RATE OF WEAK CONVERGENCE FOR HUBER-DUTTER ESTIMATORS IN A LINEAR MODEL WITH FCA PROCESSES

XU LIFENG and ZHANG SHAOYI

School of Mathematics and Computer Science
Hubei University
Wuhan 430062
P. R. China
e-mail: xlf1106@163.com

Abstract

Consider the following linear regression model 
\[ y = X^T \beta + e, \]
where the error \( e \) is a functional coefficient autoregressive (FCA) processes. In this paper, we prove that the Huber-Dutter (HD) estimators for unknown parameters in the above model week converge to the true values with rate \( n^{-\frac{1}{2}} \).

1. Introduction

Consider the following regression model:
\[ y_t = x_t^T \beta + e_t, \quad t = 1, 2, \cdots, n, \]  
(1.1a)
where \( y_t \in \mathbb{R}, x_t = (x_{t1}, \cdots, x_{td})^T \in \mathbb{R}^d, \beta \) is a \( d \)-dimensional unknown parameter, and \( e_t \)s are functional coefficient autoregressive processes given as

2010 Mathematics Subject Classification: 62F11, 62F12.
Keywords and phrases: linear regression model, functional coefficient autoregressive process, Huber-Dutter estimators, martingale difference.
Supported by Hubei Provincial Department of Education (No: Q20122202).
Received December 16, 2013

© 2014 Scientific Advances Publishers
where $\eta_i$s are non-degenerate iid random errors with zero mean and finite variance $\sigma^2$, $\theta$ is a one-dimensional unknown parameter and $f_i(\theta)$ is a real valued function defined on a compact set $\Theta$, which contains the true value $\theta_0$ as an inner point.

Model (1.1) includes many special models, such as an ordinary linear regression model (when $f_i(\theta) \equiv 0$), a linear regressive processes (when $f_i(\theta) = 0$, see Maller [7], Pere [8]), time-dependent autoregressive processes (when $f_i(\theta) = a_t$, see Carsoule and Franses [3], Azrak and Melard [2]), $e_t$ is the same as (1.1b), see Hu [5], Kwoun and Yajima [6].

A number of robust estimators have been proposed and investigated over the last 40 years. One of them is the Huber-Dutter (HD) estimators, see, for example, Silvapulle [10], Douglas [4], Tong et al. [12]. This paper will investigate the weak convergence rate of HD estimators for unknown parameters in the model (1.1) under some regularity conditions.

2. Estimation Method

Let $\lambda_0 = \{\beta_0^T, \sigma_0, \theta_0\}$ be the “true” value of unknown parameter. Let $f_i(\theta_0) = \frac{df_i(\theta)}{d\theta} |_{\theta=\theta_0} \neq 0$ and define $\prod_{i=0}^{t-1} f_{t-i}(\theta_0) = 1$, then from (1.1b),

$$e_t = \sum_{j=0}^{t-1} \prod_{i=0}^{j-1} f_{t-i}(\theta_0) \eta_{t-j}, \quad E e_t = 0, \quad V a r e_t = \sigma_0^2 \sum_{j=0}^{t-1} \prod_{i=0}^{j-1} f_{t-i}^2(\theta_0), \quad (2.1)$$

$$y_t = x_t^T \beta + f_t(\theta)(y_{t-1} - x_{t-1}^T \beta) + \eta_t, \quad y_0 = x_0 = 0. \quad (2.2)$$

Write the Huber-Dutter loss function by

$$Q(\beta, \sigma, \theta) = \sum_{t=1}^{n} \rho \left( \frac{y_t - x_t^T \beta - f_t(\theta)(y_{t-1} - x_{t-1}^T \beta)}{\sigma} \right) \sigma + A_{\rho} \sigma, \quad (2.3)$$
where $\rho \geq 0$ is convex, $\rho(0) = 0$, $\rho(t)/|t| \to k$ as $|t| \to \infty$ for some $k > 0$, and $\{A_n\}$ is a suitably choose sequence of constants. We will obtain the HD estimators for $\lambda = (\beta^T, \alpha, 0)$ by minimizing $Q(\beta, \sigma, 0)$ and we denote it by $\hat{\lambda}_n = (\hat{\beta}_n^T, \hat{\sigma}_n, \hat{\theta}_n)$. Then the estimators, if they exist (see the following Proposition 2.1), will satisfy

\[
\sum_{t=1}^{n} \psi \left( \frac{\hat{e}_t - f_t(\hat{\theta}_n)\hat{e}_{t-1}}{\hat{\sigma}_n} \right)(x_t - f_t(\hat{\theta}_n)x_{t-1}) = 0, \tag{2.4}
\]

