C. R. Acad. Sci. Paris, Ser. I 348 (2010) 211–215

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Statistics

Optimal and superoptimal convergence rate of the local linear estimator of nonparametric regression function in continuous time ☆

L'optimalité et la suroptimalité du lissage localement linéaire de la fonction de régression

Jia Shen, Shuguang Sun

Department of Statistics, Fudan University, Shanghai 200433, China

ARTICLE INFO

Article history:
Received 9 June 2009
Accepted after revision 4 December 2009
Available online 30 December 2009

Presented by Paul Deveuvels

ABSTRACT

We consider the estimation problem of the nonparametric regression in continuous time by the local linear estimator in the asymptotic quadratic error sense. In suitable conditions of strongly mixing and that of irregularity, we obtained optimal and superoptimal convergence rate of the estimator.

RÉSUMÉ

Nous étudions l'estimation non paramétrique de la fonction de régression par le lissage localement linéaire au sens d'erreur quadratique asymptotique. Sous des conditions de forte mélangeance et d'irrégularité, nous obtenons des vitesses de convergence optimale et suroptimale de l'estimateur pour l'erreur quadratique asymptotique.

Version française abrégée

Considérons la relation $E[Y \mid X = x] = m(x)$ entre deux processus en temps continu à valeurs réelles et strictement stationnaires $\{X_t, t \in R\}$ et $\{Y_t, t \in R\}$. Nous construisons l'estimateur non paramétrique de la fonction de régression par le lissage localement linéaire (2) qui minimise la quantité (1) par rapport aux paramètres a et b, à partir de l'observation d'une trajectoire $\{(X_t, Y_t), t \in [0, T]\}$.

L'objectif de cette Note est d'obtenir les vitesses de convergence en moyenne quadratique optimale et suroptimale de l'estimateur (2) dans un cadre similaire que Cheze–Payaud [9] a étudié pour l'estimatuer par noyau. Pour ce faire, nous supposons que les processus sont fortement géometriquement mélangeants et satisfont à certaines conditions d'irrégularité. Les démonstrations ont été faites suivant les décompositions de Fan [10] et les techniques de majoration en cas de processus en temps continu de Bosq [4]. Finalement, nous donnons un exemple qui vérifie les hypothèses.

1. Local linear estimator of nonparametric regression: Notations and assumptions

Suppose the following relationship between two real processes $\{X_t, t \in R\}$ and $\{Y_t, t \in R\}$:

☆ Supported partly by Natural Science Foundation of China (Project Number: 70832005) and Shanghai Leading Academic Discipline Project, Project Number: B210.

E-mail address: jiashen@fudan.edu.cn (J. Shen).

1631-073X/$ – see front matter © 2009 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.crma.2009.12.005

$$E(Y \mid X = x) = m(x).$$

In discrete time case, the local linear estimator is studied by number of authors, see Fan [10,11] and Fan and Gijbels [12]. In continuous time the estimation problem by kernel estimator and orthogonal projection method is studied by Cheze [8], Bosq and Cheze [6], Bosq [3,4], Bosq and Blanke [5]. In comparison with the discrete time case, that of continuous time has the particularity of superoptimal convergence rate under suitable conditions.

For the continuous time sample $\{(X_t, Y_t), t \in [0, T]\}$, and the nonparametric regression

$$m(x) = E[Y \mid X = x],$$

consider the minimizer of the following function with relation to parameters a and b

$$Q(a, b) = \int_0^T \left(Y_t - a - b(x - X_t)\right)^2 K\left(\frac{x - X_t}{h_T}\right) \mathrm{d}t, \tag{1}$$

where $K(\cdot)$ is the kernel function and h_T is the bandwidth. We obtain then an estimator of the unknown function $m(x)$:

$$\hat{a} = \frac{\int_0^T Y_t w_t \,\mathrm{d}t}{\int_0^T w_t \,\mathrm{d}t},$$

where

$$w_t = K\left(\frac{x - X_t}{h_T}\right)\left[s_{T,2} - (x - X_t)s_{T,1}\right],$$

$$s_{T,l} = \int_0^T K\left(\frac{x - X_t}{h_T}\right)(x - X_t)^l \,\mathrm{d}t, \quad l = 0, 1, 2.$$

It is easy to see that $\int_0^T w_t \,\mathrm{d}t = s_{T,0}s_{T,2} - (s_{T,1})^2 \geqslant 0$. For avoiding the vanishing of the denominator, we replace the estimator $\hat{a}$ by

$$\widehat{m}(x) = \frac{\int_0^T w_t Y_t \,\mathrm{d}t}{\int_0^T w_t \,\mathrm{d}t + T^{-2}}, \tag{2}$$

where T^{-2} in the denominator does not affect the asymptotic behavior of the estimator.

