应用概率统计 第 36 卷
第 1 期 2020 年 2 月

Chinese Journal of Applied Probability and Statistics
Feb., 2020, Vol. 36, No. 1, pp. 26-40
doi: 10.3969/j.issn.1001-4268.2020.01.003

A Semiparametric Estimation of a Regression Function in the Partially Linear Autoregressive Model *

WANG Minghui[1,2] LIU Xiangdong[2⋆] LI Yin[1]

([1] *School of Mathematics and Statistics, Shaoguan University, Shaoguan, 512005, China*)
([2] *School of Economics, Jinan University, Guangzhou, 510632, China*)

Abstract: In this paper, semiparametric estimation of a regression function in the third order partially linear autoregressive model with first order autoregressive errors is mainly studied. We suppose that the regression function has a parametric framework, and use the conditional least squares method to obtain the parameter estimators. Then semiparametric estimators of the regression function can be given by combining with the nonparametric kernel function adjustment. Furthermore, under certain conditions, the consistency of the estimators is proved. Finally, simulation research is presented to evaluate the effectiveness of the proposed method.

Keywords: partially linear autoregressive models; first order autoregressive errors; conditional least squares method; kernel function adjustment; semiparametric estimation

2010 Mathematics Subject Classification: 62M10

Citation: WANG M H, LIU X D, LI Y. A semiparametric estimation of a regression function in the partially linear autoregressive model [J]. Chinese J Appl Probab Statist, 2020, 36(1): 26–40.

§1. Introduction

A partially linear model is a kind of important statistical model developed in 1980s. Due to the introduction of nonparametric components that represent model errors or other systematic errors, the partially linear model contains both parametric and nonparametric components; possess the common advantages of a parametric model and a nonparametric model, while the adaptability is far beyond both of them; and it also has stronger explanatory power. Meanwhile, the so-called “curse of dimensionality” can be avoided. Owing to the briefness, intuition, easy to be understood and the good statistical properties of a parametric estimator, the linear model is widely used when we study the correlation between variables. However, the parametric estimator of a linear model is obtained under

*The project was supported by the National Natural Science Foundation of China (Grant Nos. 71471075, 11501373, 11701380) and 2015 Scientific Research Project of Shaoguan University (Grant No. S201501017).

⋆Corresponding author, E-mail: tliuxd@jnu.edu.cn.

Received April 23, 2018. Revised March 18, 2019.

the hypothesis that the mean value of errors term is zero, the variance is homogeneous and unrelated to each other. Statistical inference has high accuracy, and the modeling method is superior only when these hypotheses are established or approximated. However, both empirical and theoretical studies show that the above hypothesis has a large gap compared with the actual situation in many cases. The nonparametric method is based on this consideration, and the distribution of the errors is weakened or not set, and the statistical inference is carried out only through the sample.

An important application of a partially linear model is to deal with the time series data, for example, the partially linear single-index model:

$$Y_t = U_t^{\mathsf{T}}\beta + \psi(V_t^{\mathsf{T}}\alpha) + \varepsilon_t, \tag{1}$$

where $(\alpha^{\mathsf{T}}, \beta^{\mathsf{T}})^{\mathsf{T}}$ are unknown parameter vectors, $\psi(\cdot)$ is an unknown regression function over $(-\infty, \infty)$, and ε_t represents random error. The linear part of the model depicts the linear relationship between the response variable Y_t and the explanatory variable U_t, and the nonlinear part reflects the nonlinear relationship between Y_t and V_t.

For the model (1), Carroll et al.[1] approximated $\psi(\cdot)$ locally by a linear function, and fitted a parametric generalized linear model to obtain initial parameter estimators. The final estimator was obtained by maximizing the local quasi-likelihood function until convergence. Finally, they presented a class of asymptotically optimal estimators of the unknown parameters and proved the asymptotic normality of parameter estimators. When ε_t is a random variable with $\mathsf{E}(\varepsilon_t) = 0$ and $\mathsf{E}(\varepsilon_t^2) = \sigma^2$, Xia et al.[2] considered another equivalent form $Y_t = X_t^{\mathsf{T}}\theta + \psi(X_t^{\mathsf{T}}\eta) + \varepsilon_t$ of the model (1) and used Nadaraya-Watson and crossvalidation method to estimate $(\alpha^{\mathsf{T}}, \beta^{\mathsf{T}})^{\mathsf{T}}$ and $\psi(\cdot)$. They further had the following asymptotic normality results and the $\sqrt{n}$ consistency about the estimators. The theory and application of this model can be found in the paper [3].

In view of the complexity of the model (1), many scholars have studied several special cases. For example, Gao and Liang[4] obtained piecewise polynomial approximator of nonparametric components and pseudo-LS estimator of parametric components for the second order partially linear autoregressive model. Gao and Yee[5] also considered the second order partially linear autoregressive model and constructed kernel-based estimators for both the parametric and nonparametric components. The proposed estimation not only had good asymptotic properties, but also was effective for both simulation research and empirical analysis. Yu et al.[6] studied semiparametric estimation of regression functions in autoregressive models. They supposed that the regression function had a parametric framework. After the parameter was estimated through conditional least squares method, they adjusted it by a nonparametric factor, and furthermore proved the consistency of the estimators. Farnoosh and Mortazavi [7] considered the first order nonlinear autoregressive model with dependent errors which were defined as first order autoregressive AR(1). Farnoosh et al. [8] studied the semiparametric estimation and properties of regression functions in the second order partially linear autoregressive model. Amiri et al. [9] considered semiparametric estimation of the first order functional autoregressive model. They used the conditional nonlinear least squares method to construct the parametric estimators, and took the recursive smooth kernel approach to estimate regression adjustment. Furthermore, they investigated weak consistencies of the estimators.

This paper mainly studies the third order partially linear autoregressive model with first order autoregressive errors:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + f(Y_{t-3}) + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t,\ |\phi_2| < 1,\ \phi_2 \pm \phi_1 < 1,\ |\rho| < 1, \tag{2}$$

where $\{\varepsilon_t\}_{t=0}^{\infty}$ is a sequence of random variables which is independent and also identically distributed (i.i.d.) with zero mean and variance σ^2, and Y_t is independent with ε_t. We assume that the regression function $f(x)$ has a parametric framework $\{g(x,\theta);\theta \in \Theta\}$, and through the conditional least squares method, we can get the parametric estimators $(\widehat{\theta}, \widehat{\phi}_1, \widehat{\phi}_2, \widehat{\rho})$. In order to avoid a misleading judgment, through multiplying by $\widehat{\xi}(x)$ we can obtain $\widehat{f}(x) = g(x,\widehat{\theta})\widehat{\xi}(x)$ to make a adjustment to $g(x,\widehat{\theta})$. Semiparametric estimator $\widetilde{f}(x) = g(x,\widehat{\theta})\widetilde{\xi}(x)$ of $f(x)$ can be also gained. Related ideas and methods can be found in [6, 10–12]. Furthermore, under certain conditions, we prove the strong consistency of $(\widehat{\theta}, \widehat{\phi}_1, \widehat{\phi}_2, \widehat{\rho})$ and the weak consistency of $\widehat{f}(x)$ and $\widetilde{f}(x)$. At the same time, the statistical simulation method is used to verify the validity of the theoretical results.

