Adaptive estimators for nonparametric heteroscedastic regression models

Jean-Yves Brua

Journal of Nonparametric Statistics
Vol. 00, No. 00, Month 200x, 1–14

RESEARCH ARTICLE

Adaptive estimators for nonparametric heteroscedastic regression models

J.-Y. BRUA†

IRMA, Département de Mathématique, Université Louis Pasteur, 7 rue René Descartes, F67084, Strasbourg Cedex, France

(*Received 00 Month 200x; in final form 00 Month 200x*)

This paper deals with the estimation of a regression function at a fixed point in nonparametric heteroscedastic regression models with Gaussian noise. We assume that the variance of the noise depends on the regressor and on the regression function. We make use of the minimax absolute error risk taken over a local weak Hölder class of regression functions. As the smoothness of the regression function is supposed to be unknown, we construct an adaptive kernel estimator which attains the minimax rate. More precisely we give an asymptotic upper bound for its risk and an asymptotic lower bound for the minimax risk.

Keywords: Adaptive estimation; Heteroscedastic regression; Kernel estimator; Minimax; Nonparametric regression

MSC: primary 62G08; secondary 62G05, 62G20

† Phone number: +33 390 240 278. Email: brua@math.u-strasbg.fr

ISSN: 1048-5252 print/ISSN 1029-0311 online

DOI: 10.1080/1048525YYxxxxxxxx
http://www.informaworld.com

1. Introduction

A lot of research studies are devoted to the problem of estimating a regression function and specially a function belonging to a Hölder class. In this direction it is known from [1] that the linear minimax estimator is a kernel estimator in the case of a quasi-Hölder regression function estimated at a single point with squared error loss. This estimator is within 17 percent of asymptotically minimax over all procedures (see [2]). Furthermore it is proved in [3, 4] that a kernel estimator is asymptotically efficient when the Hölder regression function or its kth derivative is estimated with the sup-norm global loss. The reader is referred for instance to [5–7] for other regression works.

Our regression problem is the following. Suppose we observe data from:

$$y_k = S(x_k) + g(x_k, S)\xi_k, \qquad k \in \{1, \dots, n\}, \tag{1}$$

where $x_k = k/n$, $(\xi_k)_{k\in\{1,\dots,n\}}$ are independent identically distributed standard Gaussian random variables. We are interested in the estimation of the regression function S at a given point $z_0 \in]0;1[$. We point out the fact that in this heteroscedastic regression model the variance of the noises depends on the unknown function S and on the regressors x_k. This kind of model is used in financial analysis or in medical research (see, e.g., [8]). More recently we can find such a model in [9] where the authors reduce their classical regression model with variance V depending only on the regressors to a specific regression model where the regression function is nearly V and the noise term depends on V.

We assume that the regression function belongs to a Hölder class but its smoothness parameter β remains unknown. We consider the absolute error loss and for the corresponding risk for which we aim at constructing an adaptive estimator which attains the minimax rate. Because of the adaptation, this rate differs from the one in the case where β is known. Many papers deal with adaptive problems, see for instance [10–14]. Our construction is based on those one can find in [12, 15] for adaptive estimation of the drift coefficient in diffusion processes. We propose an adaptive kernel estimator which attains the minimax rate, that is to say that its risk has an asymptotic finite upper bound whereas the minimax risk is bounded away from zero. If these two bounds coincide we say that the estimator is asymptotically efficient.

We proceed with the method developed in [16] in the homoscedastic and non adaptive case. In this paper the risk of an estimator is defined as the supremum of the absolute error loss taken over a neighborhood (called weak Hölder class) of functions that allows an arbitrary large derivative but has an additional weak Hölder condition (see (2)). It has led to the heteroscedastic case studied in [17]. In these cases it is shown that a kernel estimator is asymptotically efficient, with the minimax rate $n^{\beta/(2\beta+1)}$. Here we found the same asymptotic lower bound for the minimax risk as in the non adaptive case with the minimax rate $\left(\frac{n}{\ln n}\right)^{\beta/(2\beta+1)}$, but unfortunately not the same asymptotic upper bound for the risk of the adaptive kernel estimator.

The paper is laid out as follows. Section 2 gives the description of the problem with all assumptions needed and all definitions of necessary mathematical objects. In section 3 we construct an adaptive estimator for which an upper bound of the risk is found. The lower bound of the minimax risk is given in section 4. Appendix A contains technical results for our proofs.

2. Statement of the problem

Consider model (1) where $S \in C^1([0;1],\mathbb{R})$ and $g : [0;1] \times C^1([0;1],\mathbb{R}) \longrightarrow \mathbb{R}_+^*$ are unknown functions. We want to estimate the regression function S at a fixed point $z_0 \in]0;1[$. The problem assuming that

$$S \in \bigcup_{M,K>0} \mathcal{H}(M,K,\beta), \text{where } \beta \in [\beta_\star;\beta^\star] \subset]1;2[,$$

$$\mathcal{H}(M,K,\beta) = \left\{ S \in C^1[0;1] : ||S'|| \leq M, \sup_{x,y\in[0;1]} \frac{|S'(y)-S'(x)|}{|x-y|^\alpha} \leq K \right\},$$

with $\beta = 1+\alpha$ and $||f|| = \sup_{x\in[0;1]} |f(x)|$, remains open.

As we can find in [17] and [16] where non adaptive problems of asymptotically efficient estimation for nonparametric regression models are solved, we define the local weak Hölder class $\mathcal{U}_{z_0,\delta}$ at the point z_0 for the true value of the parameter β as:

$$\left\{ S \in C^1([0;1],\mathbb{R}) :\| S' \| \leq \delta^{-1}; \forall h \geq 0, \left| \int_{-1}^{1} \big(S(z_0+hu)-S(z_0)\big)du \right| \leq \delta h^\beta \right\}, \tag{2}$$

where $\delta \in]0;1[$.
The smoothness parameter β is supposed to be unknown whereas the interval $[\beta_\star;\beta^\star]$ is considered as known.
Notice that

$$\int_{-1}^{1} \big(S(z_0+hu)-S(z_0)\big)du = \int_{-1}^{1} \left(\int_{z_0}^{z_0+uh} (S'(t)-S'(z_0))dt \right) du, \tag{3}$$

so we have for all $S \in \mathcal{H}(M,K,\beta)$,

$$\left| \int_{-1}^{1} \big(S(z_0+hu)-S(z_0)\big)du \right| \leq \frac{2K}{\beta(\beta+1)} h^\beta.$$