Assumption H.

(1) The function $m(\cdot)$ is twice differentiable with $\sup_{x \in \mathbb{R}} |m''(x)| = B_m < \infty$;
(2) The process $Z = \{Z_t = (X_t, Y_t),\ t \in R\}$ is strictly stationary and geometrically strongly mixing (GSM):

$$\alpha(u) \triangleq \sup_{\substack{A \in \sigma(Z_s, s \leqslant 0) \\ B \in \sigma(Z_s, s \geqslant u)}} \left|P(A \cap B) - P(A)P(B)\right| \leqslant c\rho^u,$$

where $\rho \in (0, 1)$, $c, u > 0$;
(3) The density $f_X(\cdot)$ of X_0 is Hölder continuous:

$$\left|f_X(x) - f_X(y)\right| \leqslant c|x - y|^\alpha, \quad 0 < \alpha \leqslant 1;$$

(4) The process $\{\varepsilon_t = Y_t - m(X_t),\ t \in R\}$ satisfies $E[\varepsilon_t | \mathcal{B}_T] = 0$, a.s., and

$$\sup_{0 \leqslant t \leqslant T < +\infty} E\left(|\varepsilon_t|^{2+\delta} | \mathcal{B}_T\right) \leqslant M^{2+\delta}, \quad \text{a.s.},$$

for some $\delta > 0$, $M > 0$, and $\mathcal{B}_T = \sigma(X_t, t \leqslant T)$.
(5) The kernel function $K(\cdot)$ is a non-negative, symmetric and bounded density, satisfying

$$0 < \int_{-\infty}^{+\infty} y^{2r} K(y) \,\mathrm{d}y < \infty, \quad \text{for integer } r > 0.$$

The following function $g_{(X,Y)}(x,y)$ is a measure of dependence between X and Y. Let (X,Y) be $\mathbb{R}\times\mathbb{R}$ valued random variable, denote

$$g_{(X,Y)}(x,y) = f_{(X,Y)}(x,y) - f_X(x) f_Y(y); \quad x, y \in \mathbb{R}^1,$$

where f_Z is the density of Z.

Let $\{\xi_t, t \in R\} = \{(X_t, \varepsilon_t), t \in R\}$ be strictly stationary and consider

$$g^*_{s,t} = g_{(\xi_s,\xi_t)} = f_{(\xi_s,\xi_t)} - f_\xi \otimes f_\xi,$$

and

$$G^*_{s,t}(x,x') = \int_{\mathbb{R}^2} yy' g^*_{s,t}(x,y;x',y')\,\mathrm{d}y\,\mathrm{d}y',$$

where $(x,x') \in \mathbb{R}^2$, $(f\otimes f)(x,y) = f(x)f(y)$. Consider now

$$g^*_{s,t} = g^*_{|s-t|}, \quad s \neq t,$$

and let

$$H^*(x,x') = \int_{]0,+\infty[} \left|G^*_u(x,x')\right|\mathrm{d}u.$$

2. Asymptotic optimal rate of convergence

First we present the optimal convergence rate of the estimator $\widehat{m}(x)$ defined by (2).

Assumption $A^*(\Gamma, p)$. There exist $\Gamma \in \mathcal{B}_{\mathbb{R}^2}$, containing $D = \{(s,t) \in \mathbb{R}^2\colon\ s = t\}$ and $p \in]2,+\infty]$ such that

(1) for all $(s,t) \notin \Gamma$, $G^*_{s,t}$ exists;
(2) $\Delta_p(\Gamma) = \sup_{(s,t)\notin\Gamma} \|G^*_{s,t}\|_{L^p(\mathbb{R}^2)} < \infty$;
(3) $\limsup_{T\to+\infty} \frac{1}{T}\int_{[0,T]^2\cap\Gamma} \mathrm{d}s\,\mathrm{d}t = \ell_\Gamma < \infty$.