This article is an innovation and promotion of the model in the literature [6], and there are two main points: First, we transform the nonlinear effects of first order autoregressive term into linear effects, and add second order autoregressive term to the linear factors and third order autoregressive term to the nonlinear factors. Because there are short-term correlation in many time series. The use of nonlinear models with only first order autoregressive terms is not sufficient to illustrate the problem completely. Second, we improve the independent and identically distributed errors to the first order autoregressive errors. The reason is that the former is ideal, but there is a certain dependence between the error items in the financial or economic series. We have an innovation in the difficulty of estimating methods. The semiparametric estimators $\widehat{f}(x)$ and $\widetilde{f}(x)$ don't have accurately distributions, it can only be proved that their have the weak consistency (according to the probability distribution). And we also use the stationary and reversible conditions in the autoregressive model to extend the independent and identically distributed errors to the first order autoregressive errors.

§2. Estimation of Parameters and Regression Function

For the model (2),

$$u_t = Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - f(Y_{t-3}), \qquad u_{t-1} = Y_{t-1} - \phi_1 Y_{t-2} - \phi_2 Y_{t-3} - f(Y_{t-4}),$$

it is clear that

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + f(Y_{t-3}) + \rho[Y_{t-1} - \phi_1 Y_{t-2} - \phi_2 Y_{t-3} - f(Y_{t-4})] + \varepsilon_t.$$

We need to estimate ϕ_1, ϕ_2, ρ, and the regression function $f(\cdot)$.

We assume that $f(\cdot)$ has a parametric framework,

$$f(x) \in \{g(x,\theta); \theta \in \Theta\}.$$

g is a known function with a limited number of unknown parameters θ, where $\Theta \subseteq R^p$ is the parameter space. So we should estimate θ, ϕ_1, ϕ_2, ρ. Set

$$\begin{aligned}\theta_0 = \mathop{\arg\min}_{\theta\in\Theta,\, |\phi_2|<1,\, \phi_2\pm\phi_1<1,\, |\rho|<1} & \mathsf{E}\{[Y_t - \mathsf{E}_\theta(Y_t \,|\, Y_{t-1}, Y_{t-2}, Y_{t-3})] \\ & - \rho[Y_{t-1} - \mathsf{E}_\theta(Y_{t-1} \,|\, Y_{t-2}, Y_{t-3}, Y_{t-4})]\}^2,\end{aligned}$$

where

$$\begin{aligned}\mathsf{E}_\theta(Y_t \,|\, Y_{t-1}, Y_{t-2}, Y_{t-3}) &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + f(Y_{t-3}) = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + g(Y_{t-3}, \theta), \\ \mathsf{E}_\theta(Y_{t-1} \,|\, Y_{t-2}, Y_{t-3}, Y_{t-4}) &= \phi_1 Y_{t-2} + \phi_2 Y_{t-3} + f(Y_{t-4}) = \phi_1 Y_{t-2} + \phi_2 Y_{t-3} + g(Y_{t-4}, \theta).\end{aligned}$$

Through the conditional least squares method, we now define

$$\begin{aligned}Q_n(\theta, \phi_1, \phi_2, \rho) = \sum_{t=4}^{n} & \{[Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - g(Y_{t-3}, \theta)] \\ & - \rho[Y_{t-1} - \phi_1 Y_{t-2} - \phi_2 Y_{t-3} - g(Y_{t-4}, \theta)]\}^2,\end{aligned}$$

and can obtain

$$(\widehat{\theta}_n, \widehat{\phi}_{1n}, \widehat{\phi}_{2n}, \widehat{\rho}_n) = \mathop{\arg\min}_{\theta\in\Theta,\, |\phi_2|<1,\, \phi_2\pm\phi_1<1,\, |\rho|<1} Q_n(\theta, \phi_1, \phi_2, \rho).$$

Furthermore, we define

$$\begin{cases} \partial Q_n(\theta,\phi_1,\phi_2,\rho)/\partial\theta=0; \\ \partial Q_n(\theta,\phi_1,\phi_2,\rho)/\partial\phi_1=0; \\ \partial Q_n(\theta,\phi_1,\phi_2,\rho)/\partial\phi_2=0; \\ \partial Q_n(\theta,\phi_1,\phi_2,\rho)/\partial\rho=0, \end{cases}$$

and can obtain

$$\begin{cases} 1=\dfrac{\sum\limits_{t=4}^{n}[(Y_t-\widehat{\phi}_{1n}Y_{t-1}-\widehat{\phi}_{2n}Y_{t-2})-\widehat{\rho}_n(Y_{t-1}-\widehat{\phi}_{1n}Y_{t-2}-\widehat{\phi}_{2n}Y_{t-3})]\Big[\dfrac{\partial g(Y_{t-3},\widehat{\theta}_n)}{\partial\theta}-\dfrac{\widehat{\rho}_n\partial g(Y_{t-4},\widehat{\theta}_n)}{\partial\theta}\Big]}{\sum\limits_{t=4}^{n}[g(Y_{t-3},\widehat{\theta}_n)-\widehat{\rho}_n g(Y_{t-4},\widehat{\theta}_n)]\Big[\dfrac{\partial g(Y_{t-3},\widehat{\theta}_n)}{\partial\theta}-\dfrac{\widehat{\rho}_n\partial g(Y_{t-4},\widehat{\theta}_n)}{\partial\theta}\Big]}; \\ \widehat{\phi}_{1n}=\dfrac{\sum\limits_{t=4}^{n}\{[Y_t-\widehat{\phi}_{2n}Y_{t-2}-g(Y_{t-3},\widehat{\theta}_n)]-\widehat{\rho}_n[Y_{t-1}-\widehat{\phi}_{2n}Y_{t-3}-g(Y_{t-4},\widehat{\theta}_n)]\}(Y_{t-1}-\widehat{\rho}_nY_{t-2})}{\sum\limits_{t=4}^{n}(Y_{t-1}-\widehat{\rho}_nY_{t-2})^2}; \\ \widehat{\phi}_{2n}=\dfrac{\sum\limits_{t=4}^{n}\{[Y_t-\widehat{\phi}_{1n}Y_{t-1}-g(Y_{t-3},\widehat{\theta}_n)]-\widehat{\rho}_n[Y_{t-1}-\widehat{\phi}_{1n}Y_{t-2}-g(Y_{t-4},\widehat{\theta}_n)]\}(Y_{t-2}-\widehat{\rho}_nY_{t-3})}{\sum\limits_{t=4}^{n}(Y_{t-2}-\widehat{\rho}_nY_{t-3})^2}; \\ \widehat{\rho}_n=\dfrac{\sum\limits_{t=4}^{n}[Y_t-\widehat{\phi}_{1n}Y_{t-1}-\widehat{\phi}_{2n}Y_{t-2}-g(Y_{t-3},\widehat{\theta}_n)][Y_{t-1}-\widehat{\phi}_{1n}Y_{t-2}-\widehat{\phi}_{2n}Y_{t-3}-g(Y_{t-4},\widehat{\theta}_n)]}{\sum\limits_{t=4}^{n}[Y_{t-1}-\widehat{\phi}_{1n}Y_{t-2}-\widehat{\phi}_{2n}Y_{t-3}-g(Y_{t-4},\widehat{\theta}_n)]^2}. \end{cases}$$

