ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Statistics & Probability Letters ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Consistency for the least squares estimator in nonlinear regression model☆

Hu Shuhe

Department of Mathematics, Anhui University, Hefei 230039, China

Received December 2002

Abstract

The consistency problems of the least-squares estimator θ_n for parameter θ in nonlinear regression model are resolved perfectly. Assuming that the tth absolute moments of the model errors are finite, for $t \geqslant 2$ and the errors satisfy general dependent conditions, we obtain the same probability inequality as that in Ivanov (Theory Probab. Appl. 21 (1976) 557) which has independent identically distributed errors; for $1 < t < 2$, we first obtain weak consistency and weak consistency rate of θ_n.

MSC: primary 62J02; secondary 62F12

Keywords: Nonlinear regression model; Least-squares estimator; Dependent error; Consistency; Consistency rate

1. Introduction

For the consistency problem of the least-squares (LS) estimator θ_n for unknown parameter θ in nonlinear regression model, because it has widen statistic applied background, it has been studied by many statisticians since the 1970s. Ivanov (1976) obtained a probability inequality of deviation for θ_n from the true parameter θ when the model errors $\{\varepsilon_n\}$ is independent identically distributed (i.i.d.), $E|\varepsilon_n|^t < \infty$ for some integer $t \geqslant 2$. Prakasa Rao (1984) generalized the result to the case when $\{\varepsilon_n\}$ is φ-mixing or strong mixing sequences, but the assumptions for the moment and the mixing coefficients of $\{\varepsilon_n\}$ are somewhat strict. Recently, Hu (2002) obtained the strong consistency and strong consistency rate of θ_n when the errors satisfy general dependent conditions and $\sup_n E|\varepsilon_n|^t < \infty$ for some $t > 2$. So far as I have known, there are following unsolved problems: (1) Whether θ_n are the consistent estimators when $E(\varepsilon_n^2) = \infty$? (2) Whether the result of Ivanov (1976) still holds true

☆ This paper was supported by the National Natural Science Foundation of China (Grant No: 10271001).
E-mail address: hushuhe@263.net (Hu Shuhe).

doi:10.1016/j.spl.2003.11.020

when $\{\varepsilon_n\}$ is dependent and $\sup_n E|\varepsilon_n|^2 < \infty$? We give a definite answer to the problems by using theory of stochastic process and the model structure.

Consider the nonlinear regression model

$$X_n = g_n(\theta) + \varepsilon_n, \quad n \geqslant 1, \tag{1.1}$$

where $\{g_n(\theta)\}$ is a known sequence of continuous functions possibly nonlinear in $\theta \in \Theta$, a closed interval on the real line, $\{\varepsilon_n\}$ is a zero mean random error. Let

$$Q_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} w_i^2 (X_i - g_i(\theta))^2, \tag{1.2}$$

where $\{w_i\}$ is a known sequence of positive numbers. An estimator θ_n is said to be a LS estimator of θ if it satisfies

$$Q_n(\theta_n) = \inf_{\theta \in \Theta} Q_n(\theta). \tag{1.3}$$

Note that $g_n(\theta)$ are continuous functions defined on compact set Θ, Lemma 3.3 of Schmetterer (1974) shows that there exists a Borel measurable map $\theta_n(x) : R^n \to \Theta$, such that $Q_n(\theta_n(x)) = \inf_{\theta \in \Theta} Q_n(\theta)$. In the following we consider this measurable version as the LS estimator θ_n. Let θ_0 be the true parameter and suppose $\theta_0 \in$ Interior of Θ. Ivanov (1976) obtained the following result when the errors ε_n are i.i.d. and $w_i \equiv 1$ (cf. MacNeill and Umphrey, 1987, p. 144):

Theorem A. *Suppose that ε_n are i.i.d. random variables with $E|\varepsilon_1|^t < \infty$ for some integer $t \geqslant 2$. Further suppose that there exist $0 < k_1 \leqslant k_2 < \infty$ such that*

$$nk_1(\theta_1 - \theta_2)^2 \leqslant \sum_{i=1}^{n} (g_i(\theta_1) - g_i(\theta_2))^2 \leqslant nk_2(\theta_1 - \theta_2)^2, \tag{1.4}$$

for all $\theta_1, \theta_2 \in \Theta$ and for all $n \geqslant 1$. Then there exists a constant $c > 0$ independent of n and ρ such that

$$P(n^{1/2}|\theta_n - \theta_0| > \rho) \leqslant c\rho^{-t}, \tag{1.5}$$

for every $\rho > 0$ and for all $n \geqslant 1$.