That is why the class $\mathcal{U}_{z_0,\delta}$ is called a weak Hölder class.
The risk of an estimator $\hat{S}$ of $S(z_0)$ is defined over the neighborhood $\mathcal{U}_{z_0,\delta}$ by

$$\mathcal{R}_{z_0,\delta,n}(\hat{S}) = \sup_{\beta\in[\beta_\star;\beta^\star]} \sup_{S\in\mathcal{U}_{z_0,\delta}} \frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S|\hat{S}-S(z_0)|, \quad \text{where} \quad N(\beta) = \left(\frac{n}{\ln n}\right)^{\beta/(2\beta+1)}.$$

We assume that there exists two known constants $g_\star > 0$ and $g^\star < \infty$ such that

$$g_\star \leq \inf_{0\leq x\leq 1} \inf_{S\in C^1([0;1])} g(x,S) \leq \sup_{0\leq x\leq 1} \sup_{S\in C^1([0;1])} g(x,S) \leq g^\star.$$

Moreover we suppose that the function g is uniformly continuous with respect to both variables x and S.

3. Upper bound

The paper [16] handles the homoscedastic non adaptive case, considering the kernel estimator

$$S_h^*(z_0) = \frac{1}{q_n(h)} \sum_{k=1}^{n} Q\big(\frac{x_k - z_0}{h}\big) y_k,$$

where $Q = \mathbb{I}_{[-1,1]}$, $h = n^{-1/(2\beta+1)}$ and $q_n(h) = \sum_{k=1}^{n} Q\big(\frac{x_k - z_0}{h}\big)$. Taking into account the fact that β is unknown we can not use such an estimator because the bandwidth h depends on β. That is the reason why we create a partition of the interval $[\beta_\star; \beta^\star]$ in the following way:

$$\beta_l = \beta_\star + l\frac{\beta^\star - \beta_\star}{\ln n}, \quad l = 0, \dots, [\ln n],$$

where $[a]$ denotes the integral part of a number a, and we define the corresponding bandwidths $h_l = h(\beta_l) = \Big(\frac{n}{\ln n}\Big)^{-1/(2\beta_l+1)}$. Then we set

$$\hat{l} = \max\left\{0 \le l \le [\ln n] : \max_{0\le j\le l}\left(\Big|S_{h_l}^*(z_0) - S_{h_j}^*(z_0)\Big| - \frac{\lambda}{N_j}\right) \le 0\right\},$$

where $N_j = N(\beta_j)$ and $\lambda = 2 + 2\sqrt{2}g^\star\left(\frac{\beta^\star}{2\beta^\star+1} - \frac{\beta_\star}{2\beta_\star+1}\right)^{1/2}$.
Notice that $\hat{l}$ really exists because the set above contains the index 0.
The adaptive estimator is now defined as $\hat{S}_n = S_{h_{\hat{l}}}^*(z_0)$. Furthermore we associate with the unknown parameter β the unique integer $l(\beta) \in \{0, \dots, [\ln n] - 1\}$ such that $\beta_{l(\beta)} \le \beta < \beta_{l(\beta)+1}$.

The following result gives an upper bound for the risk of the adaptive estimator.

Theorem 3.1 *One has*

$$\limsup_{\delta\to 0} \limsup_{n\to\infty} \mathcal{R}_{z_0,\delta,n}(\hat{S}_n) \le \lambda \frac{e^{\beta^\star - \beta_\star}}{g_\star}.$$

Proof: Fix $\delta \in]0;1[$ and write

$$|\hat{S}_n - S(z_0)| = |\hat{S}_n - S(z_0)|\mathbb{I}_{\{\hat{l}\ge l(\beta)\}} + |\hat{S}_n - S(z_0)|\mathbb{I}_{\{\hat{l}< l(\beta)\}} =: I_1 + I_2,$$

where $I_1 = \sum_{j=l(\beta)}^{[\ln n]} |\hat{S}_n - S(z_0)|\mathbb{I}_{\{\hat{l}=j\}} = \sum_{j=l(\beta)}^{[\ln n]} |S_{h_j}^*(z_0) - S(z_0)|\mathbb{I}_{\{\hat{l}=j\}}$.

For all $j = 0, \ldots, [\ln n]$, we note $S^*_{h_j}(z_0) - S(z_0) = \dfrac{\zeta_n(\beta_j)}{\sqrt{q_n(\beta_j)}} + B_n(\beta_j)$, with

$$\begin{aligned} q_n(\beta_j) &= q_n(h(\beta_j)), \\ \zeta_n(\beta_j) &= \frac{1}{\sqrt{q_n(\beta_j)}} \sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_j}\right) g(x_k, S)\xi_k, \\ B_n(\beta_j) &= \frac{1}{q_n(\beta_j)} \sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_j}\right) (S(x_k) - S(z_0)). \end{aligned}$$

The variance of $\zeta_n(\beta_j)$ is then $\sigma_n^2(\beta_j, S) := \dfrac{1}{q_n(\beta_j)} \sum_{k=1}^{n} Q\left(\dfrac{x_k - z_0}{h_j}\right) g^2(x_k, S)$.

