J Math Chem (2008) 44:831–838
DOI 10.1007/s10910-008-9379-2

ORIGINAL PAPER

Asymptotics of kernel error density estimators in nonlinear autoregressive models

Keang Fu · Xiaorong Yang

Received: 29 October 2007 / Accepted: 26 November 2007 / Published online: 19 March 2008

Abstract The limiting distribution of the kernel error estimators in nonlinear autoregressive models is considered. It is shown that, at a fixed point, the distribution of the kernel error density estimator is normal without knowing the nonlinear autoregressive function.

Keywords Nonlinear autoregressive models · Residuals · Kernel density estimator

1 Introduction

Suppose that $\{X_i, i = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of strictly stationary real random variables satisfying the nonlinear autoregressive model of order p

$$X_i = g_\theta(X_{i-1}, \ldots, X_{i-p}) + \varepsilon_i, \quad i \geq 1, \tag{1.1}$$

for some $\theta = (\theta_1, \ldots, \theta_q)' \in \Theta \subset \Re^q$, where $g_\theta, \theta \in \Theta$, is a family of known measurable functions from $\Re^p \to \Re$. Also the ε_i''s are i.i.d. random variables with mean zero, finite variance σ^2 and common density f. Moreover, we assume that $X_{i-1}, \ldots, X_{i-p}$ are independent of $\{\varepsilon_i, i = 1, 2, \ldots\}$.

Supported by the NNSF of China(10671176&10771192).

K. Fu (✉)
Department of Mathematics, Zhejiang University, Yuquan Campus, P.O. Box 1103, Hangzhou 310027, Zhejiang Province, China
e-mail: fukeang@hotmail.com

X. Yang
School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310035, China

Let $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_q)'$ be an estimator for θ, and let

$$\hat{\varepsilon}_i := X_i - g_{\hat{\theta}}(X_{i-1}, \ldots, X_{i-p}), \quad i \geq 1 \tag{1.2}$$

denote the residuals. Based on these residuals, we construct a kernel type estimator of the error density f as follows:

$$\hat{f}_n(t) := \frac{1}{n} \sum_{i=1}^{n} K_{h_n}(t - \hat{\varepsilon}_i), \quad t \in \Re,$$

where $K_h(t) = \frac{1}{h} K(\frac{t}{h})$, h_n is a sequence of positive numbers tending to zero, and K is the kernel density function. We also need to define the kernel error density based on the true errors (which we can not observe) $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$:

$$f_n(t) := \frac{1}{n} \sum_{i=1}^{n} K_{h_n}(t - \varepsilon_i), \quad t \in \Re.$$

It is well known nonlinear autoregressive model plays an important role in time series, and its properties have been extensively studied (c.f. Koul [1]). In this paper, we study the asymptotic distribution of kernel error density estimators in nonlinear autoregressive models, and show that it is asymptotically normal which extends the result of Koul [2], Koul and Osiander [3] and Cheng [4] in parametric autoregressive models.

For the proof, we introduce some basic assumptions which will be used throughout the paper.

(**A1**) Let $U \subset \Theta \subset \Re^q$ be an open neighborhood of θ. We assume that for any $y \in \Re^p$, $\vartheta = (\vartheta_1, \ldots, \vartheta_q) \in U$, $j, k = 1, \ldots, q$,

$$\left| \frac{\partial}{\partial \vartheta_j} g_\vartheta(y) \right| \leq M_1(y), \tag{1.3}$$

$$\left| \frac{\partial^2}{\partial \vartheta_j \partial \vartheta_k} g_\vartheta(y) \right| \leq M_2(y), \tag{1.4}$$

where $\mathsf{E}\left[M_1^4(X_{i-1}, \ldots, X_{i-p})\right] < \infty$ and $\mathsf{E}\left[M_2^4(X_{i-1}, \ldots, X_{i-p})\right] < \infty$ for each $i \geq 1$.

For $1 \leq i \leq n$ and $1 \leq j \leq q$, write

$$Y_{ij} := \frac{\partial}{\partial \theta_j} g_\theta(X_{i-1}, \ldots, X_{i-p}), \tag{1.5}$$

$$Z_{ijk} := \frac{\partial^2}{\partial \theta_j \partial \theta_k} g_\theta(X_{i-1}, \ldots, X_{i-p}). \tag{1.6}$$

(**A2**) There exists $\alpha < 1$ such that Y_{ij} satisfying that

$$\sum_{i=1}^{n} Y_{ij} = O_p(n^{\alpha}), \quad 1 \leq j \leq q. \tag{1.7}$$

(**A3**) The estimator $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_q)'$ for θ satisfying that there exists a constant $C_1 (0 < C_1 < \infty)$ such that

$$\limsup_{n \to \infty} \sqrt{\frac{n}{\log \log n}} \left| \hat{\theta} - \theta \right| \leq C_1 \quad a.s., \tag{1.8}$$

where $|\hat{\theta} - \theta| = \sqrt{\sum_{j=1}^{q} (\hat{\theta}_j - \theta_j)^2}$.