Prakasa Rao (1984) generalized the result to the case when ε_n are mixing sequences not necessarily identically distributed.

Theorem B. *Suppose $\{\varepsilon_n\}$ is a φ-mixing sequence satisfying the following condition*:

(A1) $E\varepsilon_n = 0, \quad \sup_n E(\varepsilon_n^4) < \infty$,
(A2) $\sum_{i=1}^{\infty} (i+1)(\varphi(i))^{1/4} < \infty$,
(A3) *there exist constants* $k_1 > 0, k_2 < \infty$ *such that* $k_1(\theta_1 - \theta_2)^2 \leqslant \frac{1}{n}\sum_{i=1}^{n} w_i^2 (g_i(\theta_1) - g_i(\theta_2))^2 \leqslant k_2(\theta_1 - \theta_2)^2$,

for all $n \geqslant 1$ *and* $\theta_1, \theta_2 \in \Theta$,

(A4) $\sup_n\{w_n\} = O(1)$.

Then, there exists a constant $c > 0$ *such that*

$$P(n^{1/2}|\theta_n - \theta_0| > \rho) \leqslant c\rho^{-4}, \tag{1.6}$$

for every $\rho > 0$ *and for all* $n \geqslant 1$.

Suppose that there exists $\delta > 0$ *such that*

(A1)′ $E\varepsilon_n = 0, \quad \sup_n E|\varepsilon_n|^{4+2\delta} < \infty$,
(A2)′ $\sum_{i=1}^{\infty}(i+1)(\alpha(i))^{\delta/(4+\delta)} < \infty$.

Prakasa Rao (1984) pointed that (1.6) still holds if $\{\varepsilon_n\}$ is a strong mixing sequence under the conditions (A1)′, (A2)′, (A3) and (A4).

Considering the applications of the model (1.1), we generalized Theorem A to the case when $\{\varepsilon_n\}$ is dependent. Assuming that the tth absolute moments of ε_n are finite. For $t \geqslant 2$ and the ε_n satisfy general dependent conditions, we obtain the same probability inequality as that in Ivanov (1976) which has i.i.d. errors, and this improves the result of Prakasa Rao (1984). When $E(\varepsilon_n^2) = \infty$, whether there are similar results or not, we have not seen it in the references. For $1 < t < 2$, we first obtain the result which is similar to (1.5) and give the weak consistency and weak consistency rate of θ_n.

In this paper, we assume that $C_1, C_2, \ldots, C, k_1, k_2, \ldots, C_1(t), C_2(t), \ldots, C(t)$ are some positive constants (not necessarily always the same) independent of n, θ, ε.

Lemma 1.1. (cf. Strook and Varadhan, 1979, p. 49). *Let* $(E, \mathscr{F}, P)$ *be a probability space and* $\theta : [0,\infty) \times E \to R^d$ *a* $\mathscr{B}_{[0,\infty} \times \mathscr{F}$ *measurable function such that* $\theta(\cdot, q)$ *is continuous for all* $q \in E$. *If for each* $T > 0$ *there exist numbers* $\alpha = \alpha_T > 0, r = r_T > 0$ *and* $C = C_T < \infty$ *such that*

$$E|\theta(t) - \theta(s)|^r \leqslant C|t-s|^{1+\alpha}, \quad \forall\, 0 \leqslant s, t \leqslant T, \tag{1.7}$$

then for any $\gamma = \gamma_T \in (2, 2+\alpha_T)$ *and* $\lambda > 0$,

$$P\left(\sup_{0 \leqslant s < t \leqslant T} \frac{|\theta(t) - \theta(s)|}{|t-s|^{\beta}} \geqslant \frac{8\gamma}{\gamma - 2}(4\lambda)^{1/r}\right) \leqslant \frac{CA}{\lambda}, \tag{1.8}$$

where

$$\beta = \beta_T = \frac{\gamma_{T-2}}{r_T}, \quad A = A_T = \int_0^T \int_0^T |t-s|^{1+\alpha-\gamma}\, \mathrm{d}s\, \mathrm{d}t.$$

