Nonstationary Time Series

May 8, 2012

Junhui Qian

1 Basic Assumptions

Let (w_t) be an m-dimensional stochastic process and we define

$$B_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t,$$

where $r \in [0, 1]$ and $[\cdot]$ denotes the first integer less than or equal to nr. Note that $B_n(r)$ is a stochastic process constructed from w_t and that the sample path of B_n is in $D[0, 1]$, the space of cadlag (right continuous with left limit) function on [0,1].

Assumption 1. We assume that the process w is such that

$$B_n \to_d B,$$

where B is a vector Brownian motion (BM) with a well-defined covariance matrix

$$\Omega = \lim_{n\to\infty} \frac{1}{n} \mathbb{E} \left(\sum_{t=1}^{n} w_t \right) \left(\sum_{t=1}^{n} w_t \right)'.$$

Note that Ω is the long run variance of w_t. If w_t is weakly stationary with absolutely summable autocovariance function

$$\Gamma(k) = \mathbb{E} w_t w_{t-k},$$

then it follows from the Toeplitz (or Kronecker) lemma that

$$\Omega = \lim_{n\to\infty} \sum_{k=-n}^{n} \left(1 - \frac{|k|}{n}\right) \Gamma(k) = \sum_{k=-\infty}^{\infty} \Gamma(k).$$

We may decompose Ω as

$$\Omega = \Sigma + \Lambda + \Lambda', \tag{1}$$

where

$$\Sigma = \Gamma(0), \quad \text{and} \quad \Lambda = \sum_{k=1}^{\infty} \Gamma(k).$$

Let W be a standard vector BM, the covariance matrix of which is the identity matrix. We may represent B by

$$B = \Omega^{1/2} W.$$

Assumption 1 is a statement of the invariance principle (IP) or functional central limit theorem (FCLT). The way we make Assumption 1 is unusual, in the sense that we avoid making direct assumptions (e.g., iid, stationarity, etc.) on w_t. This is a convenient and flexible way of making assumptions. Assumption 1, in particular, allow w_t to be general linear processes defined as follows,

$$w_t = \phi(L)\varepsilon_t = \sum_{k=0}^{\infty} \phi_k \varepsilon_{t-k},$$

where $\sum_k k|\phi_k| < \infty$ and $\varepsilon_t \sim i.i.d.(0, \sigma^2)$. We have

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t = \phi(1) \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} \varepsilon_t + R_n(r),$$

where $\sup_{r\in[0,1]} |R_n(r)| \to_p 0$. We therefore have

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t \to_d \phi(1)\sigma W(r).$$

Of course, Assumption 1 allows more than the linear process defined above. For another example, if ε_t in a martingale difference sequence with finite fourth moment, the IP still holds.

In many applications there may be deterministic trends, say (c_t). We allow c_t to be ℓ-dimensional with i-th component c_{it}. We define $f_n = (f_{n1}, \ldots, f_{n\ell})'$, where $f_{ni} \in D[0,1]$ is given by

$$f_{ni}(r) = \frac{c_{i[nr]}}{n^{\delta_i}}, \text{ for some } \delta_i \geq 0.$$

Assumption 2. For each i, we assume that there exists $\delta_i \geq 0$ and a function $f_i \in L^2[0,1]$ of bounded variation such that

$$f_{ni} \to_{L^2} f_i.$$

2 Fundamental Results

Let $z_t = \sum_{i=1}^t w_i$, or $\Delta z_t = w_t$ with $z_0 = 0$. We first present the continuous mapping theorem,

Continuous Mapping Theorem Suppose $X_n \to_d X$ and the distribution of X is P, and let π be a functional continuous P a.s., then we have

$$\pi(X_n) \to_d \pi(X).$$

The continuous mapping theorem implies, for example,

$$\begin{aligned} \sup_{r\in[0,1]} B_n(r) &\to_d \sup_{r\in[0,1]} B(r) \\ \int_0^1 B_n(r) &\to_d \int_0^1 B(r)dr \\ B_n(1) &\to_d B(1) \end{aligned}$$

The last result can be rewritten as

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_t \to_d N(0,\Omega),$$

which is a central limit theorem (CLT). Thus the IP or FCLT generalizes and implies much more than CLT.

The following lemma presents a few results that are fundamental to the asymptotic analysis of regressions involving nonstationary time series.

