 1.4

Finally let us consider the statistics:

$$\hat{\beta}_1 = \sum_{t=1}^{T} y_{1t}\, y_{2t} \Big/ \sum_{t=1}^{T} y_{1t}^2 \,, \tag{1.64}$$

$$\hat{\beta}_2 = \sum_{t=1}^{T} (y_{1t} - \bar{y}_1)(y_{2t} - \bar{y}_2) \Big/ \sum_{t=1}^{T} (y_{1t} - \bar{y}_1)^2 \,, \tag{1.65}$$

which may be interpreted as the LSE's derived from the regression relations $y_{2t} = \hat{\beta}_1 y_{1t} + \hat{v}_{1t}$ and $y_{2t} = \hat{\alpha} + \hat{\beta}_2 y_{1t} + \hat{v}_{2t}$, respectively. These are called *spurious regressions* following Granger and Newbold (1974) because the regressor $\{y_{1t}\}$ is independent of the regressand $\{y_{2t}\}$. Nonetheless the LSE's $\hat{\beta}_1$ and $\hat{\beta}_2$ have nondegenerate limiting distributions, which we show below.

As for $\hat{\beta}_1$ put $P(\hat{\beta}_1 \le x) = P(X_{1T} \ge 0)$, where

$$\begin{aligned} X_{1T} &= \frac{x}{T^2}\sum_{t=1}^{T} y_{1t}^2 - \frac{1}{T^2}\sum_{t=1}^{T} y_{1t}\, y_{2t} \\ &= \underset{\sim}{\varepsilon}' \begin{pmatrix} \frac{x}{T^2} C'C & -\frac{1}{2T^2} C'C \\ -\frac{1}{2T^2} C'C & 0 \end{pmatrix} \underset{\sim}{\varepsilon} \,, \end{aligned} \tag{1.66}$$

with $\underset{\sim}{\varepsilon} = (\varepsilon_{11}, \cdots, \varepsilon_{1T}, \varepsilon_{21}, \cdots, \varepsilon_{2T})'$. The c.f. $\phi_{1T}(\theta; x)$ of X_{1T} in (1.66) is given (Problem 4.5) by

$$\phi_{1T}(\theta; x) = \prod_{t=1}^{T} \left[\left(1 - \frac{2i\theta a(x)}{T^2}\lambda_t\right) \left(1 - \frac{2i\theta b(x)}{T^2}\lambda_t\right) \right]^{-\frac{1}{2}} , \tag{1.67}$$

where $a(x) = (x + \sqrt{x^2+1})/2$, $b(x) = (x - \sqrt{x^2+1})/2$ and λ_t is the t-th largest eigenvalue of CC' or $C'C$ given in (1.4). It is noted from the expression for $\phi_{1T}(\theta; x)$ that

$$\mathcal{L}(X_{1T}) = \mathcal{L}\left(\frac{a(x)}{T^2} \underset{\sim}{\xi}_1' C'C \underset{\sim}{\xi}_1 + \frac{b(x)}{T^2} \underset{\sim}{\xi}_2' C'C \underset{\sim}{\xi}_2 \right) ,$$

where $\underset{\sim}{\xi}_1$ and $\underset{\sim}{\xi}_2$ are independent of each other and both follow $\mathrm{N}(0, I_T)$. Arguing as before it is an easy matter to derive

$$
\begin{aligned}
(1.68) \qquad \phi_1(\theta;x) &= \lim_{T\to\infty} \phi_{1T}(\theta;x) \\
&= \prod_{n=1}^{\infty} \left[\left(1 - \frac{2i\theta a(x)}{\left(n-\frac{1}{2}\right)^2 \pi^2} \right) \left(1 - \frac{2i\theta b(x)}{\left(n-\frac{1}{2}\right)^2 \pi^2} \right) \right]^{-\frac{1}{2}} \\
&= \left[D_1(2i\theta a(x))\, D_1(2i\theta b(x)) \right]^{-\frac{1}{2}},
\end{aligned}
$$

where $D_1(\lambda) = \cos\sqrt{\lambda}$.