Remark 1.1 It can be seen from (1.5) that ε_i and Y_{ij} are independent because in model (1.1), $X_{i-1}, \ldots, X_{i-p}$ are assumed to be independent of ε_i. The condition (1.7) can be easily fulfilled for AR(1) model (c.f. Brockwell and Davis [5]) or by imposing some fairly non-restrictive weak dependence structures such as α-mixing. Also the condition (1.8) is natural since the least square estimator under certain conditions proposed by Klimko and Nelson [6] satisfies it.

(**A4**) K is a continuous bounded kernel, K'' is bounded, and

$$\int |x| K(x) \mathrm{d}x < \infty, \int |x| K^2(x) \mathrm{d}x < \infty,$$
$$\int |x K'(x)| \mathrm{d}x < \infty, \int |x| |K'(x)|^2 \mathrm{d}x < \infty.$$

Remark 1.2 It is trivially seen that the normal density satisfies all the assumptions on K.

In the sequel, we denote with C a generic constant that may be different in each of its appearance, and $\xrightarrow{d}$ denotes convergence in distribution.

2 Main result and its proof

Theorem 2.1 *Suppose that at a fixed $t \in \Re$, there exists a constant $0 < c < \infty$ such that f satisfies*

$$|f(t) - f(t-y)| \leq c|y|, \quad \text{for all } y \in \Re. \tag{2.1}$$

Assume that h_n satisfies

$$h_n \to 0, \; \frac{n^{1/2} h_n^{5/2}}{\log \log n} \to \infty. \tag{2.2}$$

Then under the assumptions (*A*1)–(*A*4)*, we have that*

$$\frac{1}{\sqrt{\mathsf{Var}(f_n(t))}}\left(\hat{f}_n(t)-\mathsf{E}f_n(t)\right)\xrightarrow{d} N(0,1). \tag{2.3}$$

Proof By (1.1), (1.2) and Taylor's expansion, it follows that

$$\begin{aligned}\varepsilon_i-\hat{\varepsilon}_i &= g_{\hat{\theta}}(X_{i-1},\ldots,X_{i-p})-g_{\theta}(X_{i-1},\ldots,X_{i-p})\\ &=\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)Y_{ij}+1/2(\hat{\theta}-\theta)'Z_i(\hat{\theta}-\theta)\\ &=\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)Y_{ij}+1/2\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_k-\theta_k)Z_{ijk},\end{aligned}$$

where Y_{ij} is defined in (1.5) and Z_i is a $q\times q$ matrix with the jth row and kth column element Z_{ijk} defined in (1.6) and evaluated at $\theta^*:=\theta+\lambda(\hat{\theta}-\theta)$, $\lambda\in(0,1)$, i.e., θ^* takes place of θ in (1.6).

Since $\hat{\theta}$ is a strong consistent estimator of θ by (1.8) and $U\in\Theta$ is a neighborhood of θ, there exists an $N_0>0$ such that for all $n>N_0$, $\hat{\theta}\in U$ and $\theta^*\in U$ a.s. Then for $1\le i\le n$ and $1\le j,k\le q$, in conjunction with (1.3) and (1.4), it follows that

$$\begin{aligned}&\mathsf{E}(Y_{ij}^2)\le \mathsf{E}M_1^2(X_{i-1},\ldots,X_{i-p})\le(\mathsf{E}M_1^4(X_{i-1},\ldots,X_{i-p}))^{1/2}<\infty,\\ &\mathsf{E}|Z_{ijk}|\le(Z_{ijk}^2)^{1/2}\le(\mathsf{E}M_2^2(X_{i-1},\ldots,X_{i-p}))^{1/2}\\ &\quad\le(\mathsf{E}M_2^4(X_{i-1},\ldots,X_{i-p}))^{1/4}<\infty,\\ &\mathsf{E}(Y_{ij}^4)\le\mathsf{E}M_1^4(X_{i-1},\ldots,X_{i-p})<\infty,\\ &\mathsf{E}(Z_{ijk}^4)\le\mathsf{E}M_2^4(X_{i-1},\ldots,X_{i-p})<\infty.\end{aligned}$$

Note that the four inequalities above will be used throughout this paper.

