学位請求論文 “Statistical Analysis of Nonlinear Time Series” 本文要旨　2013 年 9 月

一橋大学大学院経済学研究科博士課程 ED111001 植松 良公

This dissertation consists of four chapters. Chapter 1 contains the introduction of this dissertation and the overviews of the other chapters. Chapters 2, 3 and 4 give the main contributions of the author.

Summary of Chapter 2

Chapter 2 presents a model that has an asymptotically efficient ordinary least squares (OLS) estimator, irrespective of the singularity of its limiting sample moment matrix. In the literature on time series, Grenander and Rosenblatt's result is necessary to judge the asymptotic efficiency of the OLS estimator with requiring that the regressors satisfy Grenander's conditions. Without the conditions, however, it is not obvious whether the estimator is efficient. In Chapter 2, we introduce such a model by analyzing the regression model with a *slowly varying* (SV) regressor under a quite general assumption on errors. These regressors are known to display asymptotic singularity in the sample moment matrices; that is, Grenander's condition fails.

A positive-valued function L on $\mathbb{R}_+$ is called SV if it satisfies, for any $r > 0$, $L(rn)/L(n) \to 1$ as $n \to \infty$. To deal with an SV function L, we suppose that L has the following *Karamata's representation*:

$$L(n) = c_L \exp\left(\int_B^n \frac{\varepsilon(s)}{s} ds\right) \quad \text{for} \quad n \geq B$$

for some $B > 0$. Here $c_L > 0$, ε is continuous and $\varepsilon(n) \to 0$ as $n \to \infty$. Note that any SV function is of order $o(n^\alpha)$ for all $\alpha > 0$. With some additional conditions on L, we consider the following regression model

$$y_t = \beta_0 + \beta_1 L(t) + u_t \quad \text{for} \quad t = 1, \ldots, n, \quad \text{or} \quad y = X\beta + u,$$

where $y = [y_1, \ldots, y_n]'$, $\beta = [\beta_0, \beta_1]'$ and $X = [\iota, L]$ with $\iota = [1, \ldots, 1]'$ and $L = [L(1), \ldots, L(n)]'$. The term u_t represents the regression error and is modeled to include a very wide class of stationary processes with a positive and continuous spectrum. If we write $u = [u_1, \ldots, u_n]'$, the variance is given by

$$\mathrm{Var}(u) = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{bmatrix} = [\Gamma_0, \Gamma_1, \ldots, \Gamma_{n-1}] = \Gamma,$$

where Γ_t is the tth column vector of $\mathrm{Var}(u)$. Using these notations, we may define the OLS and GLS estimators as $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$ and $\hat{\beta}_{GLS} = (X'\Gamma^{-1}X)^{-1}X'\Gamma^{-1}y$, respectively.

The OLS estimator is said to be asymptotically efficient if

$$F_n \left[\mathrm{Var}(\hat{\beta}_{OLS}) - \mathrm{Var}(\hat{\beta}_{GLS}) \right] F_n \to 0$$

is true for some common standardizing matrix F_n. In Chapter 2, the convergence is proved for the SV model. That is, if regularity conditions hold, then $\mathrm{Var}(\hat{\beta}_{OLS})$ and $\mathrm{Var}(\hat{\beta}_{GLS})$ of the model have the same asymptotic variance of the form

$$\sigma^2 \begin{bmatrix} \frac{1}{n\varepsilon(n)^2} & -\frac{1}{nL(n)\varepsilon(n)^2} \\ -\frac{1}{nL(n)\varepsilon(n)^2} & \frac{1}{nL(n)^2\varepsilon(n)^2} \end{bmatrix} (1 + o(1)),$$

implying that the OLS estimator is asymptotically efficient with the standardizing matrix $F_n = diag[\sqrt{n}\varepsilon(n), \sqrt{n}L(n)\varepsilon(n)]$.