We can deal with $\hat{\beta}_2$ in (1.65) similarly. Let us put $P(\hat{\beta}_2 \le x) = P(X_{2T} \ge 0)$, where

$$
(1.69) \qquad X_{2T} = \frac{x}{T^2} \sum_{t=1}^{T} (y_{1t} - \bar{y}_1)^2 - \frac{1}{T^2} \sum_{t=1}^{T} (y_{1t} - \bar{y}_1)(y_{2t} - \bar{y}_2).
$$

The c.f. $\phi_{2T}(\theta;x)$ of X_{2T} in (1.69) is given (Problem 4.6) by

$$
(1.70) \qquad \phi_{2T}(\theta;x) = \prod_{t=1}^{T-1} \left[\left(1 - \frac{2i\theta a(x)}{T^2}\gamma_t \right) \left(1 - \frac{2i\theta b(x)}{T^2}\gamma_t \right) \right]^{-\frac{1}{2}},
$$

where γ_t $(t = 1, \cdots, T-1)$ is the t-th largest eigenvalue of CMC' or $C'MC$ defined in (1.24). Then we have

$$
\begin{aligned}
(1.71) \qquad \phi_2(\theta;x) &= \lim_{T\to\infty} \phi_{2T}(\theta;x) \\
&= \prod_{n=1}^{\infty} \left[\left(1 - \frac{2i\theta a(x)}{n^2\pi^2} \right) \left(1 - \frac{2i\theta b(x)}{n^2\pi^2} \right) \right]^{-\frac{1}{2}} \\
&= \left[D_2(2i\theta a(x))\, D_2(2i\theta b(x)) \right]^{-\frac{1}{2}},
\end{aligned}
$$

where $D_2(\lambda) = (\sin\sqrt{\lambda})/\sqrt{\lambda}$.

Figure 1.5 draws the limiting probability densities $f_j(x)$ of $\hat{\beta}_j$, which were numerically computed from $f_j(x) = dF_j(x)/dx$, where

$$
\begin{aligned}
(1.72) \qquad F_j(x) &= \lim_{T\to\infty} P(\hat{\beta}_j \le x) = \lim_{T\to\infty} P(X_{jT} \ge 0) \\
&= \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{1}{\theta}\, \mathrm{Im}(\phi_j(\theta;x))\, d\theta.
\end{aligned}
$$

Moments of $F_j(x)$ can also be computed following the formula (1.39). In particular we have (Problem 4.7)

$$
\begin{aligned}
(1.73) \qquad \mu_1(1) &= \mu_2(1) = \mu_1(3) = \mu_2(3) = 0\,, \\
\mu_1(2) &= \frac{1}{4}\int_0^\infty \frac{u}{\sqrt{\cosh u}}du - \frac{1}{2} = 0.8907\,, \\
\mu_2(2) &= \frac{1}{12}\int_0^\infty \frac{u^{\frac{3}{2}}}{\sqrt{\sinh u}}du - \frac{1}{2} = 0.3965\,, \\
\mu_1(4) &= \frac{7}{192}\int_0^\infty \frac{u^3}{\sqrt{\cosh u}}du - \mu_1(2) - \frac{1}{8} = 3.9304\,, \\
\mu_2(4) &= \frac{1}{320}\int_0^\infty \frac{u^{\frac{7}{2}}}{\sqrt{\sinh u}}du - \mu_2(2) - \frac{1}{8} = 0.6421\,,
\end{aligned}
$$

where $\mu_j(k)$ is the k-th order raw moment of $F_j(x)$. It can also be shown (Problem 4.8) that $F_1(x)$ and $F_2(x)$ are both symmetric about the origin. Table 1.5 reports percent points for $F_j(x)$ $(x \geq 0)$.