\[
\sum_{t=1}^{n} \chi \left( \frac{\hat{e}_t - f_t(\hat{\theta}_n)\hat{e}_{t-1}}{\hat{\sigma}_n} \right) = A_n, \tag{2.5}
\]

\[
\sum_{t=1}^{n} \psi \left( \frac{\hat{e}_t - f_t(\hat{\theta}_n)\hat{e}_{t-1}}{\hat{\sigma}_n} \right)f_t'(\hat{\theta}_n)\hat{e}_{t-1} = 0, \tag{2.6}
\]

where $\psi = \rho'$, $\chi(u) = uv(u) - \rho(u) = \int_0^u x\psi(x)$, and $\hat{e}_t = \gamma_t - x_t^T\hat{\beta}_n$.

Same as in Silvapulle [10] and Tong et al. [12], in what follows, it will be assumed that $n \geq d + 2$ and $A_n > 0$.

Without loss of generality, we may assume $k = 1$. Therefore, $\psi$ is bounded and increases from $-1$ to $+1$. It will also assumed that $\chi$ is bounded. From Proposition 1 in [10], we have the following:

**Proposition 1.** Suppose that $\rho'$ is continuous and for some $A > 0$, $v < 1 - AV^{-1}$, where $v$ is the largest jump in the error distribution, $V = \chi(-\infty) \land \chi(\infty)$, then the equation $E\{\psi\left( \frac{e-U}{\sigma} \right), \chi\left( \frac{e-U}{\sigma} \right) - A \} = 0$ has a solution $(\mu(A), \sigma(A))$ with $\sigma(A) > 0$. Especially, when $A = \lim_{n \to \infty} n^{-1}A_n$, where $A_n$ are defined in (2.3). We denote it by $(\mu_0, \sigma_0)$ with $\sigma_0 > 0$. 

HD estimators include some existed estimators, for example, the least squares estimators \( \rho(u) = u^2, A_n = 0 \) and \( \sigma = 1 \), the least absolute deviation estimators \( \rho(u) = |u|, A_n = 0 \), \( M \)-estimators (cf. Tong et al. [12]). If we put \( \rho(u) = |u|^q (0 < q \neq 1), A_n = 0 \), and \( \sigma \) is a known constant, then (2.4)-(2.6) are the \( L_q \) estimate equations for a linear regression model with \( q \)-norm distribution errors (Arcones [1], Rouner [9]).

3. Main Result and Preliminary Lemmas

To obtain the result in this paper, the following conditions are sufficient (\( \| \cdot \| \) denote the Euclidean norm):

\[(A_1) \max_{1 \leq t \leq n} \| x_t \| = o(n^{1/2}) \text{ (as } n \to \infty).\]

\[(A_2) \text{ There is a constant } \alpha > 0 \text{ such that for any } t \text{ and } \theta \in \Theta,
\sum_{j=1}^{t} \left( \prod_{i=0}^{j-1} f_{t-1}^2(0) \right) \leq \alpha, \quad \sum_{j=t+1}^{n} \left| \prod_{i=0}^{j-1} f_{t-1}^2(0) \right| \leq \alpha.\]

\[(A_3) f_t(\theta), f_t'(\theta), f_t''(\theta) \text{ are bounded for any } t \text{ and } \theta \in \Theta.\]

\[(A_4) \lim_{n \to \infty} \frac{1}{n} A_n - A = 0.\]

\[(A_5) E\psi\left( \frac{\eta_1}{\sigma} \right) = 0 \text{ for any } \sigma > 0, b = E\psi'(\frac{\eta_1}{\sigma_0}) > 0, c = E[\psi''\left( \frac{\eta_1}{\sigma_0} \right) \cdot \frac{\eta_1}{\sigma_0}], \]

\[r = E[\psi'(\frac{\eta_1}{\sigma_0}) \cdot \eta_1^2], br \geq c^2 \quad \text{and} \quad \text{Var}(\psi(\frac{\eta_1}{\sigma})) < \infty, \text{Var}(\psi''(\frac{\eta_1}{\sigma})) < \infty, \]

\[\text{Var}(\psi'(\frac{\eta_1}{\sigma})) < \infty, \text{Var}(\psi'(\frac{\eta_1}{\sigma})\eta_1^2) < \infty, E\eta_1^4 < \infty.\]
Remark. The condition (A1) is weaker than the corresponding condition in Silvapulle [10], Hu [5]. The condition (A2) is used by Hu [5], Song et al. [11] and the first boundlessness condition of (A2) is used by Kwoun and Yajima [6]. The boundlessness of $f_l(\theta), f'_l(\theta), f''_l(\theta)$ in (A3) are used by [5], [11]. And the conditions (A4) and (A5) are used by Silvapulle [10], Tong et al. [12].