One has $\left|\dfrac{\zeta_n(\beta_j)}{\sqrt{q_n(\beta_j)}}\right| = |Z_n(\beta_j)| \dfrac{\sigma_n(\beta_j, S)}{\sqrt{q_n(\beta_j)}}$, with $Z_n(\beta_j) \sim \mathcal{N}(0, 1)$.
In addition, taking the decomposition used in [16, §4]:

$$B_n(\beta_j) = \frac{nh_j}{q_n(\beta_j)} \int_{-1}^{1} (S(z_0 + h_j u) - S(z_0))\, du + \frac{nh_j}{q_n(\beta_j)} R_n(\beta_j),$$

where

$$|R_n(\beta_j)| \leq \frac{6\delta^{-1}}{n}.$$

Using the weak Hölder condition we obtain $|B_n(\beta_j)| \leq \delta \dfrac{nh_j^{1+\beta}}{q_n(\beta_j)} + 6\delta^{-1} \dfrac{h_j}{q_n(\beta_j)}$, and then

$$|S^*_{h_j}(z_0) - S(z_0)| \leq \delta \frac{nh_j^{1+\beta}}{q_n(\beta_j)} + 6\delta^{-1} \frac{h_j}{q_n(\beta_j)} + |Z_n(\beta_j)| \frac{\sigma_n(\beta_j, S)}{\sqrt{q_n(\beta_j)}}. \tag{4}$$

Hence

$$\begin{aligned} I_1 &\leq \sum_{j=l(\beta)}^{[\ln n]} \left(|S^*_{h_j}(z_0) - S^*_{h_{l(\beta)}}(z_0)| \mathbb{I}_{\{\hat{l}=j\}}\right) + |S^*_{h_{l(\beta)}}(z_0) - S(z_0)| \mathbb{I}_{\{\hat{l} \geq l(\beta)\}} \\ &\leq \sum_{j=l(\beta)}^{[\ln n]} \frac{\lambda}{N_j} \mathbb{I}_{\{\hat{l}=j\}} + |S^*_{h_{l(\beta)}}(z_0) - S(z_0)| \\ &\leq \frac{\lambda}{N_{l(\beta)}} + |S^*_{h_{l(\beta)}}(z_0) - S(z_0)| \\ &\leq \frac{\lambda}{N(\beta)} e^{\beta^\star - \beta_\star} + \delta \frac{nh(\beta_{l(\beta)})^{1+\beta}}{q_n(\beta_{l(\beta)})} + 6\delta^{-1} \frac{h(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} + |Z_n(\beta_{l(\beta)})| \frac{\sigma_n(\beta_{l(\beta)}, S)}{\sqrt{q_n(\beta_{l(\beta)})}}. \end{aligned}$$

Let us show that

$$\limsup_{\delta \to 0} \limsup_{n \to \infty} \sup_{\beta \in [\beta_\star; \beta^\star]} \sup_{S \in \mathcal{U}_{z_0, \delta}} \frac{N(\beta)}{g(z_0, S)} \mathbb{E}_S I_1 \leq \lambda \frac{e^{\beta^\star - \beta_\star}}{g_\star}. \tag{5}$$

As $N(\beta)h(\beta_{l(\beta)})^{\beta} \leq 1$ and $q_n(\beta_{l(\beta)}) \sim 2nh(\beta_{l(\beta)})$, we have

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{\beta\in[\beta_\star;\beta^\star]} \sup_{S\in\mathcal{U}_{z_0,\delta}} \delta \frac{N(\beta)}{g(z_0,S)} \frac{nh(\beta_{l(\beta)})^{1+\beta}}{q_n(\beta_{l(\beta)})} = 0. \tag{6}$$

Moreover it is easy to see that

$$\limsup_{n\to\infty} \sup_{\beta\in[\beta_\star;\beta^\star]} \sup_{S\in\mathcal{U}_{z_0,\delta}} \frac{N(\beta)}{g(z_0,S)} 6\delta^{-1} \frac{h(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} = 0. \tag{7}$$

Since $\sigma_n(\beta_{l(\beta)}, S)$ is bounded by $g^\star$ and

$$\frac{N(\beta)^2}{q_n(\beta_{l(\beta)})} = \frac{nh(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} \Big(\frac{n}{\ln n}\Big)^{2(\beta-\beta_{l(\beta)})/(2\beta+1)(2\beta_{l(\beta)}+1)} \frac{1}{\ln n} \leq \frac{nh(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} e^{2(\beta^\star-\beta_\star)} \frac{1}{\ln n},$$

one has

$$\limsup_{n\to\infty} \sup_{\beta\in[\beta_\star;\beta^\star]} \sup_{S\in\mathcal{U}_{z_0,\delta}} \frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S|Z_n(\beta_{l(\beta)})| \frac{\sigma_n(\beta_{l(\beta)},S)}{\sqrt{q_n(\beta_{l(\beta)})}} = 0. \tag{8}$$

Finally from (6) – (8), we get (5).

Now recall that $I_2 = |\hat{S}_n - S(z_0)|\mathbb{I}_{\{\hat{l}<l(\beta)\}}$ and let us prove that

$$\lim_{n\to\infty} \sup_{\beta\in[\beta_\star;\beta^\star]} \sup_{S\in\mathcal{U}_{z_0,\delta}} \frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S I_2 = 0. \tag{9}$$

We have $\{\hat{l} < l(\beta)\} = \bigcup_{j=0}^{l(\beta)-1} \{\hat{l} = j\}$ and by definition of $\hat{l}$,

$$\{\hat{l} = j\} \subset \bigcup_{i=0}^{j+1} \left(\left| S^*_{h_i}(z_0) - S^*_{h_{j+1}}(z_0) \right| > \frac{\lambda}{N_i} \right).$$

Remark that for $0 \leq i \leq l \leq l(\beta)$, one has $h_i \leq h_l \leq h(\beta)$ and $q_n(\beta_i) \leq q_n(\beta_l)$. Then denoting $Z_n^* = \max_l |Z_n(\beta_l)|$ and using (4) and lemma (A.1), we obtain for $0 \leq i \leq l \leq l(\beta)$:

$$\begin{aligned}
&\left\{ \left| S^*_{h_l}(z_0) - S^*_{h_i}(z_0) \right| > \frac{\lambda}{N_i} \right\} \subset \left\{ \left| S^*_{h_l}(z_0) - S(z_0) \right| + \left| S(z_0) - S^*_{h_i}(z_0) \right| > \frac{\lambda}{N_i} \right\} \\
&\subset \left\{ 2\delta h(\beta)^\beta + \frac{12\delta^{-1}}{nh_i} h(\beta) + \frac{2g^\star}{\sqrt{q_n(\beta_i)}} Z_n^* > \frac{\lambda}{N_i} \right\} \\
&\subset \left\{ \frac{2g^\star}{\sqrt{q_n(\beta_i)}} Z_n^* > \frac{\lambda - 2\delta - 12\delta^{-1}h(\beta^\star)}{N_i} \right\},
\end{aligned}$$

because $nh_i > N_i$ and $N_i < h(\beta)^{-\beta}$.