Figure 1.5 Table 1.5

The following equivalent expressions will emerge for the weak convergence of X_{1T} in (1.66) :

$$
\begin{aligned}
(1.74) \qquad \mathcal{L}(X_{1T}) &\longrightarrow \mathcal{L}\left(a(x)\sum_{n=1}^\infty \frac{\xi_{1n}^2}{\left(n-\frac{1}{2}\right)^2\pi^2} + b(x)\sum_{n=1}^\infty \frac{\xi_{2n}^2}{\left(n-\frac{1}{2}\right)^2\pi^2} \right) \\
&= \mathcal{L}\left(\int_0^1 w'(t) \begin{pmatrix} x & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} w(t)dt \right) \\
&= \mathcal{L}\left(\int_0^1\int_0^1 (1-\max(s,t))dw'(s) \begin{pmatrix} x & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} dw(t) \right)\,,
\end{aligned}
$$

where $(\xi_{1n},\xi_{2n})' \sim \mathrm{NID}(0,I_2)$ and $w(t)=(w_1(t),w_2(t))'$ is the two-dimensional standard Brownian motion. We shall also have

$$
\begin{aligned}
(1.75) \qquad \mathcal{L}(X_{2T}) &\longrightarrow \mathcal{L}\left(a(x)\sum_{n=1}^\infty \frac{\xi_{1n}^2}{n^2\pi^2} + b(x)\sum_{n=1}^\infty \frac{\xi_{2n}^2}{n^2\pi^2} \right) \\
&= \mathcal{L}\left(\int_0^1 \tilde{w}'(t) \begin{pmatrix} x & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} \tilde{w}(t)dt \right) \\
&= \mathcal{L}\left(\int_0^1\int_0^1 (\min(s,t)-st)dw'(s) \begin{pmatrix} x & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix} dw(t) \right)\,,
\end{aligned}
$$

where $\tilde{w}(t) = w(t) - \int_0^1 w(t)dt$ is the two-dimensional *demeaned Brownian motion.* In terms of the weak convergence of $\hat{\beta}_j$ we shall have the following expressions :

$$
\begin{aligned}
(1.76) \qquad \mathcal{L}(\hat{\beta}_1) &\longrightarrow \mathcal{L}\left(\frac{\int_0^1 w_1(t)w_2(t)dt}{\int_0^1 w_1^2(t)dt} \right) \\
&= \mathcal{L}\left(\frac{\int_0^1 \int_0^1 (1-\max(s,t))dw_1(s)dw_2(t)}{\int_0^1 \int_0^1 (1-\max(s,t))dw_1(s)dw_1(t)} \right),
\end{aligned}
$$

$$
\begin{aligned}
(1.77) \qquad \mathcal{L}(\hat{\beta}_2) &\longrightarrow \mathcal{L}\left(\frac{\int_0^1 \tilde{w}_1(t)\tilde{w}_2(t)dt}{\int_0^1 \tilde{w}_1^2(t)dt} \right) \\
&= \mathcal{L}\left(\frac{\int_0^1 \int_0^1 (\min(s,t)-st)dw_1(s)dw_2(t)}{\int_0^1 \int_0^1 (\min(s,t)-st)dw_1(s)dw_1(t)} \right).
\end{aligned}
$$

The first expressions in (1.74) and (1.75) cannot be converted explicitly into the ratio form as above.

Problems

4.1 Show that the c.f. $\phi_T(\theta)$ of S_T in (1.45) is given by (1.46).

4.2 Prove that V_T in (1.50) converges in distribution to $(\xi^2+1)/2$ given in (1.51).

4.3 Show that the c.f. $\phi_{1T}(\theta)$ of W_{1T} in (1.52) is given by (1.54) and the c.f. $\phi_{2T}(\theta)$ of W_{2T} in (1.53) by (1.55).

4.4 Derive the expressions for cumulants in (1.60) and (1.61).

4.5 Show that the c.f. $\phi_{1T}(\theta; x)$ of X_{1T} in (1.66) is given by (1.67).

4.6 Show that the c.f. $\phi_{2T}(\theta; x)$ of X_{2T} in (1.69) is given by (1.70).

4.7 Obtain moments of $F_1(x)$ and $F_2(x)$ given in (1.73).

4.8 Show that $F_j(x)$ $(j = 1, 2)$ in (1.72) are symmetric about the origin. (In fact the finite sample distributions are also symmetric about the origin.)