**Theorem 3.1.** Suppose that conditions (A1)-(A5) hold. Then

$$\sqrt{n}(\lambda_n - \lambda_0)$$ is $O_p(1)$ as $n \to \infty$. \hspace{1cm} (3.1)

For the proof of Theorem 3.1, we need the lemma as following:

**Lemma 3.1.** Under the conditions (A1)-(A5), we have

(i) $\Delta_n(\lambda_0) = \sum_{t=1}^{n} f_l^2(\theta)Ee_{t-1}^2 \to \infty; \quad \Delta_n(\lambda_0) = O(n)$ as $n \to \infty$.

(ii) $Ee_t^4 \leq (En_1^4 + 6\sigma_0^4)\eta^2$.

(iii) $\text{Var}(\sum_{t=1}^{n} e_t^2) = O(n) (n \to \infty)$.

**Proof.** The proof of (i), see Hu [5]. We only prove (ii) and (iii),

$$Ee_t^4 = E\left\{\sum_{j=0}^{t-1} \left(\prod_{r=0}^{j-1} f_l(\theta_0)\right)\eta_{t-j}\right\}^4$$

$$= E\left\{\sum_{j=0}^{t-1} \left(\prod_{r=0}^{j-1} f_l(\theta_0)\right)\eta_{t-j}\right\}^2 \left(\sum_{k=0}^{t-1} \left(\prod_{r=0}^{k-1} f_l(\theta_0)\right)\eta_{t-k}\right)^2$$

$$= E\left\{\sum_{j=0}^{t-1} \left(\prod_{r=0}^{j-1} f_l(\theta_0)\right)^2 \eta_{t-j}^2\right\} \left(\sum_{k=0}^{t-1} \left(\prod_{r=0}^{k-1} f_l(\theta_0)\right)^2 \eta_{t-k}^2\right)$$
+ 4E \left\{ \sum_{0 \leq i < j \leq t-1} (\prod_{r=0}^{j-1} f_{t-r}(\theta_0))\eta_{t-j} \cdot (\prod_{r=0}^{i-1} f_{t-r}(\theta_0))\eta_{t-i} \right\} \\
\times \left\{ \sum_{0 \leq l < m \leq t-1} (\prod_{r=0}^{l-1} f_{t-r}(\theta_0))\eta_{t-l} \cdot (\prod_{r=0}^{m-1} f_{t-r}(\theta_0))\eta_{t-m} \right\} \\
:= I_1 + 4I_2. \quad (3.2)

I_1 = E \left\{ \sum_{j=0}^{t-1} \prod_{r=0}^{j-1} f_{t-r}(\theta_0)\eta_{t-j} \right\}^4 \\
+ 2E \left\{ \sum_{m=0}^{t-2} \sum_{k=m+1}^{t-1} (\prod_{r=0}^{m-1} f_{t-r}(\theta_0)\eta_{t-m})^2 (\prod_{r=0}^{k-1} f_{t-r}(\theta_0)) \right\}^2 \\
= En_1^4 \left\{ \sum_{j=0}^{t-1} \prod_{r=0}^{j-1} f_{t-r}(\theta_0) \right\}^4 \\
+ 2\sigma_0^4 \sum_{m=0}^{t-2} \sum_{k=m+1}^{t-1} (\prod_{r=0}^{m-1} f_{t-r}(\theta_0))^2 (\prod_{r=0}^{k-1} f_{t-r}(\theta_0))^2 \\
= En_1^4 \left\{ \sum_{j=0}^{t-1} \prod_{r=0}^{j-1} f_{t-r}^2(\theta_0) \right\} \left[ \sum_{j=0}^{t-1} \prod_{r=0}^{j-1} f_{t-r}^2(\theta_0) \right] \\
+ 2\sigma_0^4 \left\{ \sum_{m=0}^{t-2} \prod_{r=0}^{m-1} f_{t-r}^2(\theta_0) \cdot \sum_{k=0}^{t-1} \prod_{r=0}^{k-1} f_{t-r}^2(\theta_0) \right\} \\
\leq (En_1^4 + 2\sigma_0^4)\alpha^2. \quad (3.3)