Lemma 1: Suppose Assumption 1 and 2 hold for w_t and c_t, we have

(a) $\frac{1}{n^{\delta_i+3/2}}\sum_{t=1}^n c_{it}z_t \to_d \int_0^1 f_i(r)B(r)dr$.

(b) $\frac{1}{n^{\delta_i+1/2}}\sum_{t=1}^n c_{it}w_t \to_d \int_0^1 f_i(r)dB(r)$.

(c) $\frac{1}{n^2}\sum_{t=1}^n z_t z_t' \to_d \int_0^1 B(r)B(r)'dr$.

(d) $\frac{1}{n}\sum_{t=1}^n z_{t-1}w_t' \to_d \int_0^1 B(r)dB(r)' + \Lambda'$.

(e) $\frac{1}{n}\sum_{t=1}^n z_t w_t' \to_d \int_0^1 B(r)dB(r)' + \Delta'$.

Proof:

(a)-(c) We have

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n}\frac{c_{it}}{n^{\delta_i}}\frac{z_t}{\sqrt{n}} &= \int_0^1 f_{ni}(r)B_n(r)dr + o_p(1) \\ \sum_{t=1}^{n}\frac{c_{it}}{n^{\delta_i}}\frac{w_t}{\sqrt{n}} &= \int_0^1 f_{ni}(r)dB_n(r) + o_p(1) \\ \frac{1}{n}\sum_{t=1}^{n}\frac{z_t}{\sqrt{n}}\frac{z_t'}{\sqrt{n}} &= \int_0^1 B_n(r)dB_n(r) + o_p(1). \end{aligned}$$

The results then follow from the continuous mapping theorem.

(d) We prove the scalar case, as the vector case is much more involved. First note that

$$\sum_{t=1}^{n} z_{t-1} w_t = \sum_{t=1}^{n} \left(\sum_{i=1}^{t-1} w_i \right) w_t = \frac{1}{2} \left[\left(\sum_{t=1}^{n} w_t \right)^2 - \sum_{t=1}^{n} w_t^2 \right].$$

We have

$$\frac{1}{n} \sum_{t=1}^{n} z_{t-1} w_t = \frac{1}{2} \left(\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} w_t \right)^2 - \frac{1}{n} \sum_{t=1}^{n} w_t^2 \right) \to_d \frac{1}{2} \left(B(1)^2 - \sigma^2 \right),$$

where $\sigma^2 = \mathrm{var}(w_t)$ and B is a BM with long run variance ω^2. Applying the famed Ito's formula to W_t^2, we have $W(t)^2 = 2 \int_0^t W(r) dW(r) + t$, which implies

$$\int_0^1 W(r) dW(r) = \frac{1}{2} \left(W(1)^2 - 1 \right).$$

Since $B = \omega W$,

$$\int_0^1 B(r) dB(r) = \frac{1}{2} \left(B(1)^2 - \omega^2 \right).$$

Since $\omega^2 = \sigma^2 + 2\lambda$ by definition (The vector version is in (1)), we have

$$\frac{1}{2} \left(B(1)^2 - \sigma^2 \right) = \int_0^1 B(r) dB(r) + \lambda.$$

Note that $\int_0^1 B dB$ is zero-mean. The term λ gives the asymptotic mean of $\frac{1}{n} \sum_{t=1}^{n} z_{t-1} w_t$. For the vector case, the result may be generalized to

$$\frac{1}{n} \sum_{t=1}^{n} z_{t-1} w_t' \to_d \int_0^1 B(r) dB(r)' + \Lambda'.$$

(e) Note that

$$\frac{1}{n} \sum_{t=1}^{n} z_t w_t' = \frac{1}{n} \sum_{t=1}^{n} z_{t-1} w_t' + \frac{1}{n} \sum_{t=1}^{n} w_t w_t' \to_d \int_0^1 B(r) dB(r)' + \Lambda' + \Sigma = \int_0^1 B(r) dB(r)' + \Delta'.$$

3 Cointegration

Definition

We first introduce a notation. We say that $z_t \sim I(k)$ if $(1 - L^k)z_t$ is stationary. By convention, we say $z_t \sim I(0)$ if z_t is stationary. Let x_t be an $m-$dimensional time series and y_t be a scalar process. If both x_t and y_t are $I(1)$ and there exists an $I(0)$ process u_t and $\beta \in \mathbb{R}^m$ such that

$$y_t = x_t'\beta + u_t, \tag{2}$$

we say that x_t and y_t are cointegrated. In economic applications, cointegration is taken to be a long term stable relationship, which may fluctuate in the short term but would reassert itself under some economic force.