Summary of Chapter 3

Chapter 3 considers a unit root test in the presence of a SV regressor L. The definition of the SV function L follows from the preceding section, but the assumptions to be imposed here are somewhat different. Consider the model

$$y_t = \alpha + \beta L(t) + u_t \quad \text{and} \quad u_t = \rho u_{t-1} + v_t \quad \text{for} \quad t = 1, \ldots, n, \tag{1}$$

where $\rho = 1$ and $\{v_t\}$ is assumed to be a one-summable linear process with $E|v_t|^p < \infty$ for some $p > 2$. The regressor $L(t)$ is given by an SV function that satisfies some regularity conditions. The first result is the derivation of the limiting distribution of the OLS estimator:

$$\frac{1}{\sqrt{n}} \begin{bmatrix} \varepsilon(n)\left(\hat{\alpha}_n - \alpha\right) \\ L(n)\varepsilon(n)\left(\hat{\beta}_n - \beta\right) \end{bmatrix} \xrightarrow{d} N\left(0, \frac{2\sigma_L^2}{27} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right),$$

where σ_L^2 is the long-run variance of v_t. This weak convergence is obtained by Mynbaev's CLT (see Mynbaev 2009, 2011a), which is applicable for a weighted sum of linear processes. Because any SV function possesses an asymptotic order of $o(\sqrt{n})$, the OLS estimators cannot be consistent. This fact contrasts with the case where the simple trend t is employed. Considering models with an SV regressor, we therefore remark that the existence of a unit root leads to a meaningless regression and that testing for a unit root is indispensable.

Let $W(\cdot)$ denote the standard Brownian motion obtained in the limit of the partial sum process $\sigma_L^{-1} n^{-1/2} \sum_{t=1}^{[\cdot n]} v_t$. The next result we investigate is the asymptotic behaviors of the unit root test statistics, or estimated regression coefficient $\hat{\rho}_n$ and corresponding t-statistic $t_{\hat{\rho}_n}$, based on the regression residuals. Under the regularity conditions, we obtain

$$n\varepsilon(n)^2 (\hat{\rho}_n - 1) \xrightarrow{d} -\frac{U_1}{2V_1} \quad \text{and} \quad \varepsilon(n)^2 t_{\hat{\rho}_n} \xrightarrow{d} -\frac{\sigma_L}{\sigma_S} \frac{U_1}{2\sqrt{V_1}},$$

where σ_S^2 is the short-run variance of v_t, and U_1 and U_2 are given by

$$U_1 = \left\{ \int_0^1 (1+\log r) W(r) dr \right\}^2 \quad \text{and}$$
$$V_1 = \int_0^1 W(r)^2 dr - \left(\int_0^1 W(r) dr \right)^2 - \left\{ \int_0^1 (1+\log r) W(r) dr \right\}^2 .$$

These results are derived by an application of both Mynbaev's CLT and FCLT for a linear process. When it comes to testing a unit root by these statistics, however, it will turn out to be useless in that the finite sample approximation is poor.

To overcome the difficulty, we first consider the following no-constant model

$$y_t = \beta L(t) + u_t \quad \text{and} \quad u_t = \rho u_{t-1} + v_t \quad \text{for} \quad t = 1, \ldots, n, \tag{2}$$

where the same assumptions on $L(t)$ and v_t continue to hold. Then, we have similar weak convergence results

$$\frac{L(n)}{\sqrt{n}} \left(\hat{\beta}_n - \beta \right) \xrightarrow{d} N\left(0, \frac{\sigma_L^2}{3} \right),$$

and

$$n(\hat{\rho}_n - 1) \xrightarrow{d} \frac{U_2 - \sigma_S^2/\sigma_L^2}{2V_2} \quad \text{and} \quad t_{\hat{\rho}_n} \xrightarrow{d} \frac{\sigma_L}{\sigma_S} \frac{U_2 - \sigma_S^2/\sigma_L^2}{2\sqrt{V_2}},$$

where

$$U_2 = \left\{ W(1) - \int_0^1 W(r)dr \right\}^2 \quad \text{and} \quad V_2 = \int_0^1 W(r)^2 dr - \left(\int_0^1 W(r)dr \right)^2.$$