Lemma 1.2. *Let* $(\Omega, \mathscr{F}, P)$ *be a probability space,* $[T_1, T_2]$ *be a closed interval on the real line,* $V(\theta) = V(\omega, \theta)(\theta \in [T_1, T_2], \omega \in \Omega)$ *be a random process such that* $V(\omega, \theta)$ *is continuous for all* $\omega \in \Omega$. *If there exist numbers* $\alpha > 0, r > 0$ *and* $C = C(T_1, T_2) < \infty$ *such that,*

$$E|V(\theta_1) - V(\theta_2)|^r \leqslant C|\theta_1 - \theta_2|^{1+\alpha}, \quad \forall\, \theta_1, \theta_2 \in [T_1, T_2], \tag{1.9}$$

then for any $\varepsilon > 0, a > 0, \theta_0, \theta_0 + \varepsilon \in [T_1, T_2], \gamma \in (2, 2+\alpha)$,

$$P\left(\sup_{\theta_0 \leqslant \theta_1, \theta_2 \leqslant \theta_0+\varepsilon} |V(\theta_1) - V(\theta_2)| \geqslant a\right) \leqslant \frac{8C}{(\alpha-\gamma+2)(\alpha-\gamma+3)}\left(\frac{8\gamma}{\gamma-2}\right)^r \frac{\varepsilon^{\alpha+1}}{a^r}. \tag{1.10}$$

Proof. Consider $V(\theta), \theta \in [\theta_0, \theta_0+\varepsilon] \subset [T_1, T_2]$. we define $\tilde{V}(\theta) = V(\theta+\theta_0), 0 \leqslant \theta \leqslant \varepsilon$, then $\tilde{V}(\theta)$ is still a continuous function of θ for all $\omega \in \Omega$, and (1.9) implies that

$$E|\tilde{V}(\theta_1) - \tilde{V}(\theta_2)|^r \leqslant C|\theta_1 - \theta_2|^{1+\alpha}, \quad \forall\, 0 \leqslant \theta_1, \theta_2 \leqslant \varepsilon. \tag{1.11}$$

Since

$$\sup_{0 \leqslant \theta_1, \theta_2 \leqslant \varepsilon} |\tilde{V}(\theta_1) - \tilde{V}(\theta_2)| = \sup_{0 \leqslant \theta_1 < \theta_2 \leqslant \varepsilon}\left(\frac{|\tilde{V}(\theta_1) - \tilde{V}(\theta_2)|}{|\theta_1-\theta_2|^\beta}|\theta_1-\theta_2|^\beta\right) \leqslant \sup_{0 \leqslant \theta_1 < \theta_2 \leqslant \varepsilon} \frac{|\tilde{V}(\theta_1) - \tilde{V}(\theta_2)|}{|\theta_1-\theta_2|^\beta}\varepsilon^\beta,$$

if we take $T = \varepsilon$, $r_T = r$, $\alpha_T = \alpha$, $\gamma \in (2, 2+\alpha)$, $\beta = (\gamma-2)/r$, $\lambda = \frac{1}{4}(a(\gamma-2)/(8\gamma\varepsilon^\beta))^r$ in Lemma 1.1, then