Under the assumptions and the fact that $(a_1+a_2+\cdots+a_m)^2\le m(a_1^2+a_2^2+\cdots+a_m^2)$ for any positive integer m and real numbers $a_1,a_2,\ldots,a_m$, we have that

$$\begin{aligned}\left|\hat{f}_n(t)-f_n(t)\right| &=\left|\frac{1}{nh_n}\sum_{i=1}^{n}\left[K\left(\frac{t-\hat{\varepsilon}_i}{h_n}\right)-K\left(\frac{t-\varepsilon_i}{h_n}\right)\right]\right|\\ &=\left|\frac{1}{nh_n}\sum_{i=1}^{n}\left[\frac{\varepsilon_i-\hat{\varepsilon}_i}{h_n}K'\left(\frac{t-\varepsilon_i}{h_n}\right)+\frac{(\varepsilon_i-\hat{\varepsilon}_i)^2}{2h_n^2}K''(\eta_i(t))\right]\right|\\ &\le\frac{\left|\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)\sum_{i=1}^{n}Y_{ij}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right|}{nh_n^2}\end{aligned}$$

$$+\frac{\left|\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j-\theta_j)(\hat{\theta}_k-\theta_k)\sum_{i=1}^{n}Z_{ijk}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right|}{2nh_n^2}$$
$$+\frac{q\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)^2\sum_{i=1}^{n}Y_{ij}^2\left|K''(\eta_i(t))\right|}{nh_n^3}$$
$$+\frac{q^2\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j-\theta_j)^2(\hat{\theta}_k-\theta_k)^2\sum_{i=1}^{n}Z_{ijk}^2\left|K''(\eta_i(t))\right|}{4nh_n^3}$$
$$=:I+II+III+IV,$$

where $\eta_i(t)$ is a random quantity between $\frac{t-\hat{\varepsilon}_i}{h_n}$ and $\frac{t-\varepsilon_i}{h_n}$. Now we begin to deal with them, respectively. As to I, notice that

$$\left|\sum_{j=1}^{q}(\hat{\theta}_j-\theta_j)\sum_{i=1}^{n}Y_{ij}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right|^2\le\sum_{j=1}^{q}(\hat{\theta}_j-\theta)^2\left(\sum_{i=1}^{n}Y_{ij}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)^2$$

and

$$\left(\sum_{i=1}^{n}Y_{ij}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)^2$$
$$\le 2\left(\sum_{i=1}^{n}Y_{ij}\left(K'\left(\frac{t-\varepsilon_i}{h_n}\right)-\mathsf{E}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)\right)^2+2\left(\sum_{i=1}^{n}Y_{ij}\mathsf{E}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)^2$$
$$=:2(I_1+I_2).$$

Since Y_{ij} and ε_i are independent, by (2.1) and (A5), it follows that

$$\mathsf{Var}\left(K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)\le\mathsf{E}\left(K'\left(\frac{t-\varepsilon_i}{h_n}\right)^2\right)=h_n\int K'^2(y)f(t-h_ny)\mathrm{d}y$$
$$=h_n\int f(t)K'^2(y)\mathrm{d}y+O(h_n)=O(h_n),$$

and hence

$$\mathsf{E}I_1\le C\sum_{i=1}^{n}\mathsf{E}(Y_{ij}^2)\mathsf{Var}\left(K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)=O(nh_n),$$

which leads to that $I_1 = O_p(nh_n)$. Similarly, we can readily obtain $I_2 = O_p(n^{2\alpha}h_n^2)$ by (A2). Thus it follows from (A3) that $I = O_p\left(\frac{1}{nh_n^2}\sqrt{\frac{\log\log n}{n}}(nh_n + n^{2\alpha}h_n^2)\right) = O_p\left(\frac{\sqrt{\log\log n}}{nh_n^{3/2}}\right)$.

For II, note that

$$\left|\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j - \theta_j)(\hat{\theta}_k - \theta_k)\sum_{i=1}^{n} Z_{ijk}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right|^2$$
$$\leq q^2\sum_{j=1}^{q}\sum_{k=1}^{q}(\hat{\theta}_j - \theta_j)^2(\hat{\theta}_k - \theta_k)^2\left(\sum_{i=1}^{n} Z_{ijk}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)^2$$

and

$$\left(\sum_{i=1}^{n} Z_{ijk}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)^2$$
$$\leq 2\left(\sum_{i=1}^{n} Z_{ijk}K'\left(\frac{t-\varepsilon_i}{h_n}\right) - \mathsf{E}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)^2 + 2\left(\sum_{i=1}^{n} Z_{ijk}\mathsf{E}K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right)^2$$
$$=: 2(II_1 + II_2).$$

Since $\mathsf{E}II_1 \leq \sum_{i=1}^{n}\mathsf{E}Z_{ijk}^2\mathsf{Var}\left(K'\left(\frac{t-\varepsilon_i}{h_n}\right)\right) = O(nh_n)$, it leads to $II_1 = O_p(nh_n)$, and then we have $II_2 = O_p(n^2h_n)$. Thus we can get $II = O_p\left(\frac{\log\log n}{nh_n^{3/2}}\right)$.