1.5. Statistics from the cointegrated process

Let us consider the model

$$\begin{aligned} y_{2t} &= \beta y_{1t} + \varepsilon_{2t}\,, \\ y_{1t} &= y_{1,t-1} + \varepsilon_{1t}\,, \quad y_{10} = 0\,, \quad (t = 1, \cdots, T)\,, \end{aligned} \tag{1.78}$$

where $\beta \neq 0$, $\{y_{1t}\}$ and $\{y_{2t}\}$ are observable, and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ follows $\mathrm{NID}(0, I_2)$. This model might be thought to be equivalent to the state space model or the random walk plus noise model dealt with in Section 1, but it is not the case because the random walk $\{y_{1t}\}$ is observable. The model is also different from the two-dimensional random walk dealt with in Section 4 because the latter assumes that $\{y_{1t}\}$ and $\{y_{2t}\}$ are independent of each other.

The present model is a simplified version of the *cointegrated system* to be discussed in later chapters. Note that y_{2t} is not a random walk because $\Delta y_{2t} = y_{2t} - y_{2,t-1} = \beta\varepsilon_{1t} + \Delta\varepsilon_{2t}$ is not independent, though stationary. The process $\{y_{2t}\}$ is called an *integrated process* of order 1, which is denoted as an I(1) process. A random walk process is also a special case of I(1) processes. The implication of (1.78) is that $y_{2t} - \beta y_{1t}$, a linear combination of two I(1) processes $\{y_{1t}\}$ and $\{y_{2t}\}$, follows $\mathrm{NID}(0, 1)$. In general, following Engle and Granger (1987), a vector-valued process $\{y_t\}$ is said to be integrated of order d if $\{\Delta^d y_t\}$ is stationary, and is called a cointegrated process of order $(d,\ b)$ if there exists a linear combination $\alpha' y_t$ $(\alpha \neq 0)$, say, which is $\mathrm{I}(d - b)$.

Here we consider the estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ of the cointegration parameter β defined by

$$\hat{\beta}_1 = \sum_{t=1}^{T} y_{1t}\, y_{2t} \Big/ \sum_{t=1}^{T} y_{1t}^2\,, \tag{1.79}$$

$$\hat{\beta}_2 = \sum_{t=1}^{T} (y_{1t} - \bar{y}_1)(y_{2t} - \bar{y}_2) \Big/ \sum_{t=1}^{T} (y_{1t} - \bar{y}_1)^2\,. \tag{1.80}$$

Unlike in the spurious regressions discussed in Section 4 the estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ are consistent and $T(\hat{\beta}_j - \beta)$ $(j = 1, 2)$ have nondegenerate limiting distributions, which we now show.

Put $P(T(\hat{\beta}_1 - \beta) \le x) = P(X_{1T} \ge 0)$, where

$$
\begin{aligned}
(1.81) \qquad X_{1T} &= \frac{x}{T^2}\sum_{t=1}^{T} y_{1t}^2 - \frac{1}{T}\sum_{t=1}^{T} y_{1t}\,\varepsilon_{2t} \\
&= \underset{\sim}{\varepsilon}' \begin{pmatrix} \dfrac{x}{T^2}C'C & -\dfrac{1}{2T}C' \\ -\dfrac{1}{2T}C & 0 \end{pmatrix} \underset{\sim}{\varepsilon}
\end{aligned}
$$

with $\underset{\sim}{\varepsilon} = (\varepsilon_{11}, \cdots, \varepsilon_{1T}, \varepsilon_{21}, \cdots, \varepsilon_{2T})'$. The c.f. $\phi_{1T}(\theta; x)$ of X_{1T} in (1.81) is given (Problem 5.1) by

$$
\begin{aligned}
(1.82) \qquad \phi_{1T}(\theta; x) &= \prod_{t=1}^{T}\left(1 - (2i\theta x - \theta^2)\frac{\lambda_t}{T^2}\right)^{-\frac{1}{2}} \\
&= \prod_{t=1}^{T}[(1 - 2i\theta a_t)(1 - 2i\theta b_t)\,]^{-\frac{1}{2}} ,
\end{aligned}
$$

where λ_t is defined in (1.4), while

$$
a_t,\ b_t = \frac{1}{2}\left[\frac{x\lambda_t}{T^2} \pm \sqrt{\frac{\lambda_t}{T^2}\left(\frac{x^2\lambda_t}{T^2} + 1\right)}\right] .
$$