$$\begin{aligned}
P\left(\sup_{\theta_0 \leqslant \theta_1, \theta_2 \leqslant \theta_0+\varepsilon} |V(\theta_1) - V(\theta_2)| \geqslant a\right) &= P\left(\sup_{\theta_0 \leqslant \theta_1, \theta_2 \leqslant \theta_0+\varepsilon} |\tilde{V}(\theta_1-\theta_0) - \tilde{V}(\theta_2-\theta_0)| \geqslant a\right)\\
&= P\left(\sup_{0 \leqslant \theta_1, \theta_2 \leqslant \varepsilon} |\tilde{V}(\theta_1) - \tilde{V}(\theta_2)| \geqslant a\right) \leqslant P\left(\sup_{0 \leqslant \theta_1 < \theta_2 \leqslant \varepsilon} \frac{|\tilde{V}(\theta_1) - \tilde{V}(\theta_2)|}{|\theta_1-\theta_2|^\beta} \geqslant a\varepsilon^{-\beta}\right)\\
&= P\left(\sup_{0 \leqslant \theta_1 < \theta_2 \leqslant \varepsilon} \frac{|\tilde{V}(\theta_1) - \tilde{V}(\theta_2)|}{|\theta_1-\theta_2|^\beta} \geqslant \frac{8\gamma}{\gamma-2}(4\lambda)^{1/r}\right)\\
&\leqslant \frac{CA}{\lambda} = 4CA\left(\frac{8\gamma}{\gamma-2}\right)^r \frac{\varepsilon^{\gamma-2}}{a^r},
\end{aligned} \tag{1.12}$$

where

$$A = \int_0^\varepsilon\!\!\int_0^\varepsilon |u-s|^{1+\alpha-\gamma}\,\mathrm{d}s\,\mathrm{d}u = \frac{1}{2+\alpha-\gamma}\int_0^\varepsilon (u^{2+\alpha-\gamma} + (\varepsilon-u)^{2+\alpha-\gamma})\,\mathrm{d}u = \frac{2\varepsilon^{\alpha+3-\gamma}}{(\alpha-\gamma+2)(\alpha-\gamma+3)} \tag{1.13}$$

and (1.10) follows from (1.12) and (1.13).

Lemma 1.3. (cf. Hall and Heyde, 1980, p. 23). *If $\{S_i, \mathscr{F}_i, 1 \leqslant i \leqslant n\}$ is a martingale and $1 < p < \infty$, then there exist constants c_1 and c_2 depending only on p such that*

$$c_1 E\left(\sum_{i=1}^{n} X_i^2\right)^{p/2} \leqslant E|S_n|^p \leqslant c_2 E\left(\sum_{i=1}^{n} X_i^2\right)^{p/2}, \tag{1.14}$$

where $X_1 = S_1$ and $X_i = S_i - S_{i-1}, 2 \leqslant i \leqslant n$.

2. Main result

Theorem 2.1. *Consider the nonlinear regression model (1.1), suppose that there exist positive constants c_1, c_2, c_3, c_4, such that*

$$c_1|\theta_1 - \theta_2| \leqslant |g_i(\theta_1) - g_i(\theta_2)| \leqslant c_2|\theta_1 - \theta_2|, \tag{2.1}$$

for all $\theta_1, \theta_2 \in \Theta$ and for all $i \geqslant 1$,

$$c_3 \leqslant w_i \leqslant c_4, \quad \forall\, i \geqslant 1. \tag{2.2}$$

Further suppose that $\{\varepsilon_n\}$ satisfies the conditions: there exists $t \in (1, 2]$ such that $\sup_n E|\varepsilon_n|^t < \infty$, and for any real number sequence $\{c_{ni}\}$ there exists positive constant $C_1(t)$ such that

$$E\left|\sum_{i=1}^{n} c_{ni}\varepsilon_i\right|^t \leqslant C_1(t)\sum_{i=1}^{n}|c_{ni}|^t, \quad \forall\, n \geqslant 1. \tag{2.3}$$

Then, there exists a constant $C(t)$, such that

$$P(n^{1/2}|\theta_n - \theta_0| > \rho) \leqslant C(t)n^{1-t/2}\rho^{-t}, \tag{2.4}$$

for every $\rho > 0$ and for all $n \geqslant 1$.