Similarly, from the condition (A_2), we can obtain that

\[ I_2 = E \left\{ \sum_{0 \leq i < j \leq t-1} (\prod_{r=0}^{j-1} f_{t-r}(\theta_0))^2 \eta_{t-j}^2 (\prod_{r=0}^{i-1} f_{t-r}(\theta_0))^2 \eta_{t-i}^2 \right\} \]
\[
\begin{align*}
&= \sum_{t=1}^{n} \sum_{i=1}^{t-1} \left( \prod_{r=0}^{j-1} f_{i-r}(\theta_0) \right)^2 \left( \prod_{r=0}^{j-1} f_{i-r}(\theta_0) \right)^2 E^2 \eta_i^2 E \eta_{i-j}^2 \\
&\leq \sigma_0^4 \alpha^2. 
\end{align*}
\]

(3.4)

Hence, we get (ii) from (3.2)-(3.4).

Now we prove (iii)

\[
E(\sum_{t=1}^{n} e_i^2)^2 - (\sum_{t=1}^{n} E e_i^2)^2 = \left[ \sum_{t=1}^{n} E e_i^4 + 2 \sum_{1 \leq i < j \leq n} E e_i^2 E e_j^2 \right] \\
- \left[ \sum_{t=1}^{n} (E e_i^2)^2 + 2 \sum_{1 \leq i < j \leq n} E e_i^2 E e_j^2 \right] \\
= \sum_{t=1}^{n} [E e_i^4 - (E e_i^2)^2] + 2 \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} [E e_i^2 E e_{i+k} - E e_i^2 E e_{i+k}']. 
\]

(3.5)

From (ii), we have

\[
\sum_{t=1}^{n} [E e_i^4 - (E e_i^2)^2] = O(n). 
\]

(3.6)

By recursive computation,

\[
E e_i^2 E e_{i+1}^2 = E[e_i^2(f_{i+1}(\theta_0) e_i + \eta_{i+1})^2] = f_{i+1}^2(\theta_0) E e_i^4 + \sigma_0^2 E e_i^2, 
\]

\[
E e_i^2 E e_{i+2}^2 = E[e_i^2(f_{i+2}(\theta_0) e_i + \eta_{i+2})^2] = f_{i+2}^2(\theta_0) E e_i^4 + \sigma_0^2 E e_i^2 \\
= f_{i+2}^2(\theta_0) [f_{i+1}^2(\theta_0) E e_i^4 + \sigma_0^2 E e_i^2] + \sigma_0^2 E e_i^2 \\
= f_{i+2}^2(\theta_0) f_{i+1}^2(\theta_0) E e_i^4 + f_{i+2}^2(\theta_0) \sigma_0^2 E e_i^2 + \sigma_0^2 E e_i^2. 
\]
\[
Ee_i^2 e_i^2 = E[e_i^2 (f_i+3(\theta_0) e_{i+2} + \eta_{i+3})^2] = f_i^2(\theta_0) Ee_i^2 e_{i+2} + \sigma_0^2 Ee_i^2 \\
= f_i^2(\theta_0) [f_{i+2}^2(\theta_0) f_i^2(\theta_0) Ee_i^4 + f_i^2(\theta_0) \sigma_0^2 Ee_i^2 + \sigma_0^2 Ee_i^4] + \sigma_0^2 Ee_i^2 \\
= f_i^2(\theta_0) [f_{i+2}^2(\theta_0) f_{i+1}^2(\theta_0) Ee_i^4 + f_i^2(\theta_0) \sigma_0^2 Ee_i^2 + \sigma_0^2 Ee_{i-1}^2] + \sigma_0^2 Ee_i^2 \\
+ f_i^2(\theta_0) \sigma_0^2 Ee_i^2 + \sigma_0^2 Ee_{i+2}^2.
\]

\[
(3.7)
\]

\[
Ee_i^2 e_{i+k} = E[e_i^2 (f_{i+k}(\theta_0) e_{i+k-1} + \eta_{i+k})^2] \\
= f_{i+k}^2(\theta_0) f_{i+k-1}^2(\theta_0) \cdots f_{i+1}^2(\theta_0) Ee_i^4 + f_{i+k}^2(\theta_0) \sigma_0^2 Ee_i^2 + f_{i+k}(\theta_0) \sigma_0^2 Ee_{i+k-1}^2 \\
\cdots f_{i+2}^2(\theta_0) \sigma_0^2 Ee_i^2 + f_{i+k}(\theta_0) \sigma_0^2 Ee_{i+k-2}^2, \\
\cdots + f_{i+k}^2(\theta_0) \sigma_0^2 Ee_i^2 + \sigma_0^2 Ee_i^2.
\]

