ARTICLES

Statistical Estimation of Simple Regression Model with Nonstationary I(d) Regressor and Stationary Autoregressive Error Term

Soichi SUGIHARA*

We derive the asymptotic nonstandard distributions of OLS and GLS for the simple regression model with nonstationary I(d) regressor $(d > \frac{1}{2})$ and stationary AR(1) error term and show that these two estimators are asymptotically equivalent. The asymptotic normality of these estimators is also derived using certain normalizations.

1. Introduction

Since Granger and Joyeux[2] and Hosking [4] have proposed the idea of fractional differencing, fractional process plays an important role in Econometrics and Finance area to analyze the long memory time series. (In the present paper, fractional process of order d is referred to as I(d) process.) Specifically, Krämer [5] and Phillips and Park[7] have shown the asymptotic equivalence of OLS (Ordinary Least Squares estimators) and GLS (Generalized Least Squares estimators) for the regression model with I(1) regressor(s) and stationary error term (assuming to follow AR(1) or AR(p) process). In the present paper, we extend their model to the one with nonstationary I(d) regressor $(d > \frac{1}{2})$, derive the limiting distributions of OLS and GLS, and show the asymptotic equivalence of these two estimators.

In section 2, model and assumptions are given. In section 3, we derive the asymptotic distributions of OLS and GLS for the simple regression model with nonstationary I(d) regressor $(d > \frac{1}{2})$ and stationary AR(1) error term and show that OLS and GLS are asymptotically equivalent. Asymptotic normality of the estimators is also derived using certain normalizations. A brief summary is given in section 4.

2. The Model and Assumptions

Let us consider the following simple regression model with nonstationary I(d) regressor x_t and stationary AR(1) error term u_t, where L is a lag operator.

* School of Business Administration, Kwansei Gakuin University

$$y_t = \alpha + \beta x_t + u_t, \quad t = 1, 2, \dots, T,$$
$$(1-L)^d x_t = w_t, \quad w_t = \psi(L)\varepsilon_t,$$
$$\phi(L)u_t = v_t.$$

We make the following Assumptions 1~3.

Assumption 1. $d > \frac{1}{2}$. $\psi(L) = \sum_{i=0}^{\infty} \psi_i L^i$ $(\psi_0 = 1)$, $\sum_{i=0}^{\infty} i\,|\psi_i| < \infty$ and all roots of $\psi(z) = 0$ are outside the unit circle. $x_t = 0$ $(t \leq 0)$.

Assumption 2. $\phi(L) = 1 - \phi L$ and $|\phi| < 1$.

Assumption 3. $\begin{pmatrix} \varepsilon_t \\ v_t \end{pmatrix} \sim \mathrm{IID}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_v^2 \end{pmatrix} \right)$.

In particular, in Assumption 1 we assume $d > \frac{1}{2}$. In the present paper, however, we exclude the case when $d = \frac{1}{2}$ (i.e. just nonstationary case), which needs different normalizations. As far as the author knows, there are several papers which treat the case when d=1 (specifically, Krämer [5] and Phillips and Park [7] among others), but little is known about the case when $d > \frac{1}{2}$. We assume that w_t follows a usual linear process. In Assumption 2 we assume that the error term u_t follows stationary AR(1) process, though the extension to stationary AR(p) process can be easily done using Phillips and Park [7]. Extension of Assumption 3 to the correlated case can be done similarly to Phillips and Durlauf [6] , for example, but the analysis is more complicated.

The model can be written in conventional matrix form as

$$y = X\theta + u,$$

where y, X , u and θ are given as follows.

$$y = (y_1, y_2, \dots\dots, y_T)', \qquad u = (u_1, u_2, \dots\dots, u_T)',$$
$$X = \begin{pmatrix} 1, 1, \dots\dots, 1 \\ x_1, x_2, \dots\dots, x_T \end{pmatrix}', \qquad \theta = (\alpha, \beta)'.$$

3. Asymptotic Properties of OLS and GLS

Following Lemma is useful in the subsequent sequel.