Proof. Let

$$\psi_n(\theta_1, \theta_2) = \frac{1}{n}\sum_{i=1}^{n} w_i^2(g_i(\theta_1) - g_i(\theta_2))^2, \tag{2.5}$$

$$V_n(\theta) = \frac{1}{n^{1/2}}\sum_{i=1}^{n} \varepsilon_i(g_i(\theta) - g_i(\theta_0)), \tag{2.6}$$

$$U_n(\theta) = \frac{V_n(\theta)}{n^{1/2}\psi_n(\theta, \theta_0)}, \quad \theta \neq \theta_0, \tag{2.7}$$

for simplicity, we will assume that $w_i \equiv 1$, the general case follows from similar arguments in view of (2.2). By (2.1) we have

$$c_1^2(\theta_1 - \theta_2)^2 \leqslant \psi_n(\theta_1, \theta_2) \leqslant c_2^2(\theta_1 - \theta_2)^2, \tag{2.8}$$

for all $\theta_1, \theta_2 \in \Theta$ and for all $n \geqslant 1$. By (2.8) and the proof of Theorem 2.1 of Hu (2002), we know that

$$P(n^{1/2}|\theta_n - \theta_0| > \rho) \leqslant P\left(\sup_{|\theta-\theta_0|>\rho} |U_n(\theta)| \geqslant \frac{1}{2}\right) + P\left(\sup_{\rho n^{-1/2}<|\theta-\theta_0|\leqslant\rho} |U_n(\theta)| \geqslant \frac{1}{2}\right), \tag{2.9}$$

$$P\left(\sup_{|\theta-\theta_0|>\rho} |U_n(\theta)| \geqslant \frac{1}{2}\right) \leqslant P\left(\sup_{|\theta-\theta_0|>\rho} \frac{|V_n(\theta)|}{n^{1/2}\psi_n^{1/2}(\theta,\theta_0)} \geqslant \frac{1}{2}c_1\rho\right). \tag{2.10}$$

Notice that $1 < t \leqslant 2$, thus by C_r inequality, (2.1) and (2.8) we obtain

$$\begin{aligned}
\left|\frac{V_n(\theta)}{n^{1/2}\psi_n^{1/2}(\theta,\theta_0)}\right|^t &= \left|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\left(\frac{g_i(\theta)-g_i(\theta_0)}{\psi_n^{1/2}(\theta,\theta_0)}\right)\right|^t \\
&\leqslant \frac{1}{n^t}n^{t-1}\sum_{i=1}^{n}|\varepsilon_i|^t\frac{|g_i(\theta)-g_i(\theta_0)|^t}{\psi_n^{t/2}(\theta,\theta_0)} \\
&\leqslant \frac{C(t)}{n}\sum_{i=1}^{n}|\varepsilon_i|^t, \ \ \forall\, \theta \neq \theta_0,
\end{aligned} \tag{2.11}$$

therefore by (2.10), (2.11), Markov's inequality and $\sup_n E|\varepsilon_n|^t < \infty$, we get

$$\begin{aligned}
P\left(\sup_{|\theta-\theta_0|>\rho} |U_n(\theta)| \geqslant \frac{1}{2}\right) &\leqslant P\left(\frac{C(t)}{n}\sum_{i=1}^{n}|\varepsilon_i|^t \geqslant \left(\frac{1}{2}c_1\rho\right)^t\right) \\
&\leqslant \left(\frac{2}{c_1\rho}\right)^t\frac{C(t)}{n}\sum_{i=1}^{n}E|\varepsilon_i|^t \leqslant C_1(t)\rho^{-t}.
\end{aligned} \tag{2.12}$$

