

Research Article

Asymptotic Optimality of Estimating Function Estimator for CHARN Model

Tomoyuki Amano

Faculty of Economics, Wakayama University, 930 Sakaedani, Wakayama 640-8510, Japan

Correspondence should be addressed to Tomoyuki Amano, tomchami@center.wakayama-u.ac.jp

Received 15 February 2012; Accepted 9 April 2012

Academic Editor: Hiroshi Shiraishi

Copyright © 2012 Tomoyuki Amano. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

CHARN model is a famous and important model in the finance, which includes many financial time series models and can be assumed as the return processes of assets. One of the most fundamental estimators for financial time series models is the conditional least squares (CL) estimator. However, recently, it was shown that the optimal estimating function estimator (G estimator) is better than CL estimator for some time series models in the sense of efficiency. In this paper, we examine efficiencies of CL and G estimators for CHARN model and derive the condition that G estimator is asymptotically optimal.

1. Introduction

The conditional least squares (CL) estimator is one of the most fundamental estimators for financial time series models. It has the two advantages which can be calculated with ease and does not need the knowledge about the innovation process (i.e., error term). Hence this convenient estimator has been widely used for many financial time series models. However, Amano and Taniguchi [1] proved it is not good in the sense of the efficiency for ARCH model, which is the most famous financial time series model.

The estimating function estimator was introduced by Godambe ([2, 3]) and Hansen [4]. Recently, Chandra and Taniguchi [5] constructed the optimal estimating function estimator (G estimator) for the parameters of the random coefficient autoregressive (RCA) model, which was introduced to describe occasional sharp spikes exhibited in many fields and ARCH model based on Godambe's asymptotically optimal estimating function. In Chandra and Taniguchi [5], it was shown that G estimator is better than CL estimator by simulation. Furthermore, Amano [6] applied CL and G estimators to some important time series models (RCA, GARCH, and nonlinear AR models) and proved that G estimator is

Table 1: MSE of $\hat{\theta}_{CL}$ and $\hat{\theta}_G$ for the parameter a in (4).

a	0.1	0.2	0.3
$\hat{\theta}_{CL}(n = 100)$	0.01103311	0.01135393	0.01035005
$\hat{\theta}_G(n = 100)$	0.01096804	0.01113519	0.01006857
$\hat{\theta}_{CL}(n = 200)$	0.00596135	0.00586717	0.00549567
$\hat{\theta}_G(n = 200)$	0.00565094	0.00555699	0.00533862
$\hat{\theta}_{CL}(n = 300)$	0.00371269	0.00376673	0.00351314
$\hat{\theta}_G(n = 300)$	0.00356603	0.00359798	0.00348829

better than CL estimator in the sense of the efficiency theoretically. Amano [6] also derived the conditions that G estimator becomes asymptotically optimal, which are not strict and natural.

However, in Amano [6], G estimator was not applied to a conditional heteroscedastic autoregressive nonlinear model (denoted by CHARN model). CHARN model was proposed by Härdle and Tsybakov [7] and Härdle et al. [8], which includes many financial time series models and is used widely in the finance. Kanai et al. [9] applied G estimator to CHARN model and proved its asymptotic normality. However, Kanai et al. [9] did not compare efficiencies of CL and G estimators and discuss the asymptotic optimality of G estimator theoretically. Since CHARN model is the important and rich model, which includes many financial time series models and can be assumed as return processes of assets, more investigation of CL and G estimators for this model are needed. Hence, in this paper, we compare efficiencies of CL and G estimators and investigate the asymptotic optimality of G estimator for this model.

This paper is organized as follows. Section 2 describes definitions of CL and G estimators. In Section 3, CL and G estimators are applied to CHARN model, and efficiencies of these estimators are compared. Furthermore, we derive the condition of asymptotic optimality of G estimator. We also compare the mean squared errors of $\hat{\theta}_{CL}$ and $\hat{\theta}_G$ by simulation in Section 4. Proofs of Theorems are relegated to Section 5. Throughout this paper, we use the following notation: $|\mathbf{A}|$: Sum of the absolute values of all entries of \mathbf{A} .