The precedent inclusions are true for a sufficiently large n and we will consider this case from now to the end of the proof.

As $\dfrac{q_n(\beta_i)}{N_i^2} = \dfrac{q_n(\beta_i)}{nh_i}\ln n$, lemma A.1 brings us $\sqrt{\ln n} \le \dfrac{\sqrt{q_n(\beta_i)}}{N_i} \le \sqrt{3\ln n}$.
Setting $\Lambda_n^\star := \lambda - 2\delta - 12\delta^{-1}h(\beta^\star)$, we can write

$$(\hat{l} = j) \subset \bigcup_{i=0}^{j+1}\left(\left|S^*_{h_i}(z_0) - S^*_{h_{j+1}}(z_0)\right| > \frac{\lambda}{N_i}\right) \subset \left\{Z_n^* > \frac{\Lambda_n^\star}{2g^\star}\sqrt{\ln n}\right\} =: A_n^\star$$

and

$$I_2 \le \sum_{j=0}^{l(\beta)-1} |S^*_{h_j}(z_0) - S(z_0)|\mathbb{I}_{\{A_n^\star\}}.$$

To get (9) we write

$$\frac{N(\beta)}{g(z_0,S)}\mathbb{E}_S I_2 \le \frac{N(\beta)}{g(z_0,S)}\sum_{j=0}^{l(\beta)-1}\mathbb{E}_S\left(\frac{\delta n h_j^{1+\beta}}{q_n(\beta_j)} + \frac{6\delta^{-1}h_j}{q_n(\beta_j)} + |Z_n(\beta_j)|\frac{\sigma(\beta_j,S)}{\sqrt{q_n(\beta_j)}}\right)\mathbb{I}_{\{A_n^\star\}}$$
$$\le \sum_{j=0}^{l(\beta)-1}\frac{N(\beta)}{g(z_0,S)}\frac{\delta n h_j^{1+\beta}}{q_n(\beta_j)}\mathbb{P}_S\left(A_n^\star\right) + \sum_{j=0}^{l(\beta)-1}\frac{N(\beta)}{g(z_0,S)}\frac{6\delta^{-1}h_j}{q_n(\beta_j)}\mathbb{P}_S(A_n^\star)$$
$$+ \sum_{j=0}^{l(\beta)-1}\frac{N(\beta)}{g(z_0,S)}\frac{\sigma_n(\beta_j,S)}{\sqrt{q_n(\beta_j)}}\mathbb{E}_S\left(|Z_n(\beta_j)|\mathbb{I}_{(A_n^\star)}\right)$$

and study the asymptotic behavior of each term. Let $Z \sim \mathcal{N}(0,1)$ and $A_n := \left\{|Z| > \frac{\Lambda_n^\star}{2g^\star}\sqrt{\ln n}\right\}$. For the first term one has

$$\sum_{j=0}^{l(\beta)-1}\frac{N(\beta)}{g(z_0,S)}\frac{\delta n h_j^{1+\beta}}{q_n(\beta_j)}\mathbb{P}_S\left(A_n^\star\right) \le \frac{\delta}{g_\star}\sum_{j=0}^{l(\beta)-1}N(\beta)h_j^\beta\mathbb{P}_S\left(A_n^\star\right)$$
$$\le \frac{\delta}{g_\star}\sum_{j=0}^{l(\beta)-1}\left(\frac{n}{\ln n}\right)^{2\beta(\beta_j-\beta)/(2\beta_j+1)(2\beta+1)}\mathbb{P}_S\left(A_n^\star\right)$$
$$\le [\ln n]^2\frac{\delta}{g_\star}\left(\frac{n}{\ln n}\right)^{2\beta(\beta_{l(\beta)-1}-\beta)/(2\beta_{l(\beta)-1}+1)(2\beta+1)}\mathbb{P}_S\left(A_n\right)$$
$$\le [\ln n]^2\frac{\delta}{g_\star}\left(\frac{n}{\ln n}\right)^{2\beta(\beta_{l(\beta)-1}-\beta)/(2\beta+1)^2}2\frac{2g^\star}{\Lambda_n^\star\sqrt{\ln n}}\frac{1}{\sqrt{2\pi}}e^{-\frac{(\Lambda_n^\star)^2}{4g^{\star 2}}\ln n}$$
$$\le \frac{4\delta g^\star}{\sqrt{2\pi}g_\star\Lambda_n^\star}e^{-2\beta_\star/(2\beta^\star+1)^2}e^{2\beta^\star/(2\beta_\star+1)^2}\frac{[\ln n]^2}{\sqrt{\ln n}}n^{-\frac{(\Lambda_n^\star)^2}{4g^{\star 2}}}.$$

Since $\lambda > 2$ the last term tends to zero as n goes to infinity. We handle the second term as well.

Using the Cauchy-Schwarz inequality for the third term, we obtain

$$\frac{N(\beta)}{g(z_0,S)}\sum_{j=0}^{l(\beta)-1}\mathbb{E}_S\left(|Z_n(\beta_j)|\frac{\sigma_n(\beta_j,S)}{\sqrt{q_n(\beta_j)}}\mathbb{I}_{\{A_n^\star\}}\right)\le\frac{g^\star}{g_\star}\sum_{j=0}^{l(\beta)-1}\frac{N(\beta)}{\sqrt{q_n(\beta_j)}}\left(\mathbb{P}_S\left(A_n^\star\right)\right)^{1/2}$$

$$\le\frac{g^\star}{g_\star}\sum_{j=0}^{l(\beta)-1}\frac{N(\beta)}{\sqrt{q_n(\beta_j)}}[\ln n]^{1/2}\left(\mathbb{P}_S\left(A_n\right)\right)^{1/2}$$

$$\le\frac{2(g^\star)^{3/2}}{g_\star\sqrt{\Lambda_n^\star\sqrt{2\pi}}}\frac{[\ln n]^{3/2}}{\sqrt{\sqrt{\ln n}}\sqrt{q_n(\beta_\star)}}N(\beta^\star)n^{-(\Lambda_n^\star)^2/8g^{\star 2}}.$$

By definition of λ this term tends to zero as n goes to infinity. Eventually we have proved (9) which, connected to (5), completes the proof. □

4. Lower bound

In this section we give the lower bound for the minimax risk. We will consider a family of functions in $\mathcal{U}_{z_0,\delta}$ defined with another bandwidth $\tilde{h}(\beta) = n^{-1/2\beta+1}(\ln n)^{-2\beta/2\beta+1}$ such that $n\tilde{h}(\beta) = N^2(\beta)$ and $\tilde{h}$ is a increasing function of β.