Let

$$\theta(m) = \theta_0 + \frac{\rho}{n^{1/2}} + \frac{m\rho}{[n^{1/2}]},$$

$$\rho_m = \theta(m) - \theta_0, \text{ for } m = 0, 1, \cdots, [n^{1/2}].$$

Then by (2.8) and the proof of Theorem 2.1 of Hu (2002) we know that

$$P\left(\sup_{\rho n^{-1/2}<\theta-\theta_0\leqslant\rho} |U_n(\theta)| \geqslant \frac{1}{2}\right) \leqslant \sum_{m=0}^{[n^{1/2}]-1} P\left(\sup_{\rho_m\leqslant\theta-\theta_0\leqslant\rho_{m+1}} |V_n(\theta)| \geqslant \frac{1}{2}c_1^2\rho_m^2 n^{1/2}\right), \tag{2.13}$$

$$
\begin{aligned}
P&\left(\sup_{\rho_m \leqslant \theta-\theta_0 \leqslant \rho_{m+1}} |V_n(\theta)| \geqslant \frac{1}{2}\, c_1^2 \rho_m^2 n^{1/2}\right) \\
&\leqslant P\left(|V_n(\theta(m))| \geqslant \frac{1}{4}\, c_1^2 \rho_m^2 n^{1/2}\right) \\
&\quad + P\left(\sup_{\theta(m) \leqslant \theta_1, \theta_2 \leqslant \theta(m+1)} |V_n(\theta_2) - V_n(\theta_1)| \right. \\
&\left. \geqslant \frac{1}{4} c_1^2 \rho_m^2 n^{1/2}\right). \qquad (2.14)
\end{aligned}
$$

Notice that

$$
V_n(\theta(m)) = \frac{1}{n^{1/2}} \sum_{i=1}^{n} \varepsilon_i (g_i(\theta(m)) - g_i(\theta_0)),
$$

$$
V_n(\theta_2) - V_n(\theta_1) = \frac{1}{n^{1/2}} \sum_{i=1}^{n} \varepsilon_i (g_i(\theta_2) - g_i(\theta_1)),
$$

and by Markov's inequality, (2.1) and (2.3), we get

$$
\begin{aligned}
P\left(|V_n(\theta(m))| \geqslant \frac{1}{4}\, c_1^2 \rho_m^2 n^{1/2}\right) &\leqslant \left(\frac{4}{c_1^2 \rho_m^2 n^{1/2}}\right)^t E\left|\frac{1}{n^{1/2}} \sum_{i=1}^{n} \varepsilon_i (g_i(\theta(m)) - g_i(\theta_0))\right|^t \\
&\leqslant C_1(t) \rho_m^{-2t} n^{-t} \sum_{i=1}^{n} |g_i(\theta(m)) - g_i(\theta_0)|^t \\
&\leqslant C_2(t) \rho_m^{-2t} n^{1-t} |\theta(m) - \theta_0|^t = C_2(t) \rho_m^{-t} n^{1-t}, \qquad (2.15)
\end{aligned}
$$

$$
\begin{aligned}
E|V_n(\theta_2) - V_n(\theta_1)|^t &\leqslant C_1(t) n^{-t/2} \sum_{i=1}^{n} |g_i(\theta_2) - g_i(\theta_1)|^t \\
&\leqslant C_2(t) n^{1-t/2} |\theta_2 - \theta_1|^t \triangleq C(n,t) |\theta_2 - \theta_1|^t, \qquad (2.16)
\end{aligned}
$$

for all $\theta_1, \theta_2 \in \Theta$ and for all $n \geqslant 1$. If we take $r = t = 1 + \alpha, C = C(n,t), \varepsilon = \rho/[n^{1/2}], a = \frac{1}{4} c_1^2 \rho_m^2 n^{1/2}$, $\gamma \in (2, t+1)$ in Lemma 1.2, then

$$
\begin{aligned}
P&\left(\sup_{\theta(m) \leqslant \theta_1, \theta_2 \leqslant \theta(m+1)} |V_n(\theta_2) - V_n(\theta_1)| \geqslant \frac{1}{4} c_1^2 \rho_m^2 n^{1/2}\right) \\
&= P\left(\sup_{\theta(m) \leqslant \theta_1, \theta_2 \leqslant \theta(m) + \rho/[n^{1/2}]} |V_n(\theta_2) - V_n(\theta_1)| \geqslant \frac{1}{4} c_1^2 \rho_m^2 n^{1/2}\right)
\end{aligned}
$$

$$\leqslant \frac{8C_2(t)n^{1-t/2}}{(t+1-\gamma)(t+2-\gamma)}\left(\frac{8\gamma}{\gamma-2}\right)^t\left(\frac{\rho}{[n^{1/2}]}\right)^t\left(\frac{4}{c_1^2\rho_m^2 n^{1/2}}\right)^t$$
$$\leqslant C_3(t)\rho^t n^{1-3t/2}\rho_m^{-2t}. \tag{2.17}$$