\[
Ee_{i+1}^2 = E[f_{i+1}(\theta_0) e_i + \eta_{i+1}]^2 = f_{i+1}^2(\theta_0) Ee_i^2 + \sigma_0^2, \\
Ee_{i+2}^2 = E[f_{i+2}(\theta_0) e_{i+1} + \eta_{i+2}]^2 = f_{i+2}^2(\theta_0) Ee_{i+1}^2 + \sigma_0^2 \\
= f_{i+2}^2(\theta_0) [f_{i+1}^2(\theta_0) Ee_i^2 + \sigma_0^2] + \sigma_0^2 \\
= f_{i+2}^2(\theta_0) f_{i+1}^2(\theta_0) Ee_i^2 + f_{i+2}^2(\theta_0) \sigma_0^2 + \sigma_0^2, \\
Ee_{i+3}^2 = E[f_{i+3}(\theta_0) e_{i+2} + \eta_{i+3}]^2 = f_{i+3}^2(\theta_0) Ee_{i+2}^2 + \sigma_0^2 \\
= f_{i+3}^2(\theta_0) [f_{i+2}^2(\theta_0) f_{i+1}^2(\theta_0) Ee_i^2 + f_{i+2}^2(\theta_0) \sigma_0^2] + \sigma_0^2 \\
= f_{i+3}^2(\theta_0) f_{i+2}^2(\theta_0) f_{i+1}^2(\theta_0) Ee_i^2 + f_{i+3}^2(\theta_0) f_{i+2}^2(\theta_0) \sigma_0^2 + f_{i+3}^2(\theta_0) \sigma_0^2 + \sigma_0^2, \\
\cdots + f_{i+k}^2(\theta_0) \sigma_0^2 + \sigma_0^2. \\
Ee_{i+k}^2 = f_{i+k}^2(\theta_0) f_{i+k-1}^2(\theta_0) \cdots f_{i+1}^2(\theta_0) Ee_i^2 + f_{i+k}^2(\theta_0) f_{i+k-1}^2(\theta_0) \cdots f_{i+2}^2(\theta_0) \sigma_0^2 \\
\cdots + f_{i+k}^2(\theta_0) \sigma_0^2 + \sigma_0^2. \\
(3.8)
\]
Then,
\[
\sum_{i=1}^{n-1} \sum_{k=1}^{n-i} [Ee_i^2r_i^2 - Ee_i^2e_i^2]
\]
\[
= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} [f_{i+k}^2(\theta_0)f_{i+k-1}^2(\theta_0) \cdots f_{i+1}^2(\theta_0)] (Ee_i^4 - Ee_i^2 \cdot Ee_i^2)
\]
\[
= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \prod_{r=1}^{k} f_{i+r}^2(\theta_0) (Ee_i^4 - Ee_i^2 \cdot Ee_i^2) \leq \alpha^2 \cdot \sum_{i=1}^{n-1} (Ee_i^4 - Ee_i^2 \cdot Ee_i^2) = O(n).
\]

(3.9)

From (3.5), (3.6), and (3.9), we obtain (iii).

4. Proof of Theorem 2.1

Let \( F_n(\lambda) \) be the Hessian of \( Q(\lambda) \), that is,

\[
F_n(\lambda) = \frac{\partial^2 Q}{\partial \lambda^T \partial \lambda} = \begin{bmatrix}
\frac{\partial^2 Q}{\partial \beta^T \partial \beta} & \frac{\partial^2 Q}{\partial \beta^T \partial \sigma} & \frac{\partial^2 Q}{\partial \beta^T \partial \theta} \\
* & \frac{\partial^2 Q}{\partial \sigma^2} & \frac{\partial^2 Q}{\partial \sigma \partial \theta} \\
* & * & \frac{\partial^2 Q}{\partial \theta^2}
\end{bmatrix},
\]

(4.1)

where * indicates that the elements are filled in by symmetry.