From the last expression for $\phi_{1T}(\theta; x)$ in (1.82) we have

$$
\mathcal{L}(X_{1T}) = \mathcal{L}\left(\sum_{t=1}^{T}(a_t\xi_{1t}^2 + b_t\xi_{2t}^2)\right) ,
$$

where $(\xi_{1t}, \xi_{2t})' \sim \mathrm{NID}(0, I_2)$. It now holds that

$$
\begin{aligned}
(1.83) \qquad \phi_1(\theta; x) &= \lim_{T\to\infty} \phi_{1T}(\theta; x) \\
&= \prod_{n=1}^{\infty}\left(1 - \frac{2i\theta x - \theta^2}{\left(n - \dfrac{1}{2}\right)^2 \pi^2}\right)^{-\frac{1}{2}} \\
&= (D_1(2i\theta x - \theta^2))^{-\frac{1}{2}} ,
\end{aligned}
$$

where $D_1(\lambda) = \cos\sqrt{\lambda}$.

We also put $P(T(\hat{\beta}_2 - \beta) \le x) = P(X_{2T} \ge 0)$, where

$$
(1.84) \qquad X_{2T} = \frac{x}{T^2}\sum_{t=1}^{T}(y_{1t} - \bar{y}_1)^2 - \frac{1}{T}\sum_{t=1}^{T}(y_{1t} - \bar{y}_1)\varepsilon_{2t} .
$$

The c.f. $\phi_{2T}(\theta; x)$ of X_{2T} in (1.84) is given (Problem 5.2) by

$$
\begin{aligned}
(1.85) \qquad \phi_{2T}(\theta; x) &= \prod_{t=1}^{T-1} \left(1 - (2i\theta x - \theta^2)\frac{\gamma_t}{T^2}\right)^{-\frac{1}{2}} \\
&= \prod_{t=1}^{T-1} [(1 - 2i\theta c_t)(1 - 2i\theta d_t)]^{-\frac{1}{2}},
\end{aligned}
$$

where γ_t is defined in (1.24), while

$$
c_t,\ d_t = \frac{1}{2}\left[\frac{x\gamma_t}{T^2} \pm \sqrt{\frac{\gamma_t}{T^2}\left(\frac{x^2\gamma_t}{T^2} + 1\right)}\right].
$$

Thus we have

$$
\mathcal{L}(X_{2T}) = \mathcal{L}\left(\sum_{t=1}^{T-1}(c_t\xi_{1t}^2 + d_t\xi_{2t}^2)\right).
$$

It now holds that

$$
\begin{aligned}
(1.86) \qquad \phi_2(\theta; x) &= \lim_{T\to\infty} \phi_{2T}(\theta; x) \\
&= \prod_{n=1}^{\infty}\left(1 - \frac{2i\theta x - \theta^2}{n^2\pi^2}\right)^{-\frac{1}{2}} \\
&= (D_2(2i\theta x - \theta^2))^{-\frac{1}{2}},
\end{aligned}
$$

where $D_2(\lambda) = (\sin\sqrt{\lambda})/\sqrt{\lambda}$.

Figure 1.6 draws the limiting probability densities $f_j(x) = dF_j(x)/dx$ of $T(\hat{\beta}_j - \beta)$, where $F_j(x)$ was computed following the last line of (1.72). Moments of $F_j(x)$ can also be computed following the formula (1.39). In particular we have (Problem 5.3)

$$
\begin{aligned}
(1.87) \qquad \mu_1(1) &= \mu_2(1) = \mu_1(3) = \mu_2(3) = 0, \\
\mu_1(2) &= \int_0^{\infty} \frac{u}{\sqrt{\cosh u}} du = 5.5629, \\
\mu_2(2) &= \int_0^{\infty} \frac{u^{\frac{3}{2}}}{\sqrt{\sinh u}} du = 10.7583, \\
\mu_1(4) &= \frac{3}{64}\int_0^{\infty} \frac{u^4}{\sqrt{\cosh u}}\left(u\tanh^2 u + \frac{2}{3}\tanh u - \frac{2}{3}u\right) du = 203.4937, \\
\mu_2(4) &= \frac{3}{64}\int_0^{\infty} \frac{u^{\frac{7}{2}}}{\sqrt{\sinh u}}\left(u^2\coth^2 u - \coth u - \frac{2}{3}u^2\right) du = 558.5358,
\end{aligned}
$$

where $\mu_j(k)$ is the k-th order raw moment of $F_j(x)$. It can also be shown (Problem 5.4) that $F_1(x)$ and $F_2(x)$ are both symmetric about the origin. Table 1.6 reports percent points for $F_j(x)$ $(x \ge 0)$.