2. Definitions of CL and G Estimators

One of the most fundamental estimators for parameters of the financial time series models is the conditional least squares (CL) estimator $\hat{\theta}_{CL}$ introduced by Tjøstheim [10], and it has been widely used in the finance. $\hat{\theta}_{CL}$ for a time series model $\{X_t\}$ is obtained by minimizing the penalty function

$$Q_n(\theta) \equiv \sum_{t=m+1}^n (X_t - E[X_t | \mathbf{F}_t(m)])^2, \quad (2.1)$$

where $\mathbf{F}_t(m)$ is the σ -algebra generated by $\{X_s : t - m \leq s \leq t - 1\}$, and m is an appropriate positive integer (e.g., if $\{X_t\}$ follows k th-order nonlinear autoregressive model, we can take $m = k$). CL estimator generally has a simple expression. However, it is not asymptotically optimal in general (see Amano and Taniguchi [1]).

Hence, Chandra and Taniguchi [5] constructed G estimator $\hat{\theta}_G$ based on Godambe's asymptotically optimal estimating function for RCA and ARCH models. For the definition of $\hat{\theta}_G$, we prepare the following estimating function $G(\theta)$. Let $\{X_t\}$ be a stochastic process which is depending on the k -dimensional parameter θ_0 , then $G(\theta)$ is given by

$$G(\theta) = \sum_{t=1}^n \mathbf{a}_{t-1} h_t, \quad (2.2)$$

where \mathbf{a}_{t-1} is a k -dimensional vector depending on X_1, \dots, X_{t-1} and θ , $h_t = X_t - E[X_t | \mathbf{F}_{t-1}]$, and \mathbf{F}_{t-1} is the σ -field generated by $\{X_s, s \leq t-1\}$. The estimating function estimator $\hat{\theta}_E$ for the parameter θ_0 is defined as

$$G(\hat{\theta}_E) = \mathbf{0}. \quad (2.3)$$

Chandra and Taniguchi [5] derived the asymptotic variance of $\sqrt{n}(\hat{\theta}_E - \theta_0)$ as

$$\left(\frac{1}{n} E \left[\frac{\partial}{\partial \theta'} G(\theta_0) \right] \right)^{-1} \frac{E[G(\theta_0)G'(\theta_0)]}{n} \left(\frac{1}{n} E \left[\frac{\partial}{\partial \theta'} G(\theta_0) \right]' \right)^{-1} \quad (2.4)$$

and gave the following lemma by extending the result of Godambe [3].

Lemma 2.1. *The asymptotic variance (2.4) is minimized if $G(\theta) = G^*(\theta)$ where*

$$G^*(\theta) = \sum_{t=1}^n \mathbf{a}_{t-1}^* h_t, \quad (2.5)$$

$$\mathbf{a}_{\theta,t-1}^* = E \left[\frac{\partial h_t}{\partial \theta} \mid \mathbf{F}_{t-1} \right] E \left[h_t^2 \mid \mathbf{F}_{t-1} \right]^{-1}.$$

Based on the estimating function $G^*(\theta)$ in Lemma 2.1, Chandra and Taniguchi [5] constructed G estimator $\hat{\theta}_G$ for the parameters of RCA and ARCH models and showed that $\hat{\theta}_G$ is better than $\hat{\theta}_{CL}$ by simulation. Furthermore, Amano [6] applied $\hat{\theta}_G$ to some important financial time series models (RCA, GARCH, and nonlinear AR models) and showed that $\hat{\theta}_G$ is better than $\hat{\theta}_{CL}$ in the sense of the efficiency theoretically. Amano [6] also derived conditions that $\hat{\theta}_G$ becomes asymptotically optimal. However, in Amano [6], $\hat{\theta}_{CL}$ and $\hat{\theta}_G$ were not applied to CHARN model, which includes many important financial time series models. Hence, in the next section, we apply $\hat{\theta}_{CL}$ and $\hat{\theta}_G$ to this model and prove $\hat{\theta}_G$ is better than $\hat{\theta}_{CL}$ in the sense of the efficiency for this model. Furthermore, conditions of asymptotical optimality of $\hat{\theta}_G$ are also derived.

3. CL and G Estimators for CHARN Model

In this section, we discuss the asymptotics of $\hat{\theta}_{CL}$ and $\hat{\theta}_G$ for CHARN model.