Let

\[
R_{nl}(\lambda) := \frac{\partial}{\partial \lambda_i} F_n(\lambda) = \begin{bmatrix}
\frac{\partial^3 Q}{\partial \lambda_i \partial \beta^T \partial \beta} & \frac{\partial^3 Q}{\partial \lambda_i \partial \beta^T \partial \sigma} & \frac{\partial^3 Q}{\partial \lambda_i \partial \beta^T \partial \theta} \\
* & \frac{\partial^3 Q}{\partial \lambda_i \partial \sigma^2} & \frac{\partial^3 Q}{\partial \lambda_i \partial \sigma \partial \theta} \\
* & * & \frac{\partial^3 Q}{\partial \lambda_i \partial \theta^2}
\end{bmatrix},
\]

(4.2)

where \( \lambda_i = \beta_i (i = 1, 2, \cdots, d) \), \( \lambda_{d+l} = \sigma \), \( \lambda_{d+2} = \theta \).
First, we prove that

\[ R_{nl}(\lambda) = o_p\left(n^{\frac{3}{2}}\right), \quad l = 1, 2, \ldots, d + 2 \quad (n \to \infty); \quad (4.3) \]

\[
\frac{\partial^3 Q}{\partial \beta_1^T \partial \beta_2^T \partial \beta_3} = -\frac{1}{\sigma^2} \sum_{t=1}^{n} \psi'(\frac{\eta_t}{\sigma}) \left(x_{t,l} - f_t(\theta)x_{t-1,l}\right) \left(x_t - f_t(\theta)x_{t-1}\right) \left(x_t - f_t(\theta)x_{t-1}\right)^T
\]

\[ =: L_n^{(1)}; \]

\[ |L_n^{(1)}| \leq \frac{1}{\sigma^2} \max_{1 \leq t \leq n} \|x_t - f_t(\theta)x_{t-1}\| \sum_{t=1}^{n} |\psi'(\frac{\eta_t}{\sigma})| = o(n) \cdot \sum_{t=1}^{n} |\psi'(\frac{\eta_t}{\sigma})|. \]

Since \( \{\psi'(\frac{\eta_t}{\sigma})\} \) are iid variables,

\[ E\left[\sum_{t=1}^{n} |\psi'(\frac{\eta_t}{\sigma})|\right] = \sum_{t=1}^{n} E|\psi'(\frac{\eta_t}{\sigma})|^2 = O(n). \]

By Chebyshev inequality, \( \sum_{t=1}^{n} \psi'(\frac{\eta_t}{\sigma}) = O_p\left(n^{\frac{1}{2}}\right). \) Then,

\[ L_n^{(1)} = o_p\left(n^{\frac{3}{2}}\right). \quad (4.4) \]

Similarly,

\[ L_n^{(2)} = \frac{\partial^3 Q}{\partial \beta_1^T \partial \beta_2^T \partial \sigma} = -\frac{1}{\sigma^2} \sum_{t=1}^{n} \left\{ \psi'(\frac{\eta_t}{\sigma}) \left(x_{t,l} - f_t(\theta)x_{t-1,l}\right) \left(x_t - f_t(\theta)x_{t-1}\right) \right\}
\]

\[ = o_p\left(n^{\frac{3}{2}}\right), \quad (4.5) \]

\[ L_n^{(3)} = \frac{\partial^3 Q}{\partial \beta_1 \partial \sigma^2} = -\frac{1}{\sigma^2} \sum_{t=1}^{n} \left\{ \psi'(\frac{\eta_t}{\sigma}) \left(x_{t,l} - f_t(\theta)x_{t-1,l}\right)^2 \right\}
\]

\[ - \frac{1}{\sigma^2} \sum_{t=1}^{n} \psi'(\frac{\eta_t}{\sigma}) \eta_t \left(x_{t,l} - f_t(\theta)x_{t-1,l}\right) = o_p\left(n^{\frac{3}{2}}\right), \quad (4.6) \]
\[ L_n^{(4)} = \frac{\delta^5 Q}{\partial \beta_t \partial \beta'} \frac{\partial \beta_t}{\partial \theta} = -\frac{1}{\sigma^2} \sum_{t=1}^{n} f_t'(\theta)\psi'\left(\frac{\eta_t}{\sigma}\right) e_{t-1}(x_{t,l} - f_t(\theta)x_{t-1,l}) (x_t - f_t(\theta)x_{t-1})^T \]

\[ -\frac{1}{\sigma} \sum_{t=1}^{n} f_t'(\theta)\psi'\left(\frac{\eta_t}{\sigma}\right) (x_{t,l} - f_t(\theta)x_{t-1,l}) x_{t-1}^T \]

\[ -\frac{1}{\sigma^2} \sum_{t=1}^{n} f_t'(\theta)\psi'\left(\frac{\eta_t}{\sigma}\right) x_t (x_t - f_t(\theta)x_{t-1})^T \]

\[ := L_{n1}^{(4)} - L_{n2}^{(4)} - L_{n3}^{(4)}. \quad (4.7) \]