Theorem 4.1 *For all $\delta\in]0;1[$ the following inequality holds*

$$\liminf_{n\to\infty}\inf_{\tilde{S}}\mathcal{R}_{z_0,\delta,n}(\tilde{S})\ge\frac{\mathbb{E}|\xi|}{\sqrt{2}},\qquad\xi\sim\mathcal{N}(0,1),$$

where the infimum is taken over all estimators $\tilde{S}$ of $S(z_0)$.

Proof: For $\nu\in\left]0;\frac{1}{4}\right[$, denote $S_\nu(x)=\frac{1}{N(\beta)}V_\nu\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)$, where the function V_ν is defined by

$$V_\nu(x)=\frac{1}{\nu}\int_{-\infty}^{+\infty}\tilde{Q}_\nu(u)m\left(\frac{u-x}{\nu}\right)du,$$

$$\tilde{Q}_\nu(u)=\mathbb{I}_{\{|u|\le 1-2\nu\}}+2\mathbb{I}_{\{1-2\nu\le|u|\le 1-\nu\}}$$

and m is a nonnegative function, infinitely differentiable on $\mathbb{R}$, such that $m(z)=0$ for any $|z|\ge 1$ and $\int_{-1}^{1}m(z)dz=1$. We can easily show that for all $0<\nu<\frac{1}{4}$, one has $V_\nu(0)=1$ and $\int_{-1}^{1}V_\nu(x)dx=2$.

Now fix $b>0$, $\nu\in]0;1/4[$, $\delta\in]0;1[$ and denote $S_{\nu,u}(x)=\frac{u}{N(\beta)}V_\nu\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)$ for $x,u\in\mathbb{R}$. Thanks to lemma A.2 we can say that if $|u|\le b$ then there exists an

integer $n_{b,\delta} > 0$ such that $S_{\nu,u} \in \mathcal{U}_{z_0,\delta}$ for all $n \geq n_{b,\delta}$. Hence for $n \geq n_{b,\delta}$, we have

$$\mathcal{R}_{z_0,\delta,n}(\tilde{S}) \geq \sup_{\beta\in[\beta_\star;\beta^\star]} \sup_{|u|\leq b} \frac{N(\beta)}{g(z_0, S_{\nu,u})} \mathbb{E}_{S_{\nu,u}} |\tilde{S}(z_0) - S_{\nu,u}(z_0)|$$

$$\geq \sup_{\beta\in[\beta_\star;\beta^\star]} \frac{1}{2b} \int_{-b}^{b} \frac{N(\beta)}{g(z_0, S_{\nu,u})} \mathbb{E}_{S_{\nu,u}} v_a \left(\tilde{S}(z_0) - S_{\nu,u}(z_0)\right) du =: \sup_{\beta\in[\beta_\star;\beta^\star]} I_n(a, b, \beta)$$

where $v_a(x) = |x| \wedge a, a > 0$.
Let $\mathbb{P}_{S_{\nu,u}}$ be the law of $(y_k^{(1)})_{k=1,\ldots,n}$, where $y_k^{(1)} = S_{\nu,u}(x_k) + g(x_k, S_{\nu,u})\xi_k$, and $\mathbb{P}$ the law of $(y_k^{(0)})_{k=1,\ldots,n}$, where $y_k^{(0)} = g(x_k, S_{\nu,u})\xi_k$. These two measures are equivalent and the corresponding Radon-Nikodym derivative at the point $(y_1, \ldots, y_n)$ is

$$\rho_n(u, \beta) = \frac{d\mathbb{P}_{S_{\nu,u}}}{d\mathbb{P}}(y_1, \ldots, y_n) = \exp\left(u\varsigma_n(\beta)\eta_n(\beta) - \frac{u^2}{2}\varsigma_n^2(\beta)\right)$$

where $\varsigma_n^2(\beta) = \dfrac{1}{N^2(\beta)} \displaystyle\sum_{k=1}^{n} \frac{V_\nu^2\left(\frac{x_k - z_0}{\tilde{h}(\beta)}\right)}{g^2(x_k, S_{\nu,u})}$ and $\eta_n(\beta) = \dfrac{1}{N(\beta)\varsigma_n(\beta)} \displaystyle\sum_{k=1}^{n} \frac{V_\nu\left(\frac{x_k - z_0}{\tilde{h}(\beta)}\right)}{g^2(x_k, S_{\nu,u})} y_k$.
Under the law $\mathbb{P}$, $\eta_n(\beta)$ is a standard Gaussian random variable.

Setting $\sigma_\nu^2 = \frac{1}{g^2(z_0,0)} \displaystyle\int_{-1}^{1} V_\nu^2(z)dz$, lemma A.3 implies

$$\sup_{\beta\in[\beta_\star;\beta^\star]} |\varsigma_n(\beta) - \sigma_\nu| \xrightarrow[n\to\infty]{} 0. \tag{10}$$

Rewrite $\rho_n(u, \beta) = \exp\left\{u\sigma_\nu\eta_n(\beta) - \frac{u^2\sigma_\nu^2}{2} + r_n(\beta)\right\}$, where $r_n(\beta)$ is a Gaussian random variable with expectation $\dfrac{u^2}{2}(\sigma_\nu^2 - \varsigma_n^2(\beta))$ and variance $u^2(\varsigma_n(\beta) - \sigma_\nu)^2$. Then using (10) and Tchebychev inequality we get for $\varepsilon > 0$

$$\lim_{n\to\infty} \sup_{\beta\in[\beta_\star;\beta^\star]} \mathbb{P}\left(|r_n(\beta)| > \varepsilon\right) = 0. \tag{11}$$

As a consequence we can show that $\rho_n(u, \beta) \xrightarrow[n\to\infty]{\mathcal{L}} \rho_\infty(u) := \exp\left(u\sigma_\nu\eta - \frac{u^2\sigma_\nu^2}{2}\right)$ uniformly in β and

$$\lim_{n\to\infty} \sup_{\beta\in[\beta_\star,\beta^\star]} \mathbb{P}\left(|e^{r_n(\beta)} - 1| > \varepsilon\right) = 0.$$