Theorem 1. *Under Assumptions H and $A^*(\Gamma,p)$, let $h_T = dT^{-\beta}$, $0<\beta<1/2$, then we have*

$$\begin{aligned} E\big(\widehat{m}(x) - m(x)\big)^2 \leqslant{}& \frac{1}{4}\left(B_m \int_{-\infty}^{\infty} u^2 K(u)\,\mathrm{d}u\right)^2 h_T^4 \\ &+ \frac{1}{Th_T} f_X^{-2}(x) E\left[\frac{\varepsilon_0^2}{h_T} K^2\left(\frac{x - X_0}{h_T}\right)\right] \cdot \frac{1}{T} \int_{[0,T]^2\cap\Gamma} \mathrm{d}s\,\mathrm{d}t \\ &+ o\left(h_T^4 + \frac{1}{Th_T}\right), \quad \forall x \in \big\{x\colon\ f(x) > 0\big\}. \end{aligned} \tag{3}$$

If the second order of derivative of $m(\cdot)$ is continuous at x, then the bound B_m is replaced by $m''(x)$ in (3).

Remark 1. Taking $\beta = -1/5$, we have the optimal convergence rate. That is

$$T^{4/5} E\big(\widehat{m}(x) - m(x)\big)^2 = O(1).$$

This convergence rate was obtained by Cheze-Payaud [9] in a similar framework by kernel estimation method for $\mathbb{R}^d \times \mathbb{R}$ valued $\{(X_t, Y_t)\}$ with the rate $T^{4/(d+4)}$ instead of $T^{4/5}$ (see Bosq [4]).

3. Asymptotic superoptimal rate of convergence

The following two theorems state the superoptimal rate of the estimator $\widehat{m}$ and the asymptotic limit of the quadratic error. The following is the irregularity condition:

Assumption I. H^* exists, bounded and continuous at (x,x).

It is an analog of the Castellana–Leadbetter's [7] irregularity condition on the regression problem.

Theorem 2. *Under Assumptions H and I, let* $h_T = dT^{-\beta}$, $0 < \beta < 1/3$, *then there is* $0 < \tau \leqslant \delta$, *such that*

$$E\big(\widehat{m}(x) - m(x)\big)^2 = O\left(h_T^4 + \frac{h_T^{2\alpha}}{T h_T^{2-2/(2+\tau)}} + \frac{1}{T}\right), \tag{4}$$

especially, for $1/2 < \alpha \leqslant 1$, *and* $h_T = dT^{-\beta}$, $1/4 < \beta < 1/3$, *for* $x \in \{x\colon\ f(x) > 0\}$,

$$\limsup_{T\to\infty} T E\big(\widehat{m}(x) - m(x)\big)^2 \leqslant 2 f_X^{-2}(x) \int_0^{+\infty} \big|G_u^*(x,x)\big|\,\mathrm{d}u. \tag{5}$$

Theorem 3. *Under Assumption H, if* $G_{s,t}^* = G_{|s-t|}^*$, $u \mapsto \|G_u^*\|_\infty$ *is integrable over* $]0,+\infty[$, *and* $\forall u > 0$, G_u^* *is continuous at* (x,x), *then when* $1/2 < \alpha \leqslant 1$, *take* $h_T = o(T^{-1/4})$, *we have*

$$T E\big(\widehat{m}(x) - m(x)\big)^2 \xrightarrow[T\to+\infty]{} 2 f_X^{-2}(x) \int_0^{+\infty} G_u^*(x,x)\,\mathrm{d}u. \tag{6}$$

The proofs of the above theorems mainly follow decompositions of Fan [10] and the techniques of Bosq [4].

4. Example

Let $X = \{X_t, t \in R\}$ and $\{\varepsilon_t, t \in R\}$ be two independent, strictly stationary processes. When $\{X_t\}$ is geometrically strongly mixing and Gaussian processes, with zero-mean with autocorrelation coefficient $\rho(u)$ satisfying $\rho(u) = 1 - a|u|^\theta + o(u^\theta)$, $u \to 0$, where $0 < \theta < 2$. Consider $\{\varepsilon_t, t \in \mathbb{R}\}$ the continuous-time fractional ARMA process introduced by Viano et al. [13] (see also Blanke [1]). It is defined by $\varepsilon_t = \int_{-\infty}^t f(t-s)\,\mathrm{d}W(s)$, with the Brownian motion W and the impulse response function $f \in L^2(\mathbb{R}^+)$ and its Laplace transform $F(s) = \prod_{k=1} K(s-a_k)^{d_k}$, for $\mathrm{Re}(s) > a$, where for $D = \sum_1^K d_k$ and the index set E^* of singular points of F, $a = \max\{\mathrm{Re}(a_k), k \in E^*\}$. When $D < -1/2$ and $a < 0$, the process $\{\varepsilon_t\}$ is a zero-mean, stationary, strongly mixing and Gaussian process. It is easy to see that then,

$$G_{0,u}^*\big(x,x'\big) = f_{(X_0,X_u)}\big(x,x'\big)\,\mathrm{Cov}(\varepsilon_0, \varepsilon_u).$$