Lemma

(1) $x_T(r) \Rightarrow F_{d-1}(r)$, where $x_T(r)$ is the partial sum process defined as

$$x_T(r) = \frac{1}{T^{d-1/2}} x_t + T\left(r - \frac{t}{T}\right) \frac{1}{T^{d-1/2}} (x_t - x_{t-1}) \qquad \left(\frac{t-1}{T} \leqq r \leqq \frac{t}{T};\ t = 1, 2, \dots, T\right).$$

(2) $\frac{1}{T^{d+1/2}} \sum_{t=1}^{T} x_t \Rightarrow \int_0^1 F_{d-1}(r)dr.$

(3) $\frac{1}{T^{2d}} \sum_{t=1}^{T} x_t^2 \Rightarrow \int_0^1 F_{d-1}^2(r)dr.$

(4) $\frac{1}{T^d} \sum_{t=1}^{T} x_t u_t \Rightarrow \frac{\sigma_v}{\phi(1)} \int_0^1 F_{d-1}(r)dW_v(r).$

Here $\Rightarrow$ signifies weak convergence of the associated probability measures as $T \to \infty$, $F_{d-1}(r)$ is a (d-1) fold integrated Brownian motion defined as

$$F_{d-1}(r) = \frac{\psi(1)\sigma_\varepsilon}{\Gamma(d)} \int_0^r (r-s)^{d-1} dW_\varepsilon(s)$$

and $W_v(r)$ and $W_\varepsilon(s)$ are independent standard Brownian motions defined on the probability space $(\Omega, \Im, P)$.

The above Lemma is an extension of Tanaka [8] to the case when $d > \frac{1}{2}$.

Now let $\hat{\theta}$ be the OLS of θ given below,

$$\hat{\theta} = (\hat{\alpha}, \hat{\beta})' = (X'X)^{-1}X'y,$$

and let D be the normalizer of θ

$$D = \begin{pmatrix} T^{1/2} & 0 \\ 0 & T^d \end{pmatrix}.$$

Then $D(\hat{\theta} - \theta)$ can be written as $D(\hat{\theta} - \theta) = (D^{-1}X'XD^{-1})^{-1}D^{-1}X'u$ and using the Lemma we can show that $D^{-1}X'XD^{-1}$ and $D^{-1}X'u$ converge weakly to the following;

$$D^{-1}X'XD^{-1} = \begin{pmatrix} 1 & \frac{1}{T^{d-1/2}} \sum_{t=1}^{T} x_t \\ * & \frac{1}{T^{2d}} \sum_{t=1}^{T} x_t^2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \int_0^1 F_{d-1}(r)dr \\ * & \int_0^1 F_{d-1}^2(r)dr \end{pmatrix} \equiv H,$$

$$D^{-1}X'u = \begin{pmatrix} \frac{1}{T^{1/2}} \sum_{t=1}^{T} u_t \\ \frac{1}{T^d} \sum_{t=1}^{T} x_t u_t \end{pmatrix} \Rightarrow \frac{\sigma_v}{\phi(1)} \begin{bmatrix} W_v(1) \\ \int_0^1 F_{d-1}(r)dW_v(r) \end{bmatrix} \equiv \frac{1}{\phi(1)} K.$$

Therefore as $D(\hat{\theta} - \theta)$ converges weakly to the following;

$$D(\hat{\theta} - \theta) \Rightarrow \frac{1}{\phi(1)} H^{-1}K.$$

Let us now turn to the GLS of θ.

Let Σ be the covariance matrix of u.

$$\Sigma = E(uu')$$

$$= \frac{\sigma_v^2}{1-\phi^2}\begin{pmatrix} 1 & \phi & \cdots & \phi^{T-1} \\ \phi & 1 & \phi\cdots & \phi^{T-2} \\ \vdots & & \ddots & \vdots \\ \phi^{T-1} & \cdots & \cdots & 1 \end{pmatrix} \equiv \sigma_v^2\Omega.$$

As is well known, Ω^{-1} can be decomposed as $\Omega^{-1} = M'M$, where

$$M = \begin{pmatrix} (1-\phi^2)^{1/2} & 0 & \cdots\cdots & 0 \\ -\phi & 1 & \cdots\cdots & 0 \\ 0 & -\phi & 1\cdots\cdots & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & -\phi\ 1 \end{pmatrix}.$$

Let $\tilde{\theta}$ be the GLS of θ defined as

$$\tilde{\theta} = (\tilde{\alpha}, \tilde{\beta})' = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y.$$