Denoting $\psi_{a,n}(\tilde{S}, S_{\nu,u}) = v_a(N(\beta)(\tilde{S}(z_0) - S_{\nu,u}(z_0)))$ and $\mathbb{E}$ the expectation for the probability measure $\mathbb{P}$, one has

$$I_n(a, b, \beta) \geq \frac{1}{2b} \int_{-b}^{b} \mathbb{E}\mathbb{I}_{B_d(\beta)} \frac{\psi_{a,n}(\tilde{S}, S_{\nu,u})}{g(z_0, S_{\nu,u})} \varrho_n(u, \beta)du + \delta_n(a, b, \beta) \tag{12}$$

$$=: J_n(a, b, \beta) + \delta_n(a, b, \beta), \tag{13}$$

where

$$B_d(\beta) = \{|\eta_n(\beta)| \leq d\} \quad \text{and} \quad d = \sigma_\nu(b - \sqrt{b}), b > 1,$$
$$\varrho_n(u,\beta) = \exp\left(u\sigma_\nu\eta_n(\beta) - \frac{u^2\sigma_\nu^2}{2}\right),$$
$$\delta_n(a,b,\beta) = \frac{1}{2b}\int_{-b}^{b} \mathbb{E}\mathbb{I}_{B_d(\beta)} \frac{\psi_{a,n}(\tilde{S}, S_{\nu,u})}{g(z_0, S_{\nu,u})}\theta_n(u,\beta)du,$$
$$\theta_n(u,\beta) = \rho_n(u,\beta) - \varrho_n(u,\beta) = \exp\left(u\sigma_\nu\eta_n(\beta) - \frac{u^2\sigma_\nu^2}{2}\right)(e^{r_n(\beta)} - 1).$$

Let us show that

$$\inf_{\tilde{S}} \sup_{\beta\in[\beta_\star;\beta^\star]} \delta_n(a,b,\beta) \xrightarrow[n\to\infty]{} 0. \tag{14}$$

We can easily prove that $\mathbb{E}\rho_\infty(u) = 1$ and $\mathbb{E}\rho_n(u,\beta) = 1$. Then the sequence $\{\rho_n(u,\beta), n \geq 1\}$ is uniformly integrable (see [18, p.32]). As $\varrho_n(u,\beta)$ is bounded on $B_d(\beta)$, we obtain the uniform integrability of $\left\{\mathbb{I}_{B_d(\beta)}\frac{\psi_{a,n}(\tilde{S},S_{\nu,u})}{g(z_0,S_{\nu,u})}\theta_n(u,\beta), n \geq 1\right\}$. We remark that $\exp\left(u\sigma_\nu\eta_n(\beta) - \frac{u^2\sigma_\nu^2}{2}\right)$ is bounded by a constant independent of β on $B_d(\beta)$ and that $\psi_{a,n}(\tilde{S}, S_{\nu,u})$ is bounded by a. Then

$$\lim_{n\to\infty} \sup_{\beta\in[\beta_\star;\beta^\star]} \mathbb{P}\left(\mathbb{I}_{B_d(\beta)}\frac{\psi_{a,n}(\tilde{S},S_{\nu,u})}{g(z_0,S_{\nu,u})}|\theta_n(u,\beta)| > \varepsilon\right) = 0.$$

It follows from this and from the uniform integrability of $\left\{\mathbb{I}_{B_d(\beta)}\frac{\psi_{a,n}(\tilde{S},S_{\nu,u})}{g(z_0,S_{\nu,u})}\theta_n(u,\beta), n \geq 1\right\}$ that

$$\lim_{n\to\infty} \sup_{\beta\in[\beta_\star;\beta^\star]} \mathbb{E}\left|\mathbb{I}_{B_d(\beta)}\frac{\psi_{a,n}(\tilde{S},S_{\nu,u})}{g(z_0,S_{\nu,u})}\theta_n(u,\beta)\right| = 0.$$

Furthermore there exists a constant $K > 0$ such that

$$\sup_{\beta\in[\beta_\star;\beta^\star]}\left|\mathbb{E}\frac{\mathbb{I}_{B_d}}{g(z_0,S_{\nu,u})}\psi_{a,n}(\tilde{S},S_{\nu,u})\theta_n(u,\beta)\right| \leq K \sup_{\beta\in[\beta_\star;\beta^\star]}\mathbb{E}|e^{r_n(\beta)} - 1| \leq Ke^{u^2\sigma_n\nu^2}.$$

By bounded convergence we obtain finally

$$\sup_{\beta\in[\beta_\star;\beta^\star]}\delta_n(a,b,\beta) \leq \frac{1}{2b}\int_{-b}^{b}\sup_{\beta\in[\beta_\star;\beta^\star]}\mathbb{E}\left|\frac{\mathbb{I}_{B_d}}{g(z_0,S_{\nu,u})}\psi_{a,n}(\tilde{S},S_{\nu,u})\theta_n(u)\right|du \xrightarrow[n\to\infty]{} 0,$$

and then (14).

Now we are interested in the term $J_n(a,b,\beta)$ in (13).
Firstly rewrite $\varrho_n(u,\beta) = \zeta_n(\beta)e^{-\sigma_\nu^2(u-\tilde{\eta}_n(\beta))^2/2}$ where $\zeta_n(\beta) = e^{\eta_n^2(\beta)/2}$ and $\tilde{\eta}_n(\beta) = \frac{\eta_n(\beta)}{\sigma_\nu}$, then

$$J_n(a,b,\beta) = \frac{1}{2b}\int_{-b}^{b}\mathbb{E}\mathbb{I}_{B_d(\beta)}\zeta_n(\beta)\frac{v_a(u-(t_n(\beta))}{g(z_0,S_{\nu,u})}\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\eta}_n(\beta))^2\right)du,$$

with $t_n(\beta) = N(\beta)\tilde{S}(z_0)$.
Further if $\xi \sim \mathcal{N}(0,1)$ and if we set $\tilde{\xi} = \dfrac{\xi}{\sigma_\nu}$, $\zeta = e^{\xi^2/2}$, $\tilde{B}_d = \{|\xi| \leq d\}$ and $\tilde{\mathbb{E}}$ the expectation for the probability measure of ξ, we have

$$J_n(a,b,\beta) = \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b}\frac{v_a(u-t_n(\beta))}{g(z_0,S_{\nu,u})}\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)du.$$