Notice that $\rho_0=\rho n^{-1/2}$, $\rho_m > m\rho n^{-1/2}$, and combining (2.13)–(2.15), (2.17), we obtain that

$$P\left(\sup_{\rho n^{-1/2}<\theta-\theta_0\leqslant\rho}|U_n(\theta)|\geqslant\frac{1}{2}\right)\leqslant\sum_{m=0}^{[n^{1/2}]-1}\left(C_2(t)\rho_m^{-t}n^{1-t}+C_3(t)\rho^t n^{1-3t/2}\rho_m^{-2t}\right)$$
$$\leqslant \frac{C_2(t)}{n^{t/2-1}\rho^t}+\frac{C_3(t)}{n^{t/2-1}\rho^t}+\frac{1}{n^{t/2-1}\rho^t}\sum_{m=1}^{[n^{1/2}]-1}\left(\frac{C_2(t)}{m^t}+\frac{C_3(t)}{m^{2t}}\right)$$
$$\leqslant C_4(t)n^{1-t/2}\rho^{-t}. \tag{2.18}$$

Similarly we can get

$$P\left(\sup_{\rho n^{-1/2}<\theta_0-\theta\leqslant\rho}|U_n(\theta)|\geqslant\frac{1}{2}\right)\leqslant C_5(t)n^{1-t/2}\rho^{-t}, \tag{2.19}$$

and (2.4) follows from (2.9), (2.12), (2.18) and (2.19). □

3. The application of Theorem 2.1

Using Theorem 2.1, we can get weak consistency and weak consistency rate of θ_n when the ε_n are some dependent sequences.

Theorem 3.1. *Consider the nonlinear regression model* (1.1), *we assume that conditions (2.1) and (2.2) are satisfied and the* $\{\varepsilon_n\}$ *is independent or a martingale difference sequence, there exists* $t\in(1,2]$ *such that* $\sup_n E|\varepsilon_n|^t<\infty$, *then* (2.4) *holds.*

Proof. Since independent random sequence $\{\varepsilon_n\}$ with zero mean is a martingale difference sequence, we only need to prove (2.4) for a martingale difference sequence. Assume that $\{\varepsilon_n\}$ is a martingale difference sequence, then it is easy to see that $\{c_{ni}\varepsilon_i, i=1,2,\cdots,n\}$ is also a martingale difference sequence. Thus by Lemma 1.3, C_r inequality and $\sup_n E|\varepsilon_n|^t<\infty$, we have

$$E\left|\sum_{i=1}^n c_{ni}\varepsilon_i\right|^t\leqslant C(t)E\left(\sum_{i=1}^n (c_{ni}\varepsilon_i)^2\right)^{t/2}$$
$$\leqslant C(t)\sum_{i=1}^n|c_{ni}|^t E|\varepsilon_i|^t\leqslant C_1(t)\sum_{i=1}^n|c_{ni}|^t,\quad \forall\, n\geqslant 1, \tag{3.1}$$

i.e. (2.3) holds, and (2.4) follows from Theorem 2.1. □

When $\sup_n E(\varepsilon_n^2) < \infty$, the condition (2.1) of the model (1.1) can be weakened into (2.8). In fact, we have the following result.