Then $D(\tilde{\theta}-\theta)$ can be written as

$$D(\tilde{\theta}-\theta) = (D^{-1}X'M'MXD^{-1})^{-1}\, D^{-1}X'M'Mu$$

and using the above Lemma, we can similarly show that $D^{-1}X'M'MXD^{-1}$ and $D^{-1}X'M'Mu$ converge weakly to the following

$$D^{-1}X'M'MXD^{-1} = \begin{pmatrix} \dfrac{(1-\phi^2)+(1-\phi)^2(T-1)}{T} & \dfrac{(1-\phi^2)x_1+(1-\phi)\sum_{t=2}^{T}(1-\phi L)x_t}{T^{1/2+d}} \\ * & \dfrac{(1-\phi^2)x_1^2+\sum_{t=2}^{T}((1-\phi L)x_t)^2}{T^{2d}} \end{pmatrix}$$

$$\Rightarrow \phi^2(1)H.$$

$$D^{-1}X'M'Mu = \begin{pmatrix} \dfrac{(1-\phi^2)u_1+(1-\phi)\sum_{t=2}^{T}v_t}{T^{1/2}} \\ \dfrac{(1-\phi^2)x_1u_1+\sum_{t=2}^{T}((1-\phi L)x_t)v_t}{T^d} \end{pmatrix}$$

$$\Rightarrow \phi(1)K.$$

In the derivation above, we use the Beveridge-Nelson decomposition $\phi(L) = (1-\phi L) = \phi(1) + \phi(1-L)$ and we can neglect those terms that converge to zero in probability as $T \to \infty$.

Therefore after canceling $\phi(1)$ once, we get

$$D(\tilde{\theta}-\theta) \Rightarrow \frac{1}{\phi(1)} H^{-1}K.$$

To the summary, we get the next Theorem.

Theorem 1. As $T \to \infty$, $D(\hat{\theta}-\theta)$ and $D(\tilde{\theta}-\theta)$ have the same limiting distribution, that is,

$$D(\hat{\theta}-\theta) \Rightarrow \frac{1}{\phi(1)} H^{-1}K,$$

$$D(\tilde{\theta}-\theta) \Rightarrow \frac{1}{\phi(1)} H^{-1}K.$$

Notice that each element of $D(\hat{\theta}-\theta)$ and $D(\tilde{\theta}-\theta)$ converges weakly to the following.

$$T^{1/2}(\hat{\alpha}-\alpha) \text{ and } T^{1/2}(\tilde{\alpha}-\alpha) \Rightarrow \frac{\sigma_v}{\phi(1)} \frac{\int_0^1 F_{d-1}^*(r)dW_v(r)}{\int_0^1 F_{d-1}^*(r)^2 dr},$$

$$T^d(\hat{\beta}-\beta) \text{ and } T^d(\tilde{\beta}-\beta) \Rightarrow \frac{\sigma_v}{\phi(1)} \frac{\int_0^1 P(r)dW_v(r)}{\int_0^1 P^2(r)dr},$$

where

$$F_{d-1}^*(r) \equiv F_{d-1}(r) - \int_0^1 F_{d-1}(r)dr,$$

$$P(r) \equiv 1 - \left(\frac{\int_0^1 F_{d-1}(r)dr}{\int_0^1 F_{d-1}^2(r)dr}\right) F_{d-1}(r),$$

and $F_{d-1}^*(r)$ is the demeaned $F_{d-1}(r)$.

We see that Theorem 1 is an extension of Krämer [5] and Phillips and Park [7] to the case when $d \rangle \frac{1}{2}$. We also see that it is an extended version of Grenander and Rosenblatt [3], where the asymptotic equivalence of OLS and GLS of θ are shown under the following Grenander's conditions (1)~(3).

(1) $\lim_{T\to\infty} a_T(0) = \infty$ where $a_T(h) \equiv \sum_{t=1}^{T-h} x_t x_{t+h}$ $(h = 0, 1, \cdots)$; $= \sum_{t=1-h}^{T} x_t x_{t+h}$ $(h = 0, -1, \cdots)$.

(2) $\lim_{T\to\infty} \frac{x_{T+1}^2}{a_T(0)} = 0$.

(3) $\lim_{T\to\infty} \frac{a_T(h)}{a_T(0)} \equiv \rho(h)$ exists for $h = 0, \pm 1, \ldots\ldots\ldots$.

If ϕ is unknown, $\tilde{\theta}$ is not a feasible estimator. ϕ can be consistently estimated by the usual two step procedure based on the OLS residuals.