Similar to (4.5) and (4.6),

\[ L_{n2}^{(4)} = o_p(n^{-\frac{3}{2}}), \quad L_{n3}^{(4)} = o_p(n^{-\frac{3}{2}}). \quad (4.8) \]

\[ L_{n1}^{(4)} = -\frac{1}{\sigma^2} \sum_{t=1}^{n} f_t'(\theta)\psi'\left(\frac{\eta_t}{\sigma}\right) e_{t-1}(x_{t,l} - f_t(\theta)x_{t-1,l}) (x_t - f_t(\theta)x_{t-1})^T \]

\[ = -\frac{1}{\sigma^2} \sum_{t=1}^{n} f_t'(\theta) \left[ \psi'\left(\frac{\eta_t}{\sigma}\right) - E\psi'\left(\frac{\eta_t}{\sigma}\right) \right] \]

\[ \times e_{t-1}(x_{t,l} - f_t(\theta)x_{t-1,l}) (x_t - f_t(\theta)x_{t-1})^T \]

\[ -\frac{1}{\sigma} \sum_{t=1}^{n} E \left[ \psi'\left(\frac{\eta_t}{\sigma}\right) \right] e_{t-1}(x_{t,l} - f_t(\theta)x_{t-1,l}) (x_t - f_t(\theta)x_{t-1})^T \]

\[ := J_{n1}^{(4)} - J_{n2}^{(4)}. \quad (4.9) \]

Note that \( \{ f_t'(\theta)\left[ \psi'\left(\frac{\eta_t}{\sigma}\right) - E\psi'\left(\frac{\eta_t}{\sigma}\right) \right] e_{t-1} \} \) is a \( F_{t,n} := \sigma(\eta_1, \eta_2, \ldots, \eta_t), \)

\((t \leq n)\) martingale-difference array. Then,
\[E(J_{n1}^{(4)})^2 = -\frac{1}{\sigma^4} \sum_{t=1}^n [f_t'(\theta)]^2 \cdot E \left[ \left( \psi'\left( \frac{\eta}{\sigma} \right) - E\psi'\left( \frac{\eta}{\sigma} \right) \right)^2 e_{t-1}^2 \right] \]

\[\cdot (x_{t,1} - f_t(\theta)x_{t-1,1})^2 (x_t - f_t(\theta)x_{t-1})^T (x_t - f_t(\theta)x_{t-1}) \]

\[\leq \frac{1}{\sigma^4} \max_{1 \leq t \leq n} \|f_t'(\theta)\|^2 \|x_t - f_t(\theta)x_{t-1}\|^4 \cdot \sum_{t=1}^n E \left[ \psi'\left( \frac{\eta_t}{\sigma} \right) \right]^2 Ee_{t-1}^2 \]

\[= \frac{1}{\sigma^4} \max_{1 \leq t \leq n} \|f_t'(\theta)\|^2 \|x_t - f_t(\theta)x_{t-1}\|^4 E[\psi'(\frac{\eta_t}{\sigma})]^2 \cdot \sum_{t=1}^n Ee_{t-1}^2 \]

\[= o_p\left(n^{\frac{4}{3}}\right) \cdot O(n) = o_p(n^{\frac{7}{3}}).\]

By Chebyshev inequality, \(J_{n1}^{(4)} = o_p(n^{\frac{2}{3}})\), then

\[J_{n1}^{(4)} = o\left(n^{\frac{2}{3}}\right). \quad (4.10)\]

\[J_{n2}^{(4)} = \frac{2}{\sigma^4} \text{Var} \left\{ \sum_{t=1}^n e_{t-1}f_t(\theta)(x_{t,1} - f_t(\theta)x_{t-1,1})^2 (x_t - f_t(\theta)x_{t-1})^T \right\} \]

\[= \frac{b^2}{\sigma^4} \text{Var} \left\{ \sum_{t=2}^n \sum_{k=0}^{t-2} \left( \prod_{i=0}^{k-1} f_{t-i}(\theta_0) \right) h_{t-k} \right\} \left[ f_t(\theta)[x_{t,1} - f_t(\theta)x_{t-1,1}] [x_t - f_t(\theta)x_{t-1}]^T \right]. \]