Theorem 3.2. *Consider the nonlinear regression model* (1.1), *we assume that conditions* (2.2) *and* (2.8) *are satisfied,* $\sup_n E(\varepsilon_n^2) < \infty$, *and one of the following four conditions holds*:

(i) $\{\varepsilon_n\}$ *is a independent or a martingale difference sequence*;
(ii) $\{\varepsilon_n\}$ *is a* φ*-mixing sequence with* $\sum_{i=1}^{\infty}(\varphi(i))^{1/2} < \infty$;
(iii) $\{\varepsilon_n\}$ *is a negatively associated (NA) sequence*;
(iv) $\{\varepsilon_n\}$ *is a general weakly stationary linear process*: $\varepsilon_t = \sum_{j=-\infty}^{\infty} \psi_j z_{t-j}$, $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$, $\{z_l, \mathscr{F}_l\}$ *is a adapted martingale difference, there exists* $\sigma > 0$ *such that* $E(z_l^2 \mid \mathscr{F}_{l-1}) = \sigma^2$, *a.s.*; *Then, there exists a constant* $c > 0$, *such that*

$$P(n^{1/2}|\theta_n - \theta_0| > \rho) \leqslant c\rho^{-2} \tag{3.2}$$

for every $\rho > 0$ *and for all* $n \geqslant 1$.

Proof. Assume that the conditions of Theorem 3.2 hold. For any real number sequence $\{c_{ni}\}$, similarly to the proof of Theorem 3.1 and Example 4.1 of Hu (2002) (take $t = 2$ in the Ref. of Hu, 2002), we know that there exists a constant $c_1 > 0$ such that

$$E\left(\sum_{i=1}^{n} c_{ni}\varepsilon_i\right)^2 \leqslant c_1 \sum_{i=1}^{n} c_{ni}^2, \quad \forall\, n \geqslant 1, \tag{3.3}$$

based on (3.3) and Lemma 1.2 and similarly to the proof of Theorem 2.1 of Hu (2002) we can get (3.2). □

Combining Theorem 3.2, Theorem 3.1 of Hu (2002) and the proof of Example 4.1 of Hu (2002), we get the following result.

Theorem 3.3. *Suppose that* $\sup_n E|\varepsilon_n|^t < \infty$, $\sup_l E|z_l|^t < \infty$ *for some* $t \geqslant 2$ *and the conditions of* Theorem 3.2 *hold. Then there exists a constant* $c > 0$ *such that*

$$P(n^{1/2}|\theta_n - \theta_0| > \rho) \leqslant c\rho^{-t} \tag{3.4}$$

for every $\rho > 0$ *and for all* $n \geqslant 1$.

Remark. Theorem 3.3 generalizes the result of Ivanov (1976) to the case when $\{\varepsilon_n\}$ is dependent and improves the results of Prakasa Rao (1984). For any $\varepsilon > 0$, we can take M large enough such that $C(t)M^{-t} < \varepsilon$, then we take $\rho = \rho(n, M) = Mn^{1/t-1}$ and by (2.4) we have

$$P(|\theta_n - \theta_0| > Mn^{1/t-1}) = P(n^{1/2}|\theta_n - \theta_0| > Mn^{1/t-1/2}) \leqslant \frac{C(t)n^{1-t/2}}{M^t n^{1-t/2}} < \varepsilon, \tag{3.5}$$

hence θ_n is a consistent estimator of θ_n and it has the convergence rate

$$\theta_n - \theta_0 = \mathrm{O_p}\left(\frac{1}{n^{1-1/t}}\right). \tag{3.6}$$

References

Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and Its Application. Academic Press, London.

Hu, S.H., 2002. The rate of convergence for the least squares estimator in nonlinear regression model with dependent errors. Sci. China (Series A) 45, 137–146.

Ivanov, A.V., 1976. An asymptotic expansion for the distribution of the least squares estimator of the nonlinear regression parameter. Theory Probab. Appl. 21, 557–570.

MacNeill, I.B., Umphrey, G.J., 1987. Applied Probability, Stochastic Processes, and Sampling Theory. D. Reidel Publishing Company, Dordrecht, pp. 141–165.

Prakasa Rao, B.L.S., 1984. The rate of convergence of the least squares estimator in nonlinear regression model with dependent errors. J. Multivariate Anal. 14, 315–322.

Schmetterer, L., 1974. Introduction to Mathematical Statistics. Springer, Berlin.

Strook, D.W., Varadhan, S.R.S., 1979. Multidimensional Diffusion Processes. Springer, Berlin.