Similar to the ideas of [6] and [11], we set

$$\begin{aligned} q_n(x,\xi)=&\frac{1}{h_n}\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)[f(Y_{t-3})-g(Y_{t-3},\widehat{\theta}_n)\xi]^2 \\ &+\frac{1}{h_n}\sum_{t=4}^{n}K\Big(\frac{Y_{t-4}-x}{h_n}\Big)[f(Y_{t-4})-g(Y_{t-4},\widehat{\theta}_n)\xi]^2, \end{aligned}$$

where $K(\cdot)$ is a kernel function and h_n is the bandwidth. Set $\partial q_n(x,\xi)/\partial\xi=0$, and we can get

$$\widehat{\xi}(x)=\frac{\sum\limits_{t=4}^{n}\Big[K\Big(\dfrac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})g(Y_{t-3},\widehat{\theta}_n)+K\Big(\dfrac{Y_{t-4}-x}{h_n}\Big)f(Y_{t-4})g(Y_{t-4},\widehat{\theta}_n)\Big]}{\sum\limits_{t=4}^{n}\Big[K\Big(\dfrac{Y_{t-3}-x}{h_n}\Big)g^2(Y_{t-3},\widehat{\theta}_n)+K\Big(\dfrac{Y_{t-4}-x}{h_n}\Big)g^2(Y_{t-4},\widehat{\theta}_n)\Big]},$$

then

$$\widehat{f}(x)=g(x,\widehat{\theta}_n)\widehat{\xi}(x).$$

Note that

$$\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})g(Y_{t-3},\widehat{\theta}_n)$$

$$\begin{aligned}
&\approx \sum_{t=4}^{n} K\Big(\frac{Y_{t-3}-x}{h_n}\Big)(Y_t-\widehat{\phi}_{1n}Y_{t-1}-\widehat{\phi}_{2n}Y_{t-2})g(Y_{t-3},\widehat{\theta}_n),\\
&\sum_{t=4}^{n} K\Big(\frac{Y_{t-4}-x}{h_n}\Big)f(Y_{t-4})g(Y_{t-4},\widehat{\theta}_n)\\
&\approx \sum_{t=4}^{n} K\Big(\frac{Y_{t-4}-x}{h_n}\Big)(Y_{t-1}-\widehat{\phi}_{1n}Y_{t-2}-\widehat{\phi}_{2n}Y_{t-3})g(Y_{t-4},\widehat{\theta}_n),
\end{aligned}$$

and we adjust $\widehat{\xi}(x)$ by

$$\begin{aligned}
\widetilde{\xi}(x) = &\frac{\sum_{t=4}^{n} K\Big(\frac{Y_{t-3}-x}{h_n}\Big)(Y_t-\widehat{\phi}_{1n}Y_{t-1}-\widehat{\phi}_{2n}Y_{t-2})g(Y_{t-3},\widehat{\theta}_n)}{\sum_{t=4}^{n}\Big[K\Big(\frac{Y_{t-3}-x}{h_n}\Big)g^2(Y_{t-3},\widehat{\theta}_n)+K\Big(\frac{Y_{t-4}-x}{h_n}\Big)g^2(Y_{t-4},\widehat{\theta}_n)\Big]}\\
&+\frac{\sum_{t=4}^{n} K\Big(\frac{Y_{t-4}-x}{h_n}\Big)(Y_{t-1}-\widehat{\phi}_{1n}Y_{t-2}-\widehat{\phi}_{2n}Y_{t-3})g(Y_{t-4},\widehat{\theta}_n)}{\sum_{t=4}^{n}\Big[K\Big(\frac{Y_{t-3}-x}{h_n}\Big)g^2(Y_{t-3},\widehat{\theta}_n)+K\Big(\frac{Y_{t-4}-x}{h_n}\Big)g^2(Y_{t-4},\widehat{\theta}_n)\Big]},
\end{aligned}$$

therefore the semiparametric estimator of $f(x)$ can be obtained by

$$\widetilde{f}(x) = g(x,\widehat{\theta}_n)\widetilde{\xi}(x).$$

§3. Main Theoretical Results

In this section, we will prove the strong consistency of the parametric estimators $(\widehat{\theta}_n,\widehat{\phi}_{1n},\widehat{\phi}_{2n},\widehat{\rho}_n)$ the weak consistency of the semiparametric estimators $\widehat{f}(x)$ and $\widetilde{f}(x)$. In order to obtain these properties, we need the following conditions: $\theta\in\Theta$, $i,j,k=1,2,\cdots,p$, $l=3,4$,

C1 The sequence $\{Y_t\}_{t=0}^{\infty}$ is a stationary ergodic sequence of integrable random variables.

C2 $\partial g/\partial\theta_i$, $\partial^2 g/(\partial\theta_i\partial\theta_j)$, $\partial^3 g/(\partial\theta_i\partial\theta_j\partial\theta_k)$ exist and are continuous.

C3 $\mathsf{E}|(Y_t-g)\partial g/\partial\theta_i|<\infty$, $\mathsf{E}|(Y_t-g)\partial^2 g/(\partial\theta_i\partial\theta_j)|<\infty$, $\mathsf{E}|(\partial g/\partial\theta_i)\cdot(\partial g/\partial\theta_j)|<\infty$, where g and its partial derivatives are evaluated at (θ_0,Y_{t-l}).