Exchange the order of summing,

\[\text{Var} J_{n2}^{(4)} = \frac{b^2}{\sigma^4} \text{Var} \left\{ \sum_{k=2}^n \sum_{t=0}^{n-k} \left( \prod_{i=0}^{t-1} f_{t+k-i}(\theta) \right) (x_{t+k-1} - f_{t+k}(\theta)x_{t+k-1,1}) (x_{t+k} - f_{t+k}(\theta)x_{t+k-1})^T f_{t+k}'(\theta) \right\} h_k \]

\[= b^2 \sum_{k=2}^n \sum_{t=0}^{n-k} \left( \prod_{i=0}^{t-1} f_{t+k-i}(\theta_0) \right) f_{t+k}'(\theta) \]
\[
\cdot (x_{t+k,l} - f_{t+k}(\theta)x_{t+k-1,l})(x_{t+k} - f_{t+k}(\theta)x_{t+k-1})^T (x_{t+k} - f_{t+k}(\theta)x_{t+k-1}) \left[ 
\right]
\]

\[
\leq b^2 \max_j |f_j(\theta)|^2 \|x_j - f_j(\theta)x_{j-1}\|^2 \sum_{k=2}^{n} \left[ \sum_{t=0}^{n-k} \prod_{i=0}^{t-1} f_{i+k-i}(\theta) \right]^2
\]

\[
\leq b^2 \max_j |f_j(\theta)|^2 \|x_j - f_j(\theta)x_{j-1}\|^2 \left[ \sum_{k=2}^{n} \left[ \sum_{m=k+1}^{n} \prod_{i=0}^{m-k-1} f_{i-m}(\theta) \right] \right]^2
\]

\[
\leq b^2 \max_j |f_j(\theta)|^2 \|x_j - f_j(\theta)x_{j-1}\|^2 \sum_{k=2}^{n} [1 + \alpha]^2 = o\left(\frac{4}{n^3}\right) \cdot O(n) = o\left(\frac{7}{n^3}\right),
\]

\[
J_{n^2}^{(4)} = o_p\left(\frac{7}{n^6}\right). \tag{4.11}
\]

From (4.7), (4.8), (4.9), (4.10), and (4.11), we have

\[
L_n^{(4)} = o_p\left(\frac{3}{n^2}\right). \tag{4.12}
\]

For the estimation of \(\frac{\partial^3 Q}{\partial \beta_l \partial \sigma \partial \theta}, \frac{\partial^3 Q}{\partial \beta_l \partial \sigma \partial \theta^2}, \frac{\partial^3 Q}{\partial \sigma \partial \beta \partial \sigma \partial \beta_l}, \frac{\partial^3 Q}{\partial \sigma \partial \beta \partial \sigma \partial \theta}, \ldots, \frac{\partial^3 Q}{\partial \theta^3},\)

we can use the methods seminary to that in (4.4) or (4.12) to obtain (4.3).

Let \(T_n(\lambda) = n^{-\frac{1}{2}} \frac{\partial Q}{\partial \lambda}, B_n(\lambda) = n^{-1} \frac{\partial^2 Q}{\partial \lambda^T \partial \lambda},\) we will prove that

\[
T_n(\lambda_0 + n^{-\frac{1}{2}}r) = T_n(\lambda_0) + B_n(\lambda_0)r + p_n(r), \tag{4.13}
\]

where \(\sup_{\|P\| \leq K} \|P_n(r)\| \to 0\) as \(n \to \infty,\) for any \(K > 0.\)

In fact, by Taylor’s theorem,

\[
T_n(\lambda_0 + n^{-\frac{1}{2}}r) = T_n(\lambda_0) + B_n(\lambda_0)r + n^{-\frac{3}{2}}R_n(r, s), \tag{4.14}
\]

where \(R_n(r, s)\) has \(l\)-th element.
\[ R_{n,l}(r, s_l) = r^T \left[ R_{nl}(\lambda_0 + n^{-\frac{1}{2}} s_l r) \right] r, \quad S_l \in [0, 1]. \]

For \( \|r\| \leq K \), we have
\[
|R_{n,l}(r, s_l)| \leq K^2 \sum_{a,b=1}^{d+2} \frac{\partial^2}{\partial \lambda_a \partial \lambda_b} \left( \frac{\partial}{\partial \lambda_l} (\bar{f}) \right) \bigg|_{\lambda=\lambda_0+n^{-\frac{1}{2}} s_l}.
\]

Then by (4.3), \( n^{-\frac{3}{2}} R_n(r, s) = o_p(1) \). So (4.13) holds.

From Lemma 2 and Lemma 3(a) in Douglas [4], we can obtain the result of Theorem 2.1 basing on the expansion (4.13).

References


http://dx.doi.org/10.1155/2012/862398