Journal of the Korean Statistical Society 44 (2015) 632–645

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

Sharp optimality for regression with real-time data

Zhi-Ming Luo [a], Jeongcheol Ha [a,*], Tae Yoon Kim [a], Inho Park [b]

[a] Department of Statistics, Keimyung University, Daegu 704-701, South Korea

[b] Department of Statistics, Pukyong National University, Busan 608-737, South Korea

ARTICLE INFO

Article history:
Received 5 November 2014
Accepted 22 April 2015
Available online 14 May 2015

AMS 2000 subject classifications:
primary 62G20
secondary 62G05

Keywords:
Minimax
Optimal rate of convergence
Exact constant of convergence
Real-time data

ABSTRACT

A minimax estimator for a nonparametric regression model is proposed when real-time data are assumed and its asymptotic behavior of minimax risk in the sup-norm for the Hölder function class is studied. The optimal rate of convergence and exact minimax constant are found for the estimator.

1. Introduction and the main result

Minimax estimation is one of the most powerful statistical theories which seeks the best estimator in terms of minimax risk for a given model setting. The optimal rates of convergence in terms of minimax risk have been obtained for various model settings (Ibragimov & Hasminskii, 1981, 1982; Koo, 1993; Stone, 1982; Tsybakov, 2004), however, only a few exact constants of convergence in a 'sharp' asymptotical minimax sense are obtained. The first exact constant is the Pinsker constant (Pinsker, 1980) in L_2 minimax risk for Sobolev function class under the Gaussian white noise model. Korostelev (1993) obtained the exact constant of optimal convergence for the Gaussian nonparametric regression model with fixed design. Later Korostelev and Nussbaum (1999) obtained the constant for nonparametric density estimation and Bertin (2004) obtained the constant for the nonparametric regression model with random design.

The nonparametric regression model considered by Korostelev (1993) is

$$Y_i = f(i/n) + \xi(i), \quad i = 1, \dots, n, \tag{1}$$

where the ξ's are independent Gaussian $(0, \sigma^2)$ random variables. In this model, the design points become closer among themselves as the sample size increases; yet, the error process remains the same. As Kim and Luo (2010) pointed out, the model (1) has some limitations, however. Indeed (i) it assumes a fixed bounded time domain [0, 1] and hence does not allow the time flow to infinity, and (ii) the error dependence structure is independent of the distance between the design points. They notice that these limitations are to be resolved by employing the regression model with real-time data because real-time data are defined as a sequence of observations made on the finer grid of time going to infinity. By keeping this in mind, they proposed an extended model

$$Y_i = f(i/n^a) + \xi(i/n^{a\theta}), \quad i = 1, \dots, n, \tag{2}$$

* Corresponding author.
E-mail addresses: zhimingluo@kmu.ac.kr (Z.-M. Luo), jeicy@kmu.ac.kr (J. Ha), tykim@kmu.ac.kr (T.Y. Kim), ipark@pknu.ac.kr (I. Park).

http://dx.doi.org/10.1016/j.jkss.2015.04.004
1226-3192/

where $f(\cdot)$ is an unknown nonparametric regression function with domain $(0, \infty)$ and $\{\xi(t)\}$ is a zero-mean strictly stationary Gaussian sequence with a variance–covariance structure, $\mathrm{cov}[\xi(i/n^{a\theta}), \xi(j/n^{a\theta})] = \sigma^2\rho^{|j-i|/n^{a\theta}}$ where $0 < \rho < 1$, $0 < a \le 1$, and $0 \le \theta < 1$. As an example of the processes under consideration of this paper, the following is provided.

Example. For given n, consider the Ornstein–Uhlenbeck velocity process X_t on $(0, \infty)$ which satisfies the stochastic differential equation (Arnold, 1974)

$$dX_t = -n^{(1-\theta)a}X_t dt + \sqrt{2n^{(1-\theta)a}\sigma^2}dW_t$$

with an initial random variable $X_0 \sim N(0, \sigma^2)$, where W_t is the standard Wiener process. Then X_t is a stationary Gaussian process with so-called colored noise

$$EX_t = 0,$$
$$EX_tX_s = \sigma^2 e^{-n^{(1-\theta)a}|t-s|}.$$

Now the Ornstein–Uhlenbeck position process Y_t is defined as

$$Y_t = Y_0 + \int_0^t f(s)ds + \int_0^t X_s ds,$$

where the initial random variable Y_0 is a constant and f is a smooth function. For a small increment $\Delta > 0$ we have

$$\frac{Y_{t+\Delta} - Y_t}{\Delta} = \frac{1}{\Delta}\int_t^{t+\Delta} f(s)ds + \frac{1}{\Delta}\int_t^{t+\Delta} X_s ds.$$