CHARN model of order m is defined as

$$X_t = F_{\theta}(X_{t-1}, \dots, X_{t-m}) + H_{\theta}(X_{t-1}, \dots, X_{t-m})u_t, \quad (3.1)$$

where $F_{\theta}, H_{\theta} : \mathbf{R}^m \rightarrow \mathbf{R}$ are measurable functions, and $\{u_t\}$ is a sequence of i.i.d. random variables with $E u_t = 0$, $E[u_t^2] = 1$ and independent of $\{X_s; s < t\}$. Here, the parameter vector $\theta = (\theta_1, \dots, \theta_k)'$ is assumed to be lying in an open set $\Theta \subset \mathbf{R}^k$. Its true value is denoted by θ_0 .

First we estimate the true parameter θ_0 of (3.1) by use of $\hat{\theta}_{CL}$, which is obtained by minimizing the penalty function

$$\begin{aligned} Q_n(\theta) &= \sum_{t=m+1}^n (X_t - E[X_t | \mathbf{F}_t(m)])^2 \\ &= \sum_{t=m+1}^n (X_t - F_{\theta})^2. \end{aligned} \quad (3.2)$$

For the asymptotics of $\hat{\theta}_{CL}$, we impose the following assumptions.

Assumption 3.1. (1) u_t has the probability density function $f(u) > 0$ a.e. $u \in \mathbf{R}$.

(2) There exist constants $a_i \geq 0, b_i \geq 0, 1 \leq i \leq m$, such that for $\mathbf{x} \in \mathbf{R}^m$ with $|x| \rightarrow \infty$,

$$\begin{aligned} |F_{\theta}(\mathbf{x})| &\leq \sum_{i=1}^m a_i |x_i| + o(|x|), \\ |H_{\theta}(\mathbf{x})| &\leq \sum_{i=1}^m b_i |x_i| + o(|x|). \end{aligned} \quad (3.3)$$

(3) $H_{\theta}(\mathbf{x})$ is continuous and symmetric on \mathbf{R}^m , and there exists a positive constant λ such that

$$H_{\theta}(\mathbf{x}) \geq \lambda \quad \text{for } \forall \mathbf{x} \in \mathbf{R}^m. \quad (3.4)$$

(4) Consider the following

$$\left\{ \sum_{i=1}^m a_i + E|u_1| \sum_{i=1}^m b_i \right\} < 1. \quad (3.5)$$

Assumption 3.1 makes $\{X_t\}$ strict stationary and ergodic (see [11]). We further impose the following.

Assumption 3.2. Consider the following

$$\begin{aligned} E_{\theta} |F_{\theta}(X_{t-1}, \dots, X_{t-m})|^2 &< \infty, \\ E_{\theta} |H_{\theta}(X_{t-1}, \dots, X_{t-m})|^2 &< \infty, \end{aligned} \quad (3.6)$$

for all $\theta \in \Theta$.

Assumption 3.3. (1) F_{θ} and H_{θ} are almost surely twice continuously differentiable in Θ , and their derivatives $\partial F_{\theta}/\partial\theta_j$ and $\partial H_{\theta}/\partial\theta_j$, $j = 1, \dots, k$, satisfy the condition that there exist square-integrable functions A_j and B_j such that

$$\begin{aligned} \left| \frac{\partial F_{\theta}}{\partial\theta_j} \right| &\leq A_j \\ \left| \frac{\partial H_{\theta}}{\partial\theta_j} \right| &\leq B_j, \end{aligned} \quad (3.7)$$

for all $\theta \in \Theta$.

(2) $f(u)$ satisfies

$$\begin{aligned} \lim_{|u| \rightarrow \infty} |u|f(u) &= 0, \\ \int u^2 f(u) du &= 1. \end{aligned} \quad (3.8)$$

(3) The continuous derivative $f'(u) \equiv \partial f(u)/\partial u$ exists on \mathbf{R} and satisfies

$$\begin{aligned} \int \left(\frac{f'}{f} \right)^4 f(u) du &< \infty, \\ \int u^2 \left(\frac{f'}{f} \right)^2 f(u) du &< \infty. \end{aligned} \quad (3.9)$$

From Tjøstheim [10], the following lemma holds.