In practice, it may not be appropriate to suppose that the true model has no constant term. However, it is worth analyzing the situation where the true model is given by (1), which possesses a constant term, but the no-constant model (2) is employed for regression. Then, we still have the same asymptotic result based on the no-constant model with the effect of a constant term declining at the rate $O(n^{-1/2})$. This manipulation brings about a significant improvement in terms of the size and power in finite sample situations and a unit root test based on this procedure is recommended. Applying the result, we finally give general Phillips and Perron type test statistics.

Summary of Chapter 4

Chapter 4 studies estimation and inference for nonlinear regression models with integrated time series by quantile regression. Suppose that the scalar-valued random variable y_t is generated from the following nonlinear model

$$y_t = \alpha_0 + g(x_t, \beta_0) + u_t \quad \text{for} \quad t = 1, \ldots, n, \tag{3}$$

where $g : \mathbb{R} \times \mathbb{R}^\ell \to \mathbb{R}$ is a known regression function, and x_t and u_t are the covariate and error, respectively. In particular, x_t is specified as a simple AR(1) unit root model $x_t = x_{t-1} + v_t$, where v_t is stationary. The ℓ-dimensional true parameter vector $\theta_0 = (\alpha_0, \beta_0')'$ is assumed to lie in the parameter set $\Theta = A \times B \subset \mathbb{R} \times \mathbb{R}^\ell$. Moreover, let F and f denote the cumulative distribution function (CDF) and pobability density function (PDF) of u_t, respectively. Note that the τth quantile of u_t for a fixed $\tau \in (0,1)$ is simply denoted by $F^{-1}(\tau)$ under some regularity conditions on u_t. Let $u_{t\tau} = u_t - F^{-1}(\tau)$ for $t = 1, \ldots, n$. We may also rewrite the parameter so that $\alpha_0(\tau) = \alpha_0 + F^{-1}(\tau)$ in response to the error term $u_{t\tau}$ and define the new parameter vector $\theta_0(\tau) = \big(\alpha_0(\tau), \beta_0'\big)'$.

The regression function $(x, \beta) \mapsto g(x, \beta)$ is classified into two functional classes as in park and Phillips (2001). The first is the class of *H-regular* functions, which are defined by

$$g(\lambda x, \beta) = \kappa(\lambda) h(x, \beta) + R(x, \lambda, \beta),$$

where the functions κ and h are said to be the *asymptotic order* and *limit homogeneous function* of g, respectively. The last term, $R(x,\lambda,\beta)$, is a remainder. Polynomial functions, distribution-like functions and logarithmic function are included in this H*-regular* class. The second is the class of *I-regular* functions, which are characterized as bounded and integrable functions with respect to x with sufficient smoothness in β.

With this setting, we may obtain the nonlinear quantile regression (NQR) estimator $\hat{\theta}_n(\tau) = \left(\hat{\alpha}_n(\tau), \hat{\beta}_n(\tau)'\right)'$ of $\theta_0(\tau)$ by solving the minimization problem

$$\hat{\theta}_n(\tau) = \arg\min_{\theta\in\Theta} \sum_{t=1}^{n} \rho_\tau(y_t - \alpha - g(x_t,\beta)),$$

where $\rho_\tau(u) = u\psi_\tau(u)$ with $\psi_\tau(u) = \tau - 1(u<0)$.

To develop the analysis, we are required to make assumptions on errors u_t and v_t. We construct two partial sum processes

$$U_n^\psi(\tau,r) = \frac{1}{n^{1/2}} \sum_{t=1}^{[nr]} \psi_\tau(u_{t\tau}) \quad \text{and} \quad V_n(r) = \frac{1}{n^{1/2}} \sum_{t=0}^{[nr]} v_{t+1}.$$

The process $\{\psi_\tau(u_{t\tau})\}$ is assumed to be a martingale difference sequence. Further, we suppose that the vector $(U_n^\psi, V_n)(r)$ converges weakly to a vector Brownian motion $(U^\psi, V)(r)$ whose covariance matrix is $\Omega(r)$. In addition, we assume that v_t is a linear process for an I-regular g.