Although Theorem 1 is useful, the limiting distributions are nonstandard. We can also get the limiting normal distributions after certain normalizations described below. For this purpose, let us first consider $(D^{-1}X'XD^{-1})^{1/2}D(\hat{\theta}-\theta)$ and $(D^{-1}X'\Omega^{-1}XD^{-1})^{1/2}D(\hat{\theta}-\theta)$. Similar arguments as above derive that

$$(D^{-1}X'XD^{-1})^{1/2}D(\hat{\theta}-\theta) \Rightarrow H^{1/2}\frac{1}{\phi(1)}H^{-1}K = \frac{1}{\phi(1)}H^{-1/2}K,$$

$$(D^{-1}X'\Omega^{-1}XD^{-1})^{1/2}D(\hat{\theta}-\theta) \Rightarrow (\phi^2(1)H)^{1/2}\frac{1}{\phi(1)}H^{-1}K = H^{-1/2}K.$$

Now let $\mathfrak{I}_\varepsilon$ denote the sub σ-field generated by $\{W_\varepsilon(r) | 0 \le r \le 1\}$. Conditionally upon $\mathfrak{I}_\varepsilon$, $\frac{1}{\phi(1)}H^{-1/2}K$ and $H^{-1/2}K$ are distributed as $N(0, (\frac{\sigma_v}{\phi(1)})^2 I_2)$ and $N(0, \sigma_v^2 I_2)$ respectively, but these distributions do not depend on the realization of $W_\varepsilon(r)$. Therefore we can see that these are also the unconditional distributions of $\frac{1}{\phi(1)}H^{-1/2}K$ and $H^{-1/2}K$ respectively.

Similar arguments show that, as $T \to \infty$, $(D^{-1}X'XD^{-1})^{1/2}D(\tilde{\theta}-\theta)$ and $(D^{-1}X'\Omega^{-1}XD^{-1})^{1/2}D(\tilde{\theta}-\theta)$ also converge weakly to $N(0, (\frac{\sigma_v}{\phi(1)})^2 I_2)$ and $N(0, \sigma_v^2 I_2)$ respectively. Therefore we get the following Theorem

Theorem 2.

(1) $(D^{-1}X'XD^{-1})^{1/2}D(\hat{\theta}-\theta) \Rightarrow N(0, (\frac{\sigma_v}{\phi(1)})^2 I_2)$,

$(D^{-1}X'\Omega^{-1}XD^{-1})^{1/2}D(\hat{\theta}-\theta) \Rightarrow N(0, \sigma_v^2 I_2)$.

(2) (1) remains true if $\hat{\theta}$ is replaced with $\tilde{\theta}$.

We see that Theorem 2 is an extension of Phillips and Park [7] to the case when $d > \frac{1}{2}$.

4. Summary

In the present paper, extending Krämer [5] and Phillips and Park [7], we derive the limiting nonstandard distributions of OLS and GLS of the regression parameters for the simple regression model with nonstationary I(d) regressor ($d > \frac{1}{2}$) and stationary AR(1) error term. We show the asymptotic equivalence of OLS and GLS. We also derive the limiting normal distributions after certain normalizations of the estimators. The results in this paper are asymptotic ones and the comparison in finite samples is a task that remains for the future.

References

[1] Billingsley, P. (1968), *Convergence of Probability Measures*, John Wiley, New York.

[2] Granger, C.W.J. and Joyeux, R.(1980) "An Introduction to Long-range Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, 1, 15-30.

[3] Grenander, U. and Rosenblatt, M. (1957) *Analysis of Stationary Time Series*, John Wiley, New York.

[4] Hosking, J.R.M. (1981) "Fractional differencing," *Biometrica*, 68, 165-176.

[5] Krämer, W. (1986), "Least Squares Regression When the Independent Variable follows an ARIMA Process," *Journal of the American Statistical Association*, 81, 150-154.

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

[7] Phillips, P. C. B. and Park, J.Y. (1988), "Asymptotic Equivalence of Ordinary Least Squares and Generalized Least Squares in Regressions with Integrated Regressors," *Journal of the American Statistical Association*, 83, 111-115.

[8] Tanaka, K. (1996), *Time Series Analysis; Nonstationary and Noninvertible Distribution Theory*, John Wiley, New York.