Lemma 3.4. *Under Assumptions 3.1, 3.2, and 3.3, $\hat{\theta}_{CL}$ has the following asymptotic normality:*

$$\sqrt{n}(\hat{\theta}_{CL} - \theta_0) \xrightarrow{d} \mathbf{U}^{-1} \mathbf{W} \mathbf{U}^{-1}, \quad (3.10)$$

where

$$\begin{aligned} \mathbf{W} &= E \left[H_{\theta_0}^2 \frac{\partial}{\partial \theta} F_{\theta_0} \frac{\partial}{\partial \theta'} F_{\theta_0} \right], \\ \mathbf{U} &= E \left[\frac{\partial}{\partial \theta} F_{\theta_0} \frac{\partial}{\partial \theta'} F_{\theta_0} \right]. \end{aligned} \quad (3.11)$$

Next, we apply $\hat{\theta}_G$ to CHARN model. From Lemma 2.1, $\hat{\theta}_G$ is obtained by solving the equation

$$\sum_{t=m+1}^n \frac{1}{H_{\theta}^2} \frac{\partial}{\partial \theta} F_{\theta}(X_t - F_{\theta}) = 0. \quad (3.12)$$

For the asymptotic of $\hat{\boldsymbol{\theta}}_G$, we impose the following Assumptions.

Assumption 3.5. (1) Consider the following

$$E_{\boldsymbol{\theta}} \left\| \mathbf{a}_{\boldsymbol{\theta}, t-1}^* \right\|^2 < \infty \quad (3.13)$$

for all $\boldsymbol{\theta} \in \Theta$.

(2) $\mathbf{a}_{\boldsymbol{\theta}, t-1}^*$ is almost surely twice continuously differentiable in Θ , and for the derivatives $\partial \mathbf{a}_{\boldsymbol{\theta}, t-1}^* / \partial \theta_j$, $j = 1, \dots, k$, there exist square-integrable functions C_j such that

$$\left| \frac{\partial \mathbf{a}_{\boldsymbol{\theta}, t-1}^*}{\partial \theta_j} \right| \leq C_j, \quad (3.14)$$

for all $\boldsymbol{\theta} \in \Theta$.

(3) $\mathbf{V} = E[(1/H_{\boldsymbol{\theta}_0}^2)(\partial/\partial \boldsymbol{\theta})F_{\boldsymbol{\theta}_0}(\partial/\partial \boldsymbol{\theta}')F_{\boldsymbol{\theta}_0}]$ is $k \times k$ -positive definite matrix and satisfies

$$|\mathbf{V}| < \infty. \quad (3.15)$$

(4) For $\boldsymbol{\theta} \in \mathbf{B}$ (a neighborhood of $\boldsymbol{\theta}_0$ in Θ), there exist integrable functions $P_{\boldsymbol{\theta}}^{ijl}(\mathbf{X}^{(t-1)})$, $Q_{\boldsymbol{\theta}}^{ijl}(\mathbf{X}^{(t-1)})$, and $R_{\boldsymbol{\theta}}^{ijl}(\mathbf{X}^{(t-1)})$ such that

$$\begin{aligned} \left| \frac{\partial^2}{\partial \theta_j \partial \theta_l} \left(\mathbf{a}_{\boldsymbol{\theta}, t-1}^* \right)_i h_t \right| &\leq P_{\boldsymbol{\theta}}^{ijl}(\mathbf{X}^{(t-1)}), \\ \left| \frac{\partial}{\partial \theta_j} \left(\mathbf{a}_{\boldsymbol{\theta}, t-1}^* \right)_i \frac{\partial}{\partial \theta_l} h_t \right| &\leq Q_{\boldsymbol{\theta}}^{ijl}(\mathbf{X}^{(t-1)}), \\ \left| \left(\mathbf{a}_{\boldsymbol{\theta}, t-1}^* \right)_i \frac{\partial^2}{\partial \theta_j \partial \theta_l} h_t \right| &\leq R_{\boldsymbol{\theta}}^{ijl}(\mathbf{X}^{(t-1)}), \end{aligned} \quad (3.16)$$

for $i, j, l = 1, \dots, k$, where $\mathbf{X}^{(t-1)} = (X_1, \dots, X_{t-1})$.

From Kanai et al. [9], the following lemma holds.

Lemma 3.6. *Under Assumptions 3.1, 3.2, 3.3, and 3.5, the following statement holds:*

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_G - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \mathbf{V}^{-1}). \quad (3.17)$$

Finally we compare efficiencies of $\hat{\boldsymbol{\theta}}_{CL}$ and $\hat{\boldsymbol{\theta}}_G$. We give the following theorem.

Theorem 3.7. *Under Assumptions 3.1, 3.2, 3.3, and 3.5, the following inequality holds:*

$$\mathbf{U}^{-1} \mathbf{W} \mathbf{U}^{-1} \geq \mathbf{V}^{-1}, \quad (3.18)$$

and equality holds if and only if H_{θ_0} is constant or $\partial F_{\theta_0} / \partial \theta = 0$ (for matrices \mathbf{A} and \mathbf{B} , $\mathbf{A} \geq \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is positive definite).