Since g is uniformly continuous and bounded away from zero, the following term

$$\sup_{\beta\in[\beta_\star;\beta^\star]}\tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\frac{\zeta}{2b}\int_{-b}^{b}v_a(u-t_n(\beta))\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)\left(\frac{1}{g(z_0,S_{\nu,u})}-\frac{1}{g(z_0,0)}\right)du$$

tends toward zero as n goes to infinity.
It follows that

$$\liminf_{n\to\infty}\inf_{\tilde{S}}\sup_{\beta\in[\beta_\star;\beta^\star]}J_n(a,b,\beta) \geq$$

$$\liminf_{n\to\infty}\inf_{\tilde{S}}\sup_{\beta\in[\beta_\star;\beta^\star]}\tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b}\frac{v_a(u-t_n(\beta))}{g(z_0,0)}\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)du.$$

Moreover we can write

$$\begin{aligned}\tilde{\mathbb{E}}\frac{\mathbb{I}_{\tilde{B}_d}\zeta}{2b}\int_{-b}^{b}\frac{v_a(u-t_n(\beta))}{g(z_0,0)}e^{-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2}du &\geq \tilde{\mathbb{E}}\frac{\mathbb{I}_{\tilde{B}_d}\zeta}{2b}\int_{-\sqrt{b}}^{\sqrt{b}}\frac{v_a(t-t_n(\beta)+\tilde{\xi})}{g(z_0,0)}e^{-\frac{\sigma_\nu^2}{2}t^2}dt\\ &\geq \tilde{\mathbb{E}}\frac{\mathbb{I}_{\tilde{B}_d}\zeta}{2b}\int_{-\sqrt{b}}^{\sqrt{b}}\frac{v_a(t)}{g(z_0,0)}\exp\left(-\frac{\sigma_\nu^2}{2}t^2\right)dt,\end{aligned}$$

using Anderson's lemma for the last inequality (see [19, Chapter II, Lemma 10.1 and Corollary 10.2]).
Noticing that $\tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta = \dfrac{2\sigma_\nu(b-\sqrt{b})}{\sqrt{2\pi}}$ and limiting successively $a\to\infty$ and $b\to\infty$, we get

$$\liminf_{b\to\infty}\liminf_{a\to\infty}\liminf_{n\to\infty}\inf_{\tilde{S}}\sup_{\beta\in[\beta_\star;\beta^\star]}J_n(a,b,\beta) \geq \frac{\sigma_\nu}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{|t|}{g(z_0,0)}\exp\left(-\frac{\sigma_\nu^2}{2}t^2\right)dt.$$

Eventually using the fact that $\sigma_\nu^2 \xrightarrow[\nu\to 0]{} \dfrac{2}{g^2(z_0,0)}$ and combining the result with (14) finish the proof. □

Appendix A.

Lemma A.1 *There exists an integer $N_\star$ such that if $n \geq N_\star$, then*

$$1 \leq \frac{q_n(\beta_i)}{nh_i} \leq 3,$$

for all $i = 0,\ldots,[\ln n]$.

Proof: Writing

$$\frac{q_n(\beta_i)}{nh_i} = \frac{1}{nh_i}\sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_i}\right) = \frac{[n(z_0+h_i)] - [n(z_0-h_i)]}{nh_i},$$

the following inequalities

$$\begin{aligned}[n(z_0+h_i)] \leq\ & n(z_0+h_i) \ \leq [n(z_0+h_i)] + 1,\\ -1 - [n(z_0-h_i)] \leq\ & -n(z_0-h_i) \leq -[n(z_0-h_i)],\end{aligned}$$

show that $2 - \dfrac{1}{nh_i} \leq \dfrac{q_n(\beta_i)}{nh_i} \leq 2 + \dfrac{1}{nh_i}$.
But there exists an integer $N_\star$ such that if $n \geq N_\star$, then $nh(\beta_\star) \geq 1$. As h is an increasing function of β, we get for $n \geq N^\star$:

$$2 - \frac{1}{nh(\beta_\star)} \leq \frac{q_n(\beta_i)}{nh_i} \leq 2 + \frac{1}{nh(\beta_\star)},$$

and then the desired result. □

Lemma A.2 *For any $\nu \in]0;\frac{1}{4}[$ and $\delta \in]0;1[$ there exists an integer $n_{\nu,\delta} > 0$ such that $S_\nu \in \mathcal{U}_{z_0,\delta}$ for all $n \geq n_{\nu,\delta}$.*

Proof: Firstly for all $h \geq 0$ one has $\int_{-1}^{1} (S_\nu(z_0+uh) - S_\nu(z_0))\, du = 0$. Moreover we can write

$$|S'_\nu(x)| \leq \frac{2}{\nu N(\beta)\tilde{h}(\beta)} \int_{-1}^{1} |m'(u)|du \leq \frac{2}{\nu} n^{(1-\beta_\star)/(2\beta_\star+1)} (\ln n)^{3\beta^\star/2\beta^\star+1} \int_{-1}^{1} |m'(u)|du,$$

since

$$\left|V'_\nu\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)\right| = \left|-\nu^{-2}\int_{-\infty}^{+\infty} \tilde{Q}_\nu(u) m'\left(\frac{u - \frac{x-z_0}{\tilde{h}(\beta)}}{\nu}\right) du\right| \leq \frac{2}{\nu}\int_{-1}^{1} |m'(u)|du$$

Hence for any fixed $\delta \in]0;1[$, there exists $n_{\nu,\delta} \in \mathbb{N}^*$ such that if $n \geq n_{\nu,\delta}$, then $|S'_\nu(x)| \leq \delta^{-1}$ and $S_\nu \in \mathcal{U}_{z_0,\delta}$. □

Lemma A.3 *The following limit holds*

$$\sup_{\beta\in[\beta_\star;\beta^\star]} \left|\varsigma_n^2(\beta) - \frac{1}{g^2(z_0,0)}\int_{-1}^{1} V_\nu^2(z)dz\right| \xrightarrow[n\to\infty]{} 0. \tag{A1}$$