C4 There are functions $H^{(0)}(Y_{t-l})$, $H_i^{(1)}(Y_{t-l})$, $H_{ij}^{(2)}(Y_{t-l})$, $H_{ijk}^{(3)}(Y_{t-l})$ such that

$$|g|\leqslant H^{(0)},\quad \Big|\frac{\partial g}{\partial\theta_i}\Big|\leqslant H_i^{(1)},\quad \Big|\frac{\partial^2 g}{\partial\theta_i\partial\theta_j}\Big|\leqslant H_{ij}^{(2)},\quad \Big|\frac{\partial^3 g}{\partial\theta_i\partial\theta_j\partial\theta_k}\Big|\leqslant H_{ijk}^{(3)},$$

and

$$\mathsf{E}|Y_t\cdot H_{ijk}^{(3)}(Y_{t-l})|<\infty,\quad \mathsf{E}[H^{(0)}(Y_{t-l})\cdot H_{ijk}^{(3)}(Y_{t-l})]<\infty,\quad \mathsf{E}[H_i^{(1)}(Y_{t-l})\cdot H_{jk}^{(2)}(Y_{t-l})]<\infty.$$

C5 $\mathsf{E}(Y_t\,|\,Y_{t-1},Y_{t-2},\cdots,Y_0)=\mathsf{E}(Y_t\,|\,Y_{t-1},Y_{t-2},\cdots,Y_{t-m})$ a.s., $t\geqslant m$

$$\mathsf{E}\Big[U_t^2(\theta_0)\Big|\frac{\partial g(Y_{t-l},\theta_0)}{\partial\theta_i}\cdot\frac{\partial g(Y_{t-l},\theta_0)}{\partial\theta_j}\Big|\Big]<\infty,$$

where

$$\begin{aligned}U_t(\theta_0)&=[Y_t-\mathsf{E}(Y_t\,|\,Y_{t-1},Y_{t-2},Y_{t-3})]-\rho[Y_{t-1}-\mathsf{E}(Y_{t-1}\,|\,Y_{t-2},Y_{t-3},Y_{t-4})]\\&=Y_t-\phi_1Y_{t-1}-\phi_2Y_{t-2}-f(Y_{t-3})-\rho[Y_{t-1}-\phi_1Y_{t-2}-\phi_2Y_{t-3}-f(Y_{t-4})].\end{aligned}$$

We assume that the matrices

$$\begin{aligned}V_l&=\Big[\mathsf{E}\Big(\frac{\partial g(Y_{t-l},\theta_0)}{\partial\theta_i}\cdot\frac{\partial g(Y_{t-l},\theta_0)}{\partial\theta_j}\Big)\Big],\\W_l&=\Big[\mathsf{E}\Big(U_t^2(\theta_0)\frac{\partial g(Y_{t-l},\theta_0)}{\partial\theta_i}\cdot\frac{\partial g(Y_{t-l},\theta_0)}{\partial\theta_j}\Big)\Big]\end{aligned}$$

are positive definite.

C6 The sequence $\{Y_t\}_{t\in N}$ is α-mixing.

C7 Y_0 and Y_1 have the same distribution $\pi(\cdot)$, and the density $\mu(\cdot)$ is bounded, continuous and strictly positive in a neighborhood of x.

C8 $f(x)$ and $g(x,\theta)$ are bounded and continuous with respect to x, away from 0 in a neighborhood of x. Set $g_0(x)=g(x,\theta_0)$.

C9 $g(x,\theta)$ has a continuous derivative with respect to θ, and the derivative at θ_0 is uniformly bounded with respect to x.

C10 The kernel function $K:R^1\to R^+$ is compactly symmetric bounded, such that $K(\cdot)>0$ in a set of positive Lebesgue measures.

C11 $h_n=\beta n^{-1/5}$, where $\beta>0$.

Remark 1 The rationality of the conditions of C1 – C5 can be referred to the literature [13]. The proof of the following Lemma 3 is shown by its Corollary 2.2.

Remark 2 The rationality of the conditions of C6 – C11 can be referred to the literatures [6] and [14].

Lemma 3 Under the conditions of C1 – C5, there exists a sequence estimators $\{\widehat{\theta}_n\}$ that can satisfy $\widehat{\theta}_n\xrightarrow{\text{a.s.}}\theta_0$. Suppose that for any nonzero vector of constants $c'=(c_1,c_2,\cdots,c_p)$,

$$\limsup_{n\to\infty}\frac{\sum\limits_{t=3}^n h(Y_{t-l},\theta_0,c)U_t(\theta_0)}{(2n\delta_l^2\ln\ln n)^{1/2}}=1\qquad\text{a.s.},$$

where

$$h(Y_{t-l},\theta_0,c)=\sum_{i=1}^{p}c_i\frac{\partial g(Y_{t-l},\theta_0)}{\partial\theta_i},\qquad \delta_l^2=c'V_l^{-1}W_lV_l^{-1}c,\ l=3,4,$$

then

$$\limsup_{n\to\infty}\frac{n^{1/2}c'(\widehat{\theta}_n-\theta_0)}{(2\delta_l^2\ln\ln n)^{1/2}}=1\qquad\text{a.s..}$$

Lemma 4 Under the conditions of C1 − C11, we have

(i) $n^{-4/5}\sum_{t=4}^{n}K\Big(\dfrac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})g(Y_{t-3},\widehat{\theta}_n)\xrightarrow{\mathsf{P}}\beta\mu(x)f(x)g_0(x)$,

(ii) $n^{-4/5}\sum_{t=4}^{n}K\Big(\dfrac{Y_{t-4}-x}{h_n}\Big)f(Y_{t-4})g(Y_{t-4},\widehat{\theta}_n)\xrightarrow{\mathsf{P}}\beta\mu(x)f(x)g_0(x)$,

(iii) $n^{-4/5}\sum_{t=4}^{n}K\Big(\dfrac{Y_{t-3}-x}{h_n}\Big)g^2(Y_{t-3},\widehat{\theta}_n)\xrightarrow{\mathsf{P}}\beta\mu(x)g_0^2(x)$,

(iv) $n^{-4/5}\sum_{t=4}^{n}K\Big(\dfrac{Y_{t-4}-x}{h_n}\Big)g^2(Y_{t-4},\widehat{\theta}_n)\xrightarrow{\mathsf{P}}\beta\mu(x)g_0^2(x)$.