This theorem is proved by use of Kholevo inequality (see Kholevo [12]). From this theorem, we can see that the magnitude of the asymptotic variance of $\hat{\theta}_G$ is smaller than that of $\hat{\theta}_{CL}$, and the condition that these asymptotic variances coincide is strict. Therefore, $\hat{\theta}_G$ is better than $\hat{\theta}_{CL}$ in the sense of the efficiency. Hence, we evaluate the condition that $\hat{\theta}_G$ is asymptotically optimal based on local asymptotic normality (LAN). LAN is the concept of local asymptotic normality for the likelihood ratio of general statistical models, which was established by Le Cam [13]. Once LAN is established, the asymptotic optimality of estimators and tests can be described in terms of the LAN property. Hence, its Fisher information matrix Γ is described in terms of LAN, and the asymptotic variance of an estimator has the lower bound Γ^{-1} . Now, we prepare the following Lemma, which is due to Kato et al. [14].

Lemma 3.8. *Under Assumptions 3.1, 3.2, and 3.3, CHARN model has LAN, and its Fisher information matrix Γ is*

$$E \left[\frac{1}{H_{\theta_0}^2} \begin{pmatrix} -\frac{\partial H_{\theta_0}}{\partial \theta} & \frac{\partial F_{\theta_0}}{\partial \theta} \\ c & b \end{pmatrix} \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} -\frac{\partial H_{\theta_0}}{\partial \theta'} \\ \frac{\partial F_{\theta_0}}{\partial \theta'} \end{pmatrix} \right], \quad (3.19)$$

where

$$\begin{aligned} a_t &= u_t (f'(u_t) / f(u_t)) + 1, & b_t &= -(f'(u_t) / f(u_t)), \\ a &= E[a_t^2], & b &= E[b_t^2], & c &= E[a_t b_t]. \end{aligned} \quad (3.20)$$

From this Lemma, the asymptotic variance of $\hat{\theta}_G \mathbf{V}^{-1}$ has the lower bound Γ^{-1} , that is,

$$\mathbf{V}^{-1} \geq \Gamma^{-1}. \quad (3.21)$$

The next theorem gives the condition that \mathbf{V}^{-1} equals Γ^{-1} , that is $\hat{\theta}_G$ becomes asymptotically optimal.

Theorem 3.9. *Under Assumptions 3.1, 3.2, 3.3, and 3.5, if $\partial H_{\theta_0} / \partial \theta = \mathbf{0}$ and u_t is Gaussian, then $\hat{\theta}_G$ is asymptotically optimal, that is,*

$$\mathbf{V}^{-1} = \Gamma^{-1}. \quad (3.22)$$

Finally, we give the following example which satisfies the assumptions in Theorems 3.7 and 3.9.

Example 3.10. CHARN model includes the following nonlinear AR model:

$$X_t = F_{\theta}(X_{t-1}, \dots, X_{t-m}) + u_t, \quad (3.23)$$

where $F_\theta: \mathbf{R}^m \rightarrow \mathbf{R}$ is a measurable function, and $\{u_t\}$ is a sequence of i.i.d. random variables with $E u_t = 0$, $E[u_t^2] = 1$ and independent of $\{X_s; s < t\}$, and we assume Assumptions 3.1, 3.2, 3.3, and 3.5 (for example, we define $F_\theta = \sqrt{a_0 + a_1 X_{t-1}^2 + \dots + a_m X_{t-m}^2}$, where $a_0 > 0$, $a_j \geq 0$, $j = 1, \dots, m$, $\sum_{j=1}^m a_j < 1$). In Amano [6], it was shown that the asymptotic variance of $\hat{\theta}_{\text{CL}}$ attains that of $\hat{\theta}_{\text{G}}$. Amano [6] also showed under the condition that u_t is Gaussian, $\hat{\theta}_{\text{G}}$ is asymptotically optimal.