One of the main contributions of Chapter 4 is that the asymptotic distribution of the NQR estimator $\hat{\theta}_n(\tau)$ is derived with restricting our attention to the class of H-regular functions. Note that the class of I-regular functions cannot be treated in the same framework due to the irregular convergence rate $n^{1/4}$; this class is investigated later with a restriction to model (3). The regression function $g(x,\cdot)$ is always supposed to be twice continuously differentiable. Define notation of the first and second order derivatives as $\dot{g}(x,\beta) = \partial g(x,\beta)/\partial\beta$ and $\ddot{G}(x,\beta) = \partial^2 g(x,\beta)/\partial\beta\partial\beta'$, and we further write $\ddot{g} = \mathrm{vec}(\ddot{G})$. Corresponding to the ℓ-dimensional vector $\dot{g}$ and ℓ^2-dimensional vector $\ddot{g}$, the asymptotic order matrices $\dot{\kappa}_n$ ($\ell\times\ell$) and $\ddot{\kappa}_n$ ($\ell^2\times\ell^2$) and the vector of the limit homogeneous functions $\dot{h}$ and $\ddot{h}$ are introduced when $\dot{g}$ and $\ddot{g}$ are H-regular. We further let $\tilde{g} = (1,\dot{g}')'$, $\tilde{h} = (1,\dot{h}')'$ and $\tilde{\kappa}_n = \mathrm{diag}(1,\dot{\kappa}_n)$. To obtain the limiting distribution, we need to suppose an additional assumption on the parameter vector θ so that $\theta = \theta_0(\tau) + n^{-1/2}\tilde{\kappa}_n'^{-1}\pi$, where π lies in a compact set $\Pi \subset \mathbb{R}\times\mathbb{R}^\ell$.

Define the derivative from the right of the objective function

$$z_t(\theta) = \tilde{g}_t(\beta)\psi_\tau(y_t - \alpha - g_t(\beta)).$$

Utilizing this function, we may derive the limiting distribution by considering the "first order condition"

$$n^{-1/2}\tilde{\kappa}_n^{-1}\sum_{t=1}^{n} z_t(\hat{\theta}_n(\tau)) = o_p(1).$$

This estimating equation leads to the Bahadur representation of the NQR estimator $\hat{\theta}_n(\tau)$. In consequence, we obtain the result. Let $\dot{g}$ and $\ddot{g}$ be H-regular on B. Then, under some regularity conditions, we have

$$\begin{aligned} & n^{1/2}\tilde{\kappa}_n'(\hat{\theta}_n(\tau) - \theta_0(\tau)) \\ & \xrightarrow{d} \frac{1}{f(F^{-1}(\tau))}\left[\int_0^1 \tilde{h}(V(r),\beta_0)\tilde{h}(V(r),\beta_0)'dr\right]^{-1}\int_0^1 \tilde{h}(V(r),\beta_0)dU^{\psi}(r). \end{aligned}$$

Note that the limiting distribution is not usually (mixed) normal because of the possible nonzero correlation between U^{ψ} and V. Therefore standard inferences are not applicable in this case.