Proof Observe that

$$\begin{aligned}
&n^{-4/5}\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})g(Y_{t-3},\widehat{\theta}_n)\\
&=n^{-4/5}\sum_{t=4}^{n}\Big\{K\Big(\frac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})[g(Y_{t-3},\widehat{\theta}_n)-g(Y_{t-3},\theta_0)]\Big\}\\
&\quad+n^{-4/5}\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})g(Y_{t-3},\theta_0)\\
&\triangleq A_n+B_n.
\end{aligned}$$

With Lemma 3 and C8 − C10, we know

$$\max_{3\leqslant t\leqslant n}|g(Y_{t-3},\widehat{\theta}_n)-g(Y_{t-3},\theta_0)|=O\Big(\Big(\frac{\ln\ln n}{n}\Big)^{1/2}\Big),$$

and $f_0>0$, $K_0>0$ separately meet the conditions of $|f(x)|\leqslant f_0$, $|K(x)|\leqslant K_0$, then

$$\begin{aligned}
|A_n|&=n^{-4/5}\Big|\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})[g(Y_{t-3},\widehat{\theta}_n)-g(Y_{t-3},\theta_0)]\Big|\\
&\leqslant n^{-4/5}\sum_{t=4}^{n}\Big|K\Big(\frac{Y_{t-3}-x}{h_n}\Big)\Big||f(Y_{t-3})||g(Y_{t-3},\widehat{\theta}_n)-g(Y_{t-3},\theta_0)|\\
&\leqslant n^{-4/5}\cdot n\cdot K_0\cdot f_0\cdot O\Big(\Big(\frac{\ln\ln n}{n}\Big)^{1/2}\Big)\\
&=O\Big(\frac{(\ln\ln n)^{1/2}}{n^{3/10}}\Big),
\end{aligned}$$

thus when $n\to\infty$, $A_n\xrightarrow{\text{a.s.}}0$. According to C6,

$$n^{-4/5}\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)f(Y_{t-3})g(Y_{t-3},\theta_0)-n^{1/5}\mathsf{E}\Big[K\Big(\frac{Y_0-x}{h_n}\Big)f(Y_0)g(Y_0,\theta_0)\Big]\xrightarrow{\mathsf{P}}0.$$

Set $u = (y-x)/h_n$, then

$$\begin{aligned}
n^{1/5}\mathsf{E}\Big[K\Big(\frac{Y_0-x}{h_n}\Big)f(Y_0)g(Y_0,\theta_0)\Big] &= \frac{\beta}{h_n}\int K\Big(\frac{y-x}{h_n}\Big)f(y)g(y,\theta_0)\mu(y)\,\mathrm{d}y \\
&= \beta\int K(u)f(h_nu+x)g(h_nu+x,\theta_0)\mu(h_nu+x)\,\mathrm{d}u \\
&= \beta\int K(u)f(x)g(x,\theta_0)\mu(x)[1+o(1)]\,\mathrm{d}u \\
&= \beta\mu(x)f(x)g_0(x)[1+o(1)],
\end{aligned}$$

so $B_n \xrightarrow{\mathsf{P}} \beta\mu(x)f(x)g_0(x)$. Combining these results, we can obtain (i). The proof of (ii) is similar to the proof of (i).

Note that

$$\begin{aligned}
&n^{-4/5}\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)g^2(Y_{t-3},\widehat{\theta}_n) \\
={}& n^{-4/5}\sum_{t=4}^{n}\Big\{K\Big(\frac{Y_{t-3}-x}{h_n}\Big)g(Y_{t-3},\widehat{\theta}_n)[g(Y_{t-3},\widehat{\theta}_n)-g(Y_{t-3},\theta_0)]\Big\} \\
&+n^{-4/5}\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)g(Y_{t-3},\widehat{\theta}_n)g(Y_{t-3},\theta_0).
\end{aligned}$$

Analogously, we can get (iii). The proof of (iv) is similar to the proof of (iii). $\square$

Theorem 5 Under the conditions of C1−C11, when $n\to\infty$, $(\widehat{\theta}_n,\widehat{\phi}_{1n},\widehat{\phi}_{2n},\widehat{\rho}_n)\xrightarrow{\text{a.s.}}(\theta,\phi_1,\phi_2,\rho)$.

Proof Suppose that $\widehat{\alpha}_n = (\widehat{\theta}_n,\widehat{\phi}_{1n},\widehat{\phi}_{2n},\widehat{\rho}_n)'$, $\alpha_0 = (\theta,\phi_1,\phi_2,\rho)'$. The Taylor expansion of $Q_n(\alpha)$ about α_0 is

$$\begin{aligned}
Q_n(\alpha) &= Q_n(\alpha_0)+(\alpha-\alpha_0)'\frac{\partial Q_n(\alpha_0)}{\partial\alpha}+\frac{1}{2}(\alpha-\alpha_0)'\frac{\partial^2 Q_n(\alpha^*)}{\partial\alpha^2}(\alpha-\alpha_0) \\
&= Q_n(\alpha_0)+(\alpha-\alpha_0)'\frac{\partial Q_n(\alpha_0)}{\partial\alpha}+\frac{1}{2}(\alpha-\alpha_0)'V_n(\alpha-\alpha_0) \\
&\quad+\frac{1}{2}(\alpha-\alpha_0)'T_n(\alpha^*)(\alpha-\alpha_0),
\end{aligned}$$

where $V_n = [\partial^2 Q_n(\alpha_0)/(\partial\alpha_i\partial\alpha_j)]$, $T_n(\alpha^*) = [\partial^2 Q_n(\alpha^*)/\partial\alpha^2]-V_n$.

The ergodic theorem yields:

$$\begin{aligned}
\frac{1}{n}(\alpha-\alpha_0)'\frac{\partial Q_n(\alpha_0)}{\partial\alpha} &\xrightarrow{\text{a.s.}} 0, \\
\frac{1}{2n}(\alpha-\alpha_0)'V_n(\alpha-\alpha_0) &\xrightarrow{\text{a.s.}} (\alpha-\alpha_0)'V(\alpha-\alpha_0),
\end{aligned}$$

where V is a positive definite matrix of constants according to the strong laws for martingales in [15]. For given $\delta>0$, $\varepsilon>0$ and let N_δ denote the open sphere of radius δ

centered at α_0. Through Egoroff's theorem we can find an event E with $\mathsf{P}(E) > 1-\varepsilon$, a positive $\delta^* < \delta$, $M > 0$ and an n_0 such that on E. For any $n > n_0$, $\alpha \in N_{\delta^*}$, there is $|(\alpha-\alpha_0)'\partial Q_n(\alpha_0)/\partial\alpha| < n\delta^3$, and the minimum eigenvalue of $V_n/(2n)$ is greater than some $\Delta > 0$. Under the conditions of C1 − C4, it also implies that $R_n = (\alpha-\alpha_0)'T_n(\alpha^*)(\alpha-\alpha_0)/2$ satisfies $\lim\limits_{n\to\infty}\sup\limits_{\delta\to 0}(|T_n(\alpha^*)_{ij}|/n\delta) < \infty$ a.s., thus $|(\alpha-\alpha_0)'T_n(\alpha^*)(\alpha-\alpha_0)|/2 < nM\delta^3$.