4. Numerical Studies

In this section, we evaluate accuracies of $\hat{\theta}_{\text{CL}}$ and $\hat{\theta}_{\text{G}}$ for the parameter of CHARN model by simulation. Throughout this section, we assume the following model:

$$X_t = aX_{t-1} + \sqrt{0.2 + 0.1X_{t-1}^2} u_t, \quad (4.1)$$

where $\{u_t\} \sim \text{i.i.d.} N(0, 1)$. Mean squared errors (MSEs) of $\hat{\theta}_{\text{CL}}$ and $\hat{\theta}_{\text{G}}$ for the parameter a are reported in the following Table 1. The simulations are based on 1000 realizations, and we set the parameter value a and the length of observations n as $a = 0.1, 0.2, 0.3$ and $n = 100, 200, 300$.

From Table 1, we can see that MSE of $\hat{\theta}_{\text{G}}$ is smaller than that of $\hat{\theta}_{\text{CL}}$. Furthermore it is seen that MSE of $\hat{\theta}_{\text{CL}}$ and $\hat{\theta}_{\text{G}}$ decreases as the length of observations n increases.

5. Proofs

This section provides the proofs of the theorems. First, we prepare the following lemma to compare the asymptotic variances of CL and G estimators (see Kholevo [12]).

Lemma 5.1. *We define $\boldsymbol{\psi}(\omega)$ and $\boldsymbol{\phi}(\omega)$ as $r \times s$ and $t \times s$ random matrices, respectively, and $h(\omega)$ as a random variable that is positive everywhere. If the matrix $E[\boldsymbol{\phi}\boldsymbol{\phi}'/h]^{-1}$ exists, then the following inequality holds:*

$$E[\boldsymbol{\psi}\boldsymbol{\psi}'h] \geq E[\boldsymbol{\psi}\boldsymbol{\phi}'] E\left[\frac{\boldsymbol{\phi}\boldsymbol{\phi}'}{h}\right]^{-1} E[\boldsymbol{\psi}\boldsymbol{\phi}']'. \quad (5.1)$$

The equality holds if and only if there exists a constant $r \times t$ matrix \mathbf{C} such that

$$h\boldsymbol{\psi} + \mathbf{C}\boldsymbol{\phi} = \mathbf{0} \quad \text{a.e.} \quad (5.2)$$

Now we proceed to prove Theorem 3.7.

Proof of Theorem 3.7. Let $\boldsymbol{\psi} = (\partial/\partial\boldsymbol{\theta})F_{\theta_0}$, $\boldsymbol{\phi} = (\partial/\partial\boldsymbol{\theta})F_{\theta_0}$, and $h = H_{\theta_0}^2$, then from the definitions of matrices \mathbf{U} , \mathbf{W} and \mathbf{V} , it can be represented as

$$\begin{aligned}\mathbf{U} &= E[\boldsymbol{\psi}\boldsymbol{\phi}'] , \\ \mathbf{W} &= E[\boldsymbol{\psi}\boldsymbol{\psi}'h], \\ \mathbf{V} &= E\left[\frac{\boldsymbol{\phi}\boldsymbol{\phi}'}{h}\right].\end{aligned}\tag{5.3}$$

Hence from Lemma 5.1, we can see that

$$\mathbf{W} \geq \mathbf{U}\mathbf{V}^{-1}\mathbf{U}.\tag{5.4}$$

From this inequality, we can see that

$$\mathbf{U}^{-1}\mathbf{W}\mathbf{U}^{-1} \geq \mathbf{V}^{-1}.\tag{5.5}$$

□

Proof of Theorem 3.9. Fisher information matrix of CHARN model based on LAN Γ can be represented as

$$\begin{aligned}\Gamma &= E\left[\frac{1}{H_{\theta_0}^2}\left(-\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}}, \frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}}\right)\begin{pmatrix} a & c \\ c & b \end{pmatrix}\begin{pmatrix} -\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}'} \\ \frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}'} \end{pmatrix}\right] \\ &= E\left[\frac{1}{H_{\theta_0}^2}\left(a\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}'} - c\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}'} - c\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}'} + b\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}'}\right)\right] \\ &= E\left[\frac{1}{H_{\theta_0}^2}\left(a\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}'} - c\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}'} - c\frac{\partial H_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}'}\right)\right] \\ &\quad + E\left[\left(\frac{f'(u_t)}{f(u_t)}\right)^2\right]E\left[\frac{1}{H_{\theta_0}^2}\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}'}\right].\end{aligned}\tag{5.6}$$