Proof: There exists an integer n_0 such that for all $n \geq n_0$ and all $\beta \in [\beta_\star;\beta^\star]$, we have $[z_0 - \tilde{h}(\beta); z_0 + \tilde{h}(\beta)] \subset [0;1]$. As a consequence for $n \geq n_0$ we get

$$\varsigma_n^2(\beta) = \frac{1}{\tilde{h}(\beta)}\int_{z_0-\tilde{h}(\beta)}^{z_0+\tilde{h}(\beta)} \frac{V_\nu^2\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)}{g^2(x,S_{\nu,u})}\mu_n(dx) = \int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)}{g^2(x,S_{\nu,u})}\nu_n(dx)$$

with $\mu_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{k/n}$ and $\nu_n = \frac{\mathbb{I}_{[z_0-\tilde{h}(\beta);z_0+\tilde{h}(\beta)]}}{\tilde{h}(\beta)}\mu_n$.
As g is uniformly continuous, we have:

$$\left|\int_0^1 \left(\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(x,0)}\right)\nu_n(dx)\right| \xrightarrow[n\to\infty]{} 0. \tag{A2}$$

We easily prove that for a real function f defined on $\mathbb{R}$ and continuous at the point z_0, one has

$$\lim_{n\to\infty}\sup_{\beta\in[\beta_\star;\beta^\star]}\left|\int_0^1 f(x)\nu_n(dx) - 2f(z_0)\right| = 0. \tag{A3}$$

Writing

$$\begin{aligned}&\int_0^1 \left(\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(z_0,0)}\right)\nu_n(dx)\\ &= \int_0^1 \left(\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(x,0)}\right)\nu_n(dx) + \int_0^1 \left(\frac{1}{g^2(x,0)} - \frac{1}{g^2(z_0,0)}\right)\nu_n(dx),\end{aligned}$$

and applying (A3) to the function $x \mapsto \dfrac{1}{g^2(x,0)}$ and using (A2), we obtain

$$\lim_{n\to\infty}\sup_{\beta\in[\beta_\star;\beta^\star]}\left|\int_0^1 \left(\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(z_0,0)}\right)\nu_n(dx)\right| = 0. \tag{A4}$$

Now we want to show that

$$\lim_{n\to\infty}\sup_{\beta\in[\beta_\star;\beta^\star]}\left|\int_0^1 V_\nu^2\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)\nu_n(dx) - \int_{-1}^1 V_\nu^2(y)dy\right| = 0. \tag{A5}$$

Let F be a function such that $F' = V_\nu^2$ and denote $u_k = \dfrac{x_k - z_0}{\tilde{h}(\beta)}$ again. For all $n \geq n_0$, one has

$$\int_{-1}^1 V_\nu^2(y)dy = F(1) - F(-1) = \sum_{k=1}^{n}(F(u_k) - F(u_{k-1})),$$

then

$$\begin{aligned}&\left|\int_0^1 V_\nu^2\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)\nu_n(dx) - \int_{-1}^1 V_\nu^2(y)dy\right|\\ &= \left|\sum_{k=1}^{n}\left(F(u_k) - F(u_{k-1}) - \frac{1}{n\tilde{h}(\beta)}F'(u_k)\right)\right| = \left|\sum_{k=1}^{n}\frac{-1}{2(n\tilde{h}(\beta))^2}F''(v_k)\right|\\ &\leq \frac{1}{2n\tilde{h}^2(\beta_\star)}(V_\nu^{2\prime})_{max},\end{aligned}$$

with $v_k \in]u_{k-1};u_k[$ and $(V_\nu^{2\prime})_{max}$ the maximum of the function $V_\nu^{2\prime}$. Limiting $n \to \infty$ yields (A5).

Finally, as we have

$$\left|\varsigma_n^2(\beta) - \int_{-1}^{1} \frac{V_\nu^2(z)}{g^2(z_0,0)}dz\right| = \left|\int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)}{g^2(x,S_{\nu,u})}\nu_n(dx) - \int_{-1}^{1} \frac{V_\nu^2(z)}{g^2(z_0,0)}dz\right|$$

$$\leq V_\nu^{max} \int_0^1 \left|\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(z_0,0)}\right| \nu_n(dx)$$

$$+ \left|\int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{\tilde{h}(\beta)}\right)}{g^2(z_0,0)}\nu_n(dx) - \int_{-1}^{1} \frac{V_\nu^2(z)}{g^2(z_0,0)}dz\right|,$$

(A4) and (A5) implies lemma A.3. □

References

[1] J. Sacks and D. Ylvisaker, *Asymptotically optimum kernels for density estimation at a point*, Ann. Statist. 9(2) (1981), pp. 334–346.
[2] D. Donoho and R. Liu, *Geometrizing rates of convergence. III*, Ann. Statist. 19 (1991), pp. 668–701.
[3] D. Donoho, *Asymptotic minimax risk for sup-norm loss: solution via optimal recovery*, Probab. Theory Related Fields 99(2) (1994), pp. 145–170.
[4] A. Korostelev, *Exact asymptotically minimax estimator for nonparametric regression in uniform norm*, Theory Probab. Appl. 38 (1993), pp. 737–743.
[5] S. Efromovich *Nonparametric Curve Estimation. Methods, Theory and Applications*, Springer, Berlin, New York, 1999.
[6] G.K. Golubev, *Asymptotically minimax estimation of a regression function in an additive model*, Problems Inform. Trans. 28 (1992), pp. 3–15.
[7] D. Donoho et al., *Wavelet shrinkage: asymptopia?*, J. Roy. Statist. Soc. Ser. B 57 (1995), pp. 301–369.
[8] S. Goldfeld and R. Quandt *Nonlinear Methods in Econometrics*, North-Holland, Amsterdam, London, 1972.
[9] T. Cai and L. Wang, *Adaptive variance function estimation in heteroscedastic nonparametric regression*, Ann. Statist., to appear (2007).
[10] A. Barron, L. Birgé, and P. Massart, *Risk bounds for model selection via penalization*, Probab. Theory Related Fields 113 (1999), pp. 301–413.
[11] L. Galtchouk and S. Pergamenshchikov, *Efficient adaptive nonparametric estimation in heteroscedastic regression models*