Using the Taylor expansion of $Q_n(\alpha)$, for α on the boundary of N_{δ^*},

$$Q_n(\alpha) \geqslant Q_n(\alpha_0) + (-n\delta^3 + n\delta^2\Delta - nM\delta^3) = Q_n(\alpha_0) + n\delta^2(\Delta - \delta - M\delta).$$

Since $\Delta - \delta - M\delta$ can be made positive by initially choosing δ sufficiently small, $Q_n(\alpha)$ must attain a minimum at some $\widehat{\alpha}_n$ in N_{δ^*}, at which point the least squares equations $\partial Q_n(\alpha)/\partial\alpha_i = 0$ must be satisfied.

Let $\varepsilon_k = 2^{-k}$ and $\delta_k = 1/k$, $k = 1, 2, \cdots$ be given, and let $\{E_k\}$ denote a sequence of events and $\{n_k\}$ is an increasing sequence. For $n_k < n \leqslant n_{k+1}$ define $\widehat{\alpha}_n$ on E_k to be a root of $\partial Q_n(\alpha)/\partial\alpha_i = 0$ within δ_k of α_0 at which Q_n attains a relative minimum and define $\widehat{\alpha}_n$ to be zero otherwise. Then $\widehat{\alpha}_n \xrightarrow{\text{a.s.}} \alpha_0$ on $\liminf E_k$ and we complete the proof. □

Theorem 6 Under the conditions of C1 − C11, when $n \to \infty$, $\widehat{f}(x) \xrightarrow{\mathsf{P}} f(x)$.

Proof We can prove this theorem by using $\widehat{\theta}_n \xrightarrow{\text{a.s.}} \theta_0$ and Lemma 4. □

Theorem 7 Under the conditions of C1 − C11, when $n \to \infty$, $\widetilde{f}(x) \xrightarrow{\mathsf{P}} f(x)$.

Proof Observe that

$$\begin{aligned}\widetilde{f}(x) - \widehat{f}(x) &= g(x,\widehat{\theta}_n)\frac{\sum\limits_{t=4}^{n}\Big[K\Big(\dfrac{Y_{t-3}-x}{h_n}\Big)u_t g(Y_{t-3},\widehat{\theta}_n) + K\Big(\dfrac{Y_{t-4}-x}{h_n}\Big)u_t g(Y_{t-4},\widehat{\theta}_n)\Big]}{\sum\limits_{t=4}^{n}\Big[K\Big(\dfrac{Y_{t-3}-x}{h_n}\Big)g^2(Y_{t-3},\widehat{\theta}_n) + K\Big(\dfrac{Y_{t-4}-x}{h_n}\Big)g^2(Y_{t-4},\widehat{\theta}_n)\Big]}\\ &\triangleq g(x,\widehat{\theta}_n)\frac{C_n + D_n}{\sum\limits_{t=4}^{n}\Big[K\Big(\dfrac{Y_{t-3}-x}{h_n}\Big)g^2(Y_{t-3},\widehat{\theta}_n) + K\Big(\dfrac{Y_{t-4}-x}{h_n}\Big)g^2(Y_{t-4},\widehat{\theta}_n)\Big]}.\end{aligned}$$

From $u_t = \varepsilon_t + \widehat{\rho}_n\varepsilon_{t-1} + \widehat{\rho}_n^2\varepsilon_{t-2} + \cdots + \widehat{\rho}_n^t\varepsilon_0$, we know

$$\begin{aligned}n^{-4/5}C_n &= n^{-4/5}\sum_{t=4}^{n}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)\sum_{i=0}^{t}\widehat{\rho}_n^{t-i}\varepsilon_i g(Y_{t-3},\widehat{\theta}_n)\\ &= n^{-4/5}\sum_{t=4}^{n}\sum_{i=0}^{t}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)\widehat{\rho}_n^{t-i}\varepsilon_i[g(Y_{t-3},\widehat{\theta}_n) - g(Y_{t-3},\theta_0)]\\ &\quad + n^{-4/5}\sum_{t=4}^{n}\sum_{i=0}^{t}K\Big(\frac{Y_{t-3}-x}{h_n}\Big)\widehat{\rho}_n^{t-i}\varepsilon_i g(Y_{t-3},\theta_0)\\ &\triangleq E_n + F_n.\end{aligned}$$

According to Theorem B in [16], when $n \to \infty$, we have $\max\limits_{1\leqslant i\leqslant n} |\varepsilon_i| = O((\ln n)^{1/2})$, then

$$\begin{aligned}
|E_n| &\leqslant n^{-4/5} \sum_{t=4}^{n}\sum_{i=0}^{t} \Big|K\Big(\frac{Y_{t-3}-x}{h_n}\Big)\Big| |\widehat{\rho}_n^{t-i}||\varepsilon_i||g(Y_{t-3},\widehat{\theta}_n) - g(Y_{t-3},\theta_0)| \\
&\leqslant n^{-4/5} \sum_{t=4}^{n}\sum_{i=0}^{t} K_0 \cdot |\widehat{\rho}_n^{t-i}| \cdot O((\ln n)^{1/2}) \cdot O\Big(\Big(\frac{\ln\ln n}{n}\Big)^{1/2}\Big) \\
&\leqslant n^{-4/5} \cdot n \cdot \frac{K_0}{1-|\widehat{\rho}_n|} \cdot O\Big(\Big(\frac{\ln n \ln\ln n}{n}\Big)^{1/2}\Big) \\
&= O\Big(\frac{(\ln n \ln\ln n)^{1/2}}{n^{3/10}}\Big),
\end{aligned}$$

thus when $n \to \infty$, $E_n \xrightarrow{\text{a.s.}} 0$. Besides,

$$\begin{aligned}
\mathsf{E}(F_n) &= n^{-4/5}\mathsf{E}\Big[\sum_{t=4}^{n}\sum_{i=0}^{t} K\Big(\frac{Y_{t-3}-x}{h_n}\Big)\widehat{\rho}_n^{t-i}\varepsilon_i g(Y_{t-3},\theta_0)\Big] = 0, \\
\mathsf{E}(F_n^2) &= n^{-8/5} \sum_{t=4}^{n}\sum_{i=0}^{t} \mathsf{E}\Big[K^2\Big(\frac{Y_{t-3}-x}{h_n}\Big)\widehat{\rho}_n^{2(t-i)}\varepsilon_i^2 g^2(Y_{t-3},\theta_0)\Big] \\
&\quad + 2n^{-8/5} \sum_{4\leqslant t_1<t_2\leqslant n}\ \sum_{0\leqslant i_1<i_2\leqslant t} \mathsf{E}\Big[K\Big(\frac{Y_{t_1-3}-x}{h_n}\Big)\widehat{\rho}_n^{t_1-i_1}\varepsilon_{i_1} g(Y_{t_1-3},\theta_0) \\
&\quad \cdot K\Big(\frac{Y_{t_2-3}-x}{h_n}\Big)\widehat{\rho}_n^{t_2-i_2}\varepsilon_{i_2} g(Y_{t_2-3},\theta_0)\Big] \\
&\leqslant n^{-8/5} \cdot n \cdot \frac{K_0^2}{1-\rho^2} \cdot \sigma^2 \cdot g_0^2 \\
&= O\Big(\frac{1}{n^{3/5}}\Big),
\end{aligned}$$

when $n \to \infty$, we can get $F_n \xrightarrow{\mathsf{P}} 0$, so $n^{-4/5}C_n \xrightarrow{\mathsf{P}} 0$. $n^{-4/5}D_n \xrightarrow{\mathsf{P}} 0$ is likewise to be proved. Using Lemma 4, we have

$$\widetilde{f}(x) - \widehat{f}(x) \xrightarrow{\mathsf{P}} 0. \qquad \square$$

§4. Simulation Research and Analysis

Under the conditions of C1 − C11, simulation research are presented to evaluate the effectiveness of the proposed estimator $\widetilde{f}(x)$ for the model (2).