From (5.6), if $\partial H_{\theta_0}/\partial\boldsymbol{\theta} = \mathbf{0}$, Γ becomes

$$\Gamma = E\left[\left(\frac{f'(u_t)}{f(u_t)}\right)^2\right]E\left[\frac{1}{H_{\theta_0}^2}\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}}\frac{\partial F_{\theta_0}}{\partial\boldsymbol{\theta}'}\right].\tag{5.7}$$

Next, we show under the Gaussianity of u_t that $E[(f'(u_t)/f(u_t))^2] = 1$. From the Schwarz inequality, it can be obtained that

$$\begin{aligned}
 E\left[\left(\frac{f'(u_t)}{f(u_t)}\right)^2\right] &= E[u_t^2]E\left[\left(\frac{f'(u_t)}{f(u_t)}\right)^2\right] \\
 &\geq \left(E\left[u_t \frac{f'(u_t)}{f(u_t)}\right]\right)^2 \\
 &= \left(\int_{-\infty}^{\infty} x f'(x) dx\right)^2 \\
 &= \left(\int_{-\infty}^{\infty} x f(x) dx - \int_{-\infty}^{\infty} f(x) dx\right)^2 \\
 &= 1.
 \end{aligned} \tag{5.8}$$

The equality holds if and only if there exists some constant c such that

$$cx = \frac{f'(x)}{f(x)}. \tag{5.9}$$

Equation (5.9) becomes, for some constant k ,

$$\begin{aligned}
 cx &= (\log f(x))', \\
 \frac{c}{2}x^2 + k &= \log f(x), \\
 f(x) &= e^{(c/2)x^2+k} = e^k e^{(c/2)x^2}.
 \end{aligned} \tag{5.10}$$

Hence, c is -1 , and $f(x)$ becomes the density function of the normal distribution. \square

Acknowledgment

The author would like to thank referees for their comments, which improved the original version of this paper.

References

- [1] T. Amano and M. Taniguchi, "Asymptotic efficiency of conditional least squares estimators for ARCH models," *Statistics & Probability Letters*, vol. 78, no. 2, pp. 179–185, 2008.
- [2] V. P. Godambe, "An optimum property of regular maximum likelihood estimation," *Annals of Mathematical Statistics*, vol. 31, pp. 1208–1211, 1960.
- [3] V. P. Godambe, "The foundations of finite sample estimation in stochastic processes," *Biometrika*, vol. 72, no. 2, pp. 419–428, 1985.
- [4] L. P. Hansen, "Large sample properties of generalized method of moments estimators," *Econometrica*, vol. 50, no. 4, pp. 1029–1054, 1982.

- [5] S. A. Chandra and M. Taniguchi, "Estimating functions for nonlinear time series models," *Annals of the Institute of Statistical Mathematics*, vol. 53, no. 1, pp. 125–141, 2001.
- [6] T. Amano, "Asymptotic efficiency of estimating function estimators for nonlinear time series models," *Journal of the Japan Statistical Society*, vol. 39, no. 2, pp. 209–231, 2009.
- [7] W. Härdle and A. Tsybakov, "Local polynomial estimators of the volatility function in nonparametric autoregression," *Journal of Econometrics*, vol. 81, no. 1, pp. 223–242, 1997.
- [8] W. Härdle, A. Tsybakov, and L. Yang, "Nonparametric vector autoregression," *Journal of Statistical Planning and Inference*, vol. 68, no. 2, pp. 221–245, 1998.
- [9] H. Kanai, H. Ogata, and M. Taniguchi, "Estimating function approach for CHARN models," *Metron*, vol. 68, pp. 1–21, 2010.
- [10] D. Tjøstheim, "Estimation in nonlinear time series models," *Stochastic Processes and Their Applications*, vol. 21, no. 2, pp. 251–273, 1986.
- [11] Z. Lu and Z. Jiang, " L_1 geometric ergodicity of a multivariate nonlinear AR model with an ARCH term," *Statistics & Probability Letters*, vol. 51, no. 2, pp. 121–130, 2001.
- [12] A. S. Kholevo, "On estimates of regression coefficients," *Theory of Probability and its Applications*, vol. 14, pp. 79–104, 1969.
- [13] L. Le Cam, "Locally asymptotically normal families of distributions. Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses," vol. 3, pp. 37–98, 1960.
- [14] H. Kato, M. Taniguchi, and M. Honda, "Statistical analysis for multiplicatively modulated nonlinear autoregressive model and its applications to electrophysiological signal analysis in humans," *IEEE Transactions on Signal Processing*, vol. 54, pp. 3414–3425, 2006.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