Asymptotic Properties of OLS Estimators

To study the statistical properties of the OLS estimator $\hat{\beta}$, we define

$$w_t = (u_t \ \Delta x_t')'.$$

We assume that w_t satisfies Assumption 1 (invariance principle). We partition all quantities ($B(r)$, Ω, etc.) conformably with u_t and Δx_t in the definition of w_t. For example, we write

$$B(r) = \begin{pmatrix} B_1(r) \\ B_2(r) \end{pmatrix}, \quad \Omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \Omega_{22} \end{pmatrix}, \quad \Delta = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \Delta_{22} \end{pmatrix}.$$

Note that partitioned components are denoted by the same letter (lower case for scalar or vector components). We have the following result,

$$n\left(\hat{\beta} - \beta\right) \to_d \left(\int_0^1 B_2(r)B_2(r)'dr\right)^{-1}\left(\int_0^1 B_2(r)dB_1(r) + \delta_{21}\right).$$

To prove this, we write

$$n\left(\hat{\beta}-\beta\right)=\left(\frac{1}{n^2}\sum_{t=1}^{n}x_tx_t'\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^{n}x_tu_t\right).$$

The conclusion then follows from the fundamental lemma in the previous section.

When there are $I(0)$ regressors, in addition to x_t, the limiting distribution of the OLS estimator of β does not change. Consider

$$y_t=v_t'\alpha+x_t'\beta+u_t,$$

where $v_t\sim I(0)$ and $\frac{1}{n}\sum_{t=1}^{n}v_tu_t\to_p 0$. The OLS estimator of β is given by

$$\hat{\beta}=\left(\sum_{t=1}^{n}\tilde{x}_t\tilde{x}_t'\right)^{-1}\left(\sum_{t=1}^{n}\tilde{x}_tu_t\right),$$

where $\tilde{x}_t=x_t-\left(\sum_{t=1}^{n}x_tv_t'\right)\left(\sum_{t=1}^{n}v_tv_t'\right)^{-1}v_t$, i.e., the residual of the regression of x_t on v_t. We have

$$\begin{aligned}\frac{1}{n^2}\sum_{t=1}^{n}\tilde{x}_t\tilde{x}_t' &= \frac{1}{n^2}\sum nx_tx_t'-\frac{1}{n}\left(\sum_{t=1}^{n}x_tv_t'\right)\left(\sum_{t=1}^{n}v_tv_t'\right)^{-1}\left(\sum_{t=1}^{n}v_tx_t'\right)\\ &= \frac{1}{n^2}\sum_{t=1}^{n}x_tx_t'+O_p\left(\frac{1}{n}\right),\end{aligned}$$

and

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\tilde{x}_tu_t &= \frac{1}{n}\sum nx_tu_t-\left(\sum_{t=1}^{n}x_tv_t'\right)\left(\sum_{t=1}^{n}v_tv_t'\right)^{-1}\left(\sum_{t=1}^{n}v_tu_t\right)\\ &= \frac{1}{n^2}\sum_{t=1}^{n}x_tu_t+o_p\left(1\right).\end{aligned}$$

We may now apply the fundamental lemma and confirm our claim.

However, when there are deterministic regressors c_t in addition to x_t, the limiting distribution of the OLS estimator of β may change. Consider

$$y_t = c_t'\alpha + x_t'\beta + u_t,$$

where c_t satisfies Assumption 2. Let $\tilde{x}_t$ be the residual of the regression of x_t on c_t, and let

$$c_t^* = D_n^{-1} c_t,$$

where $D_n = \mathrm{diag}(n^{\delta_1}, \ldots, n^{\delta_\ell})$. We have

$$\begin{aligned}
\frac{1}{n^2}\sum_{t=1}^n \tilde{x}_t\tilde{x}_t' &= \frac{1}{n^2}\sum_{t=1}^n x_t x_t' - \left(\frac{1}{n}\sum_{t=1}^n \frac{x_t}{\sqrt{n}} c_t'\right)\left(\frac{1}{n}\sum_{t=1}^n c_t c_t'\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^n c_t \frac{x_t'}{\sqrt{n}}\right) \\
&= \frac{1}{n^2}\sum_{t=1}^n x_t x_t' - \left(\frac{1}{n}\sum_{t=1}^n \frac{x_t}{\sqrt{n}} c_t^{*\prime}\right)\left(\frac{1}{n}\sum_{t=1}^n c_t^* c_t^{*\prime}\right)^{-1}\left(\frac{1}{n}\sum_{t=1}^n c_t^* \frac{x_t'}{\sqrt{n}}\right) \\
&\to_d \int_0^1 B_2 B_2' - \left(\int_0^1 B_2 f'\right)\left(\int_0^1 f f'\right)^{-1}\left(\int_0^1 f B_2'\right) \\
&= \int_0^1 \tilde{B}_2 \tilde{B}_2',
\end{aligned}$$