Set $n = 500$, and consider the following model:

$$Y_t = Y_{t-1} - 0.5Y_{t-2} + f(Y_{t-3}) + u_t, \qquad u_t = 0.4u_{t-1} + \varepsilon_t,$$

$\varepsilon_t \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, 0.15^2)$. We choose the Gaussian kernel function $\mathrm{e}^{-u^2/2}/\sqrt{2\pi}$ to examine two types of $f(x)$ below:

(i) $f(x) = 4\mathrm{e}^{-x^2} + ax$, we assume $g(x,\theta) = \theta\mathrm{e}^{-x^2}$,

(ii) $f(x) = 3\mathrm{e}^{-2x} + a\cos x$, we assume $g(x,\theta) = \theta_1\mathrm{e}^{\theta_2 x}$,

$a = 0.1,\ 0.2,\ 0.3$. We use

$$\text{Mean Square Error: MSE} = \frac{1}{n}\sum_{i=1}^{n}[\widetilde{f}(x_i) - f(x_i)]^2,$$
$$\text{Standard Error: SD} = \sqrt{\text{MSE}}$$

to measure the accuracy of $\widetilde{f}(x)$, and repeat the process of calculation for 100 times.

Tables 1 and 2 separately list the estimated values of two functional forms under different a and h_n, as well as their MSE and SD. With the increase of a, the accuracy of $\widetilde{f}(x)$ is weakened, but the overall fitting effect is better.

Table 1 The estimated value and accuracy of the parameters in $f(x) = 4\mathrm{e}^{-x^2} + ax$

a	$\widehat{\theta}$	$\widehat{\phi}_1$	$\widehat{\phi}_2$	$\widehat{\rho}$	h_n	MSE	SD
					0.04	0.000897920	0.02996531
0.1	3.981702	0.9834588	−0.4958761	0.3983633	0.06	0.000768543	0.02772261
					0.08	0.001356673	0.03683304
					0.06	0.003106583	0.05573673
0.2	3.964253	0.9756316	−0.4813675	0.3892175	0.08	0.002677549	0.05174504
					0.10	0.003066771	0.05537843
					0.06	0.013360544	0.11558782
0.3	3.958722	0.9727935	−0.4808183	0.3851976	0.08	0.007429996	0.08619742
					0.10	0.007890625	0.08882919

Table 2 The estimated value and accuracy of the parameters in $f(x)=3\mathrm{e}^{-2x}+a\cos x$

a	$\widehat{\theta}_1$	$\widehat{\theta}_2$	$\widehat{\phi}_1$	$\widehat{\phi}_2$	$\widehat{\rho}$	h_n	MSE	SD
						0.08	0.06551853	0.2559659
0.1	2.990853	−1.965372	0.9987291	−0.4944566	0.4089553	0.10	0.06550958	0.2559484
						0.12	0.06551154	0.2559522
						0.08	0.06845322	0.2616357
0.2	2.985854	−1.899653	0.9892157	−0.4909882	0.4092574	0.10	0.06844986	0.2616292
						0.12	0.06845688	0.2616427
						0.04	0.08151222	0.2855035
0.3	2.951269	−1.866859	0.9886543	−0.4897532	0.4085658	0.06	0.08145388	0.2854013
						0.08	0.08149803	0.2854786

Figures 1 and 2 are the comparison diagrams of the functions $f(x)=4\mathrm{e}^{-x^2}+0.2x$ at $h_n=0.08$ and $f(x)=3\mathrm{e}^{-2x}+0.2\cos x$ at $h_n=0.10$. The full line shows the true value, while the dotted line represents the estimated value. It also shows that the fitting effect of $\widetilde{f}(x)$ is better.

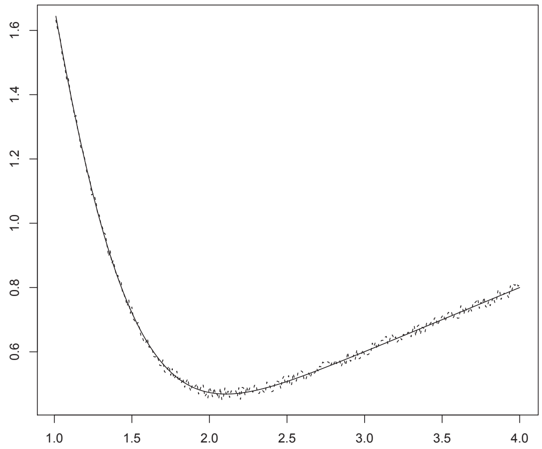

Figure 1 $f(x)=4\mathrm{e}^{-x^2}+0.2x$, $h_n=0.08$

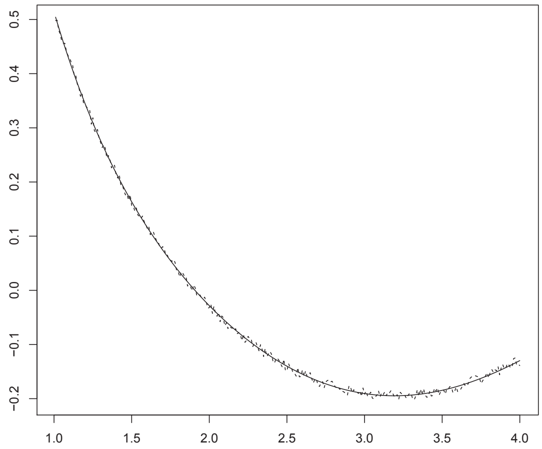

Figure 2 $f(x)=3\mathrm{e}^{-2x}+0.2\cos x$, $h_n=0.10$

Furthermore, we compare this method with the kernel-based method (KBM) in [5]. Tables 3 and 4 separately list the estimated values of two methods under different functional forms, as well as their MSE and SD. It is shown that our method is a little better than the KBM.