Since for $D \leqslant -3/2$ and $a < 0$, the autocovariance function $\sigma_\varepsilon(u)$ of the process $\{\varepsilon_t\}$ is bounded over $]0,+\infty[$, Assumption $A^*(\Gamma, p)$ and the irregularity condition I are reduced to the following assumptions on the process $\{X_t\}$. Denote $g_{s,t} = g_{(X_s,X_t)}$.

Assumption $A(\Gamma, p)$. There is $\Gamma \in \mathcal{B}_{\mathbb{R}^2}$, containing $D = \{(s,t) \in \mathbb{R}^2\colon\ s = t\}$. There exists $p \in]2,+\infty]$ such that

(1) for all $(s,t) \notin \Gamma$, $g_{s,t}$ exists;
(2) $\delta_p(\Gamma) = \sup_{(s,t)\notin\Gamma} \|g_{s,t}\|_{L^p(\mathbb{R}^2)} < \infty$;
(3) $\limsup_{T\to+\infty} \frac{1}{T}\int_{[0,T]^2\cap\Gamma} \mathrm{d}s\,\mathrm{d}t = \ell_\Gamma < \infty$.

Assumption I′. The function $H(x,x') = \int_{]0,+\infty[} |g_u(x,x')|\,\mathrm{d}u$ exists, bounde dand continuous at (x,x).

They are studied in Bosq [4], Blanke and Bosq [2] and Bosq and Blanke [5]. Especially, Blanke and Bosq [2] consider the case of $f_{(X_0,X_u)}(\cdot,\cdot) = M(\cdot,\cdot)u^{-\gamma_0}$ with either $M(\cdot,\cdot)$ bounded or $M(\cdot,\cdot) \in L^1(\mathbb{R}^{2d})$ and then the irregularity condition of the process matches with $0 < \gamma_0 < 1$.

Acknowledgements

We are thankful to Professor Denis Bosq for his very helpful suggestions and his continuous encouragements. Our gratitude also goes to all reviewers for their careful reading, corrections and suggestions which improved the manuscripts.

References

[1] D. Blanke, Estimation of local smoothness coefficients for continuous time processes, Stat. Inference Stoch. Process. 5 (1) (2002) 65–93.
[2] D. Blanke, D. Bosq, Regression estimation and prediction in continuous time, J. Japan Statist. Soc. (2008).
[3] D. Bosq, Parametric rates of noparametric estimators and predictors for continuous time processes, Ann. Statist. 25 (3) (1997) 982–1000.
[4] D. Bosq, Nonparametric Statistics for Stochastic Processes, Lecture Notes in Statist., vol. 110, Springer-Verlag, New York, 1998.
[5] D. Bosq, D. Blanke, Inference and Prediction in Large Dimensions, Wiley Ser. Probab. Stat., vol. 754, Wiley-Dunod, Chichester, England, 2007.

[6] D. Bosq, N. Cheze, Erreur quadratique asymptotique optimale de l'estimateur non paramétrique de la régression pour des observations discrétisées d'un processus stationaire à temps continu, C. R. Acad. Sci., Paris Sér. I 317 (9) (1993) 891–894.
[7] J.V. Castellana, M.R. Leadbetter, On smoothed probability density estimation or stationary processes, Stochastic Process. Appl. 21 (1986) 179–193.
[8] N. Cheze, Régression non paramétrique pour un processus à temps continu, C. R. Acad. Sci., Paris Sér. I 315 (1992) 1009–1012.
[9] Nathalie Cheze–Payaud, Régression, prédiction et discrétisation des processus à temps continu, PhD thesis, Université Pierre et Marie Curie (Paris 6), 1994.
[10] J. Fan, Design-adaptive nonparametric regression, J. Amer. Statist. Assoc. 21 (1992) 196–216.
[11] J. Fan, Local linear regression smoothers and their minimax efficiency, Ann. Statist. 21 (1993) 196–216.
[12] J. Fan, I. Gijbels, Local Polynomial Modelling and Its Applications, Chapman & Hall, London, 1996.
[13] M.C. Viano, C. Deniau, G. Oppenheim, Continuous-time fractional ARMA processes, Statist. Probab. Lett. 21 (1994) 323–336.