Also from (A3), we easily have that

$$III = O_p\left(\frac{n\log\log n}{n^2h_n^3}\right) = O_p\left(\frac{\log\log n}{nh_n^3}\right),$$

and

$$IV = O_p\left(\frac{n(\log\log n)^2}{n^3h_n^3}\right) = O_p\left(\frac{(\log\log n)^2}{n^2h_n^3}\right).$$

Hence, coupled with (2.2) it follows that

$$\sqrt{nh_n}\left|\hat{f}_n(t) - f_n(t)\right| = O_p\left(\frac{\log\log n}{n^{1/2}h_n^{5/2}}\right) = o_p(1). \tag{2.4}$$

Note that

$$\begin{aligned}\mathsf{Var}[f_n(t)] &= \frac{1}{nh_n^2}\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right] \\ &= \frac{1}{nh_n}\left[\int K^2(x)f(t-h_nx)\mathrm{d}x - h_n\left(\int K(x)f(t-h_nx)\mathrm{d}x\right)^2\right] \\ &= \frac{1}{nh_n}\left[f(t)\int K^2(x)\mathrm{d}x + O(h_n)\right], \end{aligned} \tag{2.5}$$

and this, combined with $f(t) > 0$, implies that $\mathsf{Var}[f_n(t)] = O(1/(nh_n))$. Thus in order to prove (2.3), it is sufficient to show that

$$\frac{1}{\sqrt{\mathsf{Var}[f_n(t)]}}(f_n(t) - \mathsf{E}f_n(t)) \xrightarrow{d} N(0,1). \tag{2.6}$$

Observe that

$$\frac{f_n(t) - \mathsf{E}f_n(t)}{\sqrt{\mathsf{Var}[f_n(t)]}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{K\left(\frac{t-\varepsilon_i}{h_n}\right) - \mathsf{E}K\left(\frac{t-\varepsilon_i}{h_n}\right)}{\sqrt{\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}},$$

and hence to prove the asymptotic normality, we shall use the Linderberg–Feller's central limit theorem. First, we verify the Linderberg–Feller's condition as follows. Fixing any $\varepsilon > 0$, from the boundedness of K it follows that

$$\begin{aligned}&\mathsf{E}\left[\frac{\left(K\left(\frac{t-\varepsilon_i}{h_n}\right) - \mathsf{E}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]\right)^2}{\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]} I\left(\left|\frac{K\left(\frac{t-\varepsilon_i}{h_n}\right) - \mathsf{E}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}{\sqrt{\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}}\right| > \epsilon\sqrt{n}\right)\right] \\ &\le C\frac{1}{\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}\mathsf{E}\left[I\left(\left|\frac{K\left(\frac{t-\varepsilon_i}{h_n}\right) - \mathsf{E}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}{\sqrt{\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}}\right| > \epsilon\sqrt{n}\right)\right] \\ &= C\frac{1}{\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}\mathsf{P}\left(\left|K\left(\frac{t-\varepsilon_i}{h_n}\right) - \mathsf{E}\left[K\left(t-\varepsilon_1 h_n\right)\right]\right|\right. \\ &\qquad \left. > \epsilon\sqrt{n\mathsf{Var}\left[K\left(\frac{t-\varepsilon_1}{h_n}\right)\right]}\right) \\ &\le \frac{C}{n\epsilon^2\mathsf{Var}\left(K\frac{t-\varepsilon_1}{h_n}\right)} = \frac{C}{n\epsilon^2 h_n[f(t)\int K^2(x)\mathrm{d}x + O(h_n)]} \to 0,\end{aligned}$$

where $f(t) > 0$ and $nh_n \to \infty$ implied by (2.2). Thus by applying the Lindeberg–Feller's central limit theorem, it is easily seen statement (2.6) follows, and the proof is now terminated. □

Acknowledgements The authors wish to express their deep gratitude to Prof. Li-Xin Zhang for his guidance.

References

1. H.L. Koul, *Weighted Empirical Processes in Dynamic Nonlinear Models*. Lecture Notes in Statistics, vol. 166 (Springer, New York, 2002)
2. H.L. Koul, A weak convergence result useful in robust autoregression. J. Stat. Plan. Inference **29**, 291–308 (1991)
3. H.L. Koul, M. Osiander, Weak convergence of randomly weighted residuals empiricals with application to autoregression. Ann. Stat. **22**, 540–562 (1994)
4. F.X. Cheng, Asymptotic distribution of error density estimators in first-order autoregressive models. Sankhyā. Indian J. Stat. **67**(3), 553–567 (2005)
5. P.J. Brockwell, R.A. Davis, *Time Series: Thoery and Methods*. Springer Series in Statistics (Springer-Verlag, New York, 1991)
6. L.A. Klimko, P.I. Nelson, On conditional least squares estimation for stochastic processes. Ann. Stat. **6**, 629–642 (1978)