Table 3 The estimated value and accuracy of two methods in $f(x)=4\mathrm{e}^{-x^2}+ax$

a	Method	$\widehat{\theta}$	$\widehat{\phi}_1$	$\widehat{\phi}_2$	$\widehat{\rho}$	h_n	MSE	SD
0.1	our	3.981702	0.9834588	−0.4958761	0.3983633	0.06	0.000768543	0.02772261
	KBM	3.979881	0.9800375	−0.4898236	0.3963883	0.04	0.000932357	0.03053452
0.2	our	3.964253	0.9756316	−0.4813675	0.3892175	0.08	0.002677549	0.05174504
	KBM	3.960278	0.9698756	−0.4776379	0.3805721	0.04	0.003938577	0.06275808
0.3	our	3.958722	0.9727935	−0.4808183	0.3851976	0.08	0.007429996	0.08619742
	KBM	3.949898	0.9649873	−0.4780198	0.3800988	0.05	0.009063409	0.09520194

§5. Conclusion

This paper mainly studies the third order partially linear autoregressive model with first order autoregressive errors. We suppose that the regression function $f(x)$ has a parametric framework $\{g(x,\theta);\theta\in\Theta\}$, and through the conditional least squares method, we

Table 4 The estimated value and accuracy of two methods in $f(x) = 3\mathrm{e}^{-2x} + a\cos x$

a	Method	$\widehat{\theta}_1$	$\widehat{\theta}_2$	$\widehat{\phi}_1$	$\widehat{\phi}_2$	$\widehat{\rho}$	h_n	MSE	SD
0.1	our	2.990853	−1.965372	0.9987291	−0.4944566	0.4089553	0.10	0.06550958	0.2559484
	KBM	2.983508	−1.957312	0.9877982	−0.4888532	0.4093445	0.08	0.13531774	0.3678556
0.2	our	2.985854	−1.899653	0.9892157	−0.4909882	0.4092574	0.10	0.06844986	0.2616292
	KBM	2.935761	−1.858776	0.9801572	−0.4830928	0.4099735	0.07	0.12174942	0.3489261
0.3	our	2.951269	−1.866859	0.9886543	−0.4897532	0.4085658	0.06	0.08145388	0.2854013
	KBM	2.947983	−1.860035	0.9879238	−0.4807645	0.4108756	0.07	0.14597015	0.3820604

get the parametric estimator $g(x,\widehat{\theta})$. Semiparametric estimators $g(x,\widehat{\theta})\widehat{\xi}(x)$ and $g(x,\widehat{\theta})\widetilde{\xi}(x)$ can be given under the use of nonparametric kernel function method. Furthermore, under certain conditions, we prove the consistency of the estimators. Finally, simulation research under two types of models are presented to evaluate the effectiveness of the proposed method in theory.

In recent decades, the research work of nonlinear time series has been deepened. The theoretical studies have become increasingly completed and applied research has been more widely used. However, there are still many problems need to be studied urgently, such as the variable selection in the model (1). In the modeling, whether it is missing variables or selecting explanatory variables which are weakly or even independent of response variables, or there are collinearity between selected variables, will have great influence on the explanatory power and accuracy of the models. Therefore, it is necessary to establish robust models with simple structure, clear meaning and accurate prediction in selecting appropriate variables. That is also the research area way forward.

References

[1] CARROLL R J, FAN J Q, GIJBELS I, et al. Generalized partially linear single-index models [J]. *J Amer Statist Assoc*, 1997, **92(438)**: 477–489.

[2] XIA Y C, TONG H, LI W K. On extended partially linear single-index models [J]. *Biometrika*, 1999, **86(4)**: 831–842.

[3] GAO J T. *Nonlinear Time Series: Semiparametric and Nonparametric Methods* [M]. Boca Raton: Chapman and Hall/CRC, 2007.

[4] GAO J T, LIANG H. Asymptotic normality of pseudo-LS estimator for partly linear autoregression models [J]. *Statist Probab Lett*, 1995, **23(1)**: 27–34.

[5] GAO J T, YEE T. Adaptive estimation in partially linear autoregressive models [J]. *Canad J Statist*, 2000, **28(3)**: 571–586.

[6] YU Z X, WANG D H, SHI N Z. Semiparametric estimation of regression functions in autoregressive models [J]. *Statist Probab Lett*, 2009, **79(2)**: 165–172.

[7] FARNOOSH R, MORTAZAVI S J. A semiparametric method for estimating nonlinear autoregressive model with dependent errors [J]. *Nonlinear Anal*, 2011, **74(17)**: 6358–6370.

[8] FARNOOSH R, HAJEBI M, MORTAZAVI S J. A semiparametric estimation for regression functions in the partially linear autoregressive time series mode [J]. *Appl Appl Math*, 2014, **9(2)**: 573–591.

[9] AMIRI J, FARNOOSH R, BEHZADI M H, et al. A sequential semiparametric estimation for nonlinear function in the functional autoregressive model [J]. *Indian J Sci Technol*, 2016, **9(S1)**: 1–8.

[10] HJORT N L, GLAD I K. Nonparametric density estimation with a parametric start [J]. *Ann Statist*, 1995, **23(3)**: 882–904.

[11] HJORT N L, JONES M C. Locally parametric nonparametric density estimation [J]. *Ann Statist*, 1996, **24(4)**: 1619–1647.

[12] NAITO K. Semiparametric density estimation by local L_2-fitting [J]. *Ann Statist*, 2004, **32(3)**: 1162–1191.

[13] KLIMKO L A, NELSON P I. On conditional least squares estimation for stochastic processes [J]. *Ann Statist*, 1978, **6(3)**: 629–642.

[14] DAVYDOV YU A. Mixing conditions for Markov chains [J]. *Theory Probab Appl*, 1974, **18(2)**: 312–328.

[15] STOUT W F. *Almost Sure Convergence* [M]. New York: Academic Press, 1974.

[16] BERKES I, CSÁKI E. A universal result in almost sure central limit theory [J]. *Stochastic Process Appl*, 2001, **94(1)**: 105–134.

部分线性自回归模型中回归函数的半参数估计

王明辉[1,2]　柳向东[2]　李　银[1]

([1]韶关学院数学与统计学院, 韶关, 512005;　[2]暨南大学经济学院, 广州, 510632)

摘　要: 本文主要研究具有一阶自回归误差的三阶部分线性自回归模型中回归函数的半参数估计问题. 假定回归函数来自某个参数分布族, 利用条件最小二乘法得到参数估计量, 再结合非参数核函数进行调整, 给出回归函数的半参数估计量. 并在一定条件下, 证明了估计量具有相合性. 最后, 通过模拟研究验证了此方法的有效性.

关键词: 部分线性自回归模型; 一阶自回归误差; 条件最小二乘法; 核函数调整; 半参数估计

中图分类号: O212.2