where $\tilde{B}_2(r) = B_2(r) - \int_0^1 B_2 f' \left(\int_0^1 f f'\right)^{-1} f(r) = (I - P_f)B_2$, i.e., the residual of the Hilbert space projection of B_2 on the span of f. And we have

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^n \tilde{x}_t u_t &= \frac{1}{n^2}\sum_{t=1}^n x_t u_t - \left(\frac{1}{n}\sum_{t=1}^n \frac{x_t}{\sqrt{n}} c_t'\right)\left(\frac{1}{n}\sum_{t=1}^n c_t c_t'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^n c_t u_t\right) \\
&= \frac{1}{n^2}\sum_{t=1}^n x_t u_t - \left(\frac{1}{n}\sum_{t=1}^n \frac{x_t}{\sqrt{n}} c_t^{*\prime}\right)\left(\frac{1}{n}\sum_{t=1}^n c_t^* c_t^{*\prime}\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^n c_t^* u_t\right) \\
&\to_d \left(\int_0^1 B_2 dB_1' + \delta_{21}\right) - \left(\int_0^1 B_2 f'\right)\left(\int_0^1 f f'\right)^{-1}\left(\int_0^1 f dB_1'\right) \\
&= \int_0^1 \tilde{B}_2 dB_1' + \delta_{21}
\end{aligned}$$

For example, if $c_t = t$, then $f(r) = r$ and

$$\tilde{B}_2(r) = B_2(r) - 3r \int_0^1 B_2(r) r dr,$$

which is definitely different from $B_2(r)$.

Spurious Regression

Let y_t and x_t be $I(1)$. If $y_t - x_t'\beta$ is $I(1)$ for any β, we say that the regression

$$y_t = x_t'\beta + e_t \tag{3}$$

is a spurious regression. If we estimate β in (3) using OLS, we will get inconsistent estimates. The computed R^2 will also be misleading, since it will be random and often close to 1. To see this, define

$$w_t = (\Delta y_t, \Delta x_t')',$$

which is assumed to satisfy the IP with $\Omega > 0$. Since

$$\begin{aligned} \hat{\beta} &= \left(\sum_{t=1}^n x_t x_t' \right)^{-1} \left(\sum_{t=1}^n x_t y_t \right) \\ R^2 &= \frac{\left(\sum_{t=1}^n y_t x_t'\right) \left(\sum_{t=1}^n x_t x_t'\right)^{-1} \left(\sum_{t=1}^n x_t y_t\right)}{\sum_{t=1}^n y_t^2}, \end{aligned}$$

we rewrite

$$\begin{aligned} \hat{\beta} &= \left(\sum_{t=1}^n \frac{x_t}{\sqrt{n}} \frac{x_t'}{\sqrt{n}} \right)^{-1} \left(\sum_{t=1}^n \frac{x_t}{\sqrt{n}} \frac{y_t}{\sqrt{n}} \right) \\ R^2 &= \frac{\left(\sum_{t=1}^n \frac{y_t}{\sqrt{n}} \frac{x_t'}{\sqrt{n}}\right) \left(\sum_{t=1}^n \frac{x_t}{\sqrt{n}} \frac{x_t'}{\sqrt{n}}\right)^{-1} \left(\sum_{t=1}^n \frac{x_t}{\sqrt{n}} \frac{y_t}{\sqrt{n}}\right)}{\sum_{t=1}^n \left(\frac{y_t}{\sqrt{n}}\right)^2}. \end{aligned}$$

We thus obtain

$$\begin{aligned}
\hat{\beta} &\to_d \left(\int_0^1 B_2 B_2'\right)^{-1} \int_0^1 B_2 B_1 \\
R^2 &\to_d \frac{\int_0^1 B_1 B_2' \left(\int_0^1 B_2 B_2'\right)^{-1} \int_0^1 B_2' B_1}{\int_0^1 B_1^2}.
\end{aligned}$$

4 Unit Root Tests