Figure 1.6 Table 1.6

The following equivalent expressions will emerge for the weak convergence of X_{1T} in (1.81) :

$$
\text{(1.88)} \qquad \mathcal{L}(X_{1T}) \longrightarrow \mathcal{L}\left(\sum_{n=1}^{\infty}(A_n\xi_{1n}^2 + B_n\xi_{2n}^2)\right) = \mathcal{L}\left(x\int_0^1 w_1^2(t)dt - \int_0^1 w_1(t)dw_2(t)\right),
$$

where $(\xi_{1n}, \xi_{2n})' \sim \mathrm{NID}(0, I_2)$ and $w(t) = (w_1(t), w_2(t))'$ is the two-dimensional standard Brownian motion, while

$$
A_n\,,\; B_n = \frac{1}{2}\left[\frac{x}{\left(n-\frac{1}{2}\right)^2\pi^2} \pm \sqrt{\frac{1}{\left(n-\frac{1}{2}\right)^2\pi^2}\left(\frac{x^2}{\left(n-\frac{1}{2}\right)^2\pi^2}+1\right)}\right].
$$

The double integral expression is not possible in the present case, unlike in the case of spurious regressions. We shall also have

$$
\text{(1.89)} \qquad \mathcal{L}(X_{2T}) \longrightarrow \mathcal{L}\left(\sum_{n=1}^{\infty}(C_n\xi_{1n}^2 + D_n\xi_{2n}^2)\right) = \mathcal{L}\left(x\int_0^1 \tilde{w}_1^2(t)dt - \int_0^1 \tilde{w}_1(t)d\tilde{w}_2(t)\right),
$$

where $\tilde{w}(t) = (\tilde{w}_1(t), \tilde{w}_2(t))'$ is the two-dimensional demeaned Brownian motion, while

$$
C_n\,,\; D_n = \frac{1}{2}\left[\frac{x}{n^2\pi^2} \pm \sqrt{\frac{1}{n^2\pi^2}\left(\frac{x^2}{n^2\pi^2}+1\right)}\right].
$$

As for the weak convergence of $T(\hat{\beta}_1 - \beta)$ and $T(\hat{\beta}_2 - \beta)$ we shall have the following expressions :

$$
\text{(1.90)} \qquad \mathcal{L}(T(\hat{\beta}_1 - \beta)) \longrightarrow \mathcal{L}\left(\frac{\int_0^1 w_1(t)dw_2(t)}{\int_0^1 w_1^2(t)dt}\right),
$$

$$\mathcal{L}(T(\hat{\beta}_2 - \beta)) \longrightarrow \mathcal{L}\left(\frac{\int_0^1 \tilde{w}_1(t)\, d\tilde{w}_2(t)}{\int_0^1 \tilde{w}_1^2(t) dt}\right). \tag{1.91}$$

The cointegrated system (1.78) is quite restricted. More generally components of $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ are correlated and $\{\varepsilon_t\}$ may be dependent. Then the LSE's of β will have a different distribution. We also need to test if cointegration exists among the components of multiple time series. We will discuss those topics in Chapter 11.

Problems

5.1 Show that the c.f. $\phi_{1T}(\theta; x)$ of X_{1T} in (1.81) is given by (1.82).

5.2 Show that the c.f. $\phi_{2T}(\theta; x)$ of X_{2T} in (1.84) is given by (1.85).

5.3 Obtain moments of $F_1(x)$ and $F_2(x)$ given in (1.87).

5.4 Show that the limiting distributions of $T(\hat{\beta}_1 - \beta)$ and $T(\hat{\beta}_2 - \beta)$ are both symmetric about the origin. (In fact the finite sample distributions are also symmetric.)