Letting

$$y(t) = \frac{Y_{t+\Delta} - Y_t}{\Delta} \quad \text{and} \quad \xi(t) = X_t,$$

we have an approximation

$$y(t) \approx f(t) + \xi(t).$$

Again letting $y_i = y(i/n^a)$ and $\xi_i = \xi(i/n^a)$ $(i = 1, \ldots, n)$ with frequency n^a from $[0, 1]$ we have an approximation

$$y_i \approx f(i/n^a) + \xi_i$$

which follows model (2) with variance–covariance structure, $\mathrm{cov}[\xi_i, \xi_j] = \sigma^2 e^{-|j-i|/n^{a\theta}}$.

It is easy to see that model (2) not only allows the time flow to infinity but the dependence structure might also be directly influenced by the distance between the design points. Note here that $\theta = 0$ indicates that the error dependence structure is free of the design points, whereas $\theta = 1$ indicates that the error dependence structure is entirely dependent on the design points. It turns out that the nonparametric regression estimator is not applicable when $\theta = 1$ or $a = 0$. Refer to Kim and Luo (2010).

In this paper, we will focus on the minimax property of the nonparametric regression estimator of the unknown function f for model (2). Let β and L be some positive constants and $\ell = \lfloor \beta \rfloor$ be the greatest integer strictly less than β. The Hölder class of functions $\Sigma(\beta, L)$ is defined as

$$\Sigma(\beta, L) = \left\{f : |f^{(\ell)}(x) - f^{(\ell)}(x')| \le L|x - x'|^{\beta-\ell},\ \forall x, x' \in [0, \infty]\right\}.$$

Suppose that the function $f \in \Sigma(\beta, L) \cap \mathcal{S}$ where

$$\mathcal{S} = \left\{ f(x) \left| \begin{cases} |f(x)| \le \dfrac{C}{(1-a)^{\beta/(2\beta+1)}} \left(\dfrac{\log x}{x^{a(1-\theta)/(1-a)}}\right)^{\beta/(2\beta+1)}, & \text{as } x \to \infty,\ \text{if } 0 < a < 1 \\ f(x) = 0, & \text{for } x > 1,\ \text{if } a = 1 \end{cases} \right. \right\}$$

and C above is the exact constant of convergence specified in Theorem 1. $\Sigma(\beta, L) \cap \mathcal{S}$ is a quite generalization of

$$\Sigma_1(\beta, L) = \left\{f : |f^{(\ell)}(x) - f^{(\ell)}(x')| \le L|x - x'|^{\beta-\ell},\ \forall x, x' \in [0, 1]\right\}$$

imposed by Korostelev (1993) since $\Sigma_1(\beta, L) \subset \Sigma(\beta, L) \cap \mathcal{S}$. Note that $\mathcal{S}$ consists of functions bounded by a rapidly decreasing function on $(0, \infty)$ while $\Sigma_1(\beta, L)$ is a set of integrable functions on $[0, 1]$.

Let the loss function $\omega(u)$, $u \ge 0$, be a continuous increasing function such that $\omega(0) = 0$, $\omega \ne 0$ and $\omega(u) \le W_0(1+u^\gamma)$ with some positive constants W_0 and γ, i.e., $\omega(u)$ has a polynomial upper bound. We introduce the minimax risk

$$R_n = R_n(\omega(\cdot); \beta, L; a, \theta, \rho, \sigma^2) = \inf_{\widehat{f}_n} \sup_{f \in \Sigma(\beta,L)\cap\mathcal{S}} \mathbb{E}_f\omega(\psi_n^{-1}\|\widehat{f}_n - f\|_\infty) \tag{3}$$

where the infimum is taken over all estimators of f based on observations $Y_1, \ldots, Y_n$, $\mathbb{E}_f$ means the expectation with respect to the probability measure of observations $Y_1, \ldots, Y_n$ depending on f, ψ_n is the optimal rate of convergence to be specified later and $\|\cdot\|_\infty$ is the sup-norm given by $\|f\|_\infty = \sup_{x\ge0}|f(x)|$.