To overcome the difficulty, we suggest fully-modified NQR (FM-NQR) estimator based on the results of Phillips and Hansen (1990) and de Jong (2002). The FM-NQR estimator is constructed by

$$\hat{\theta}_n^+(\tau) = \hat{\theta}_n(\tau) - \frac{n^{-1/2}\tilde{\kappa}_n^{-1}}{\widehat{f(F^{-1}(\tau))}}\frac{\hat{\omega}_{\psi v}}{\hat{\omega}_v^2}S_n^{-1}T_n,$$

where $1/\widehat{f(F^{-1}(\tau))}$, $\hat{\omega}_{\psi v}$ and $\hat{\omega}_v^2$ are consistent estimators. The statistics S_n and T_n satisfy

$$S_n \xrightarrow{p} \int_0^1 \tilde{h}(V(r),\beta_0)\tilde{h}(V(r),\beta_0)'dr \quad \text{and} \quad T_n \xrightarrow{d} \int_0^1 \tilde{h}(V(r),\beta_0)dV(r).$$

If some additional conditions of de Jong (2002) are satisfied, we then have

$$\begin{aligned} & n^{1/2}\tilde{\kappa}_n'\left(\hat{\theta}_n^+(\tau) - \theta_0(\tau)\right) \\ & \xrightarrow{d} \frac{1}{f(F^{-1}(\tau))}\left[\int_0^1 \tilde{h}(V(r),\beta_0)\tilde{h}(V(r),\beta_0)'dr\right]^{-1}\int_0^1 \tilde{h}(V(r),\beta_0)dU^{\psi+}(r), \end{aligned}$$

where $U^{\psi+}(r) = U^{\psi}(r) - \omega_{\psi v}\omega_v^{-2}V(r)$. Therefore, the mixed normality of the limiting random variable is brought to light immediately because $U^{\psi+}$ is easily found to be uncorrelated with V

and, hence, independent of V. Because of the asymptotic normality of the FM-NQR estimator, we can consider testing linear restrictions on the parameter vector.

We have considered the NQR estimator of the nonlinear model only in the case of H-regular $\dot{g}$ and $\ddot{g}$. We then investigate the so-called linear-in-parameter model obtained by confining model (3) to

$$y_t = \alpha_0 + \beta_0 g(x_t) + u_t. \tag{4}$$

The regression function g is either I-regular or H-regular and write $g_t = g(x_t)$. The parameter β_0 is allowed ℓ-dimensional, but is assumed $\ell = 1$ for the sake of brevity. Because the model is linear in parameter, the asymptotics can be derived even if g is an I-regular function as well as an H-regular one.

First, we consider model (3) under restriction (4) with I-regular regression function derivative $\dot{g}$. The limiting distribution of the NQR estimator $\hat{\theta}_n(\tau)$ is summarized as follows. Let g be I-regular on B, and suppose some regularity conditions. Then we have

$$D_n^I(\hat{\theta}_n(\tau) - \theta_0(\tau)) \xrightarrow{d} \frac{1}{f(F^{-1}(\tau))} \begin{bmatrix} U^{\psi}(1) \\ \left(L(1,0) \int_{-\infty}^{\infty} g(x)^2 dx \right)^{-1/2} W(1) \end{bmatrix},$$

where $D_n^I = \mathrm{diag}(n^{1/2}, n^{1/4})$ and the Brownian motion $W(r)$ has variance $r\tau(1-\tau)$, which is the same as the variance of $U^{\psi}(r)$. This implies that $\hat{\alpha}_n(\tau)$ and $\hat{\beta}_n(\tau)$ are asymptotically independent; consequently, the limiting joint distribution is mixed normal of the form

$$MN\left(0,\ \frac{\omega_{\psi}^2}{f(F^{-1}(\tau))^2} \begin{bmatrix} 1 & 0 \\ 0 & L(1,0)\int_{-\infty}^{\infty} g(x)^2 dx \end{bmatrix}^{-1}\right).$$

Hence, standard inferences are applicable in an asymptotic sense. For the case of H-regular functions, we certainly have the same result obtained above.

Finally, we investigate finite sample performances of the NQR estimators with $\tau = 0.5$ via comparison to nonlinear least squares (NLS) estimators by simulations. We observe from simulations that our suggested NQR estimators are preferable to the NLS estimators in terms of estimation accuracy and powers of tests when distributions of regression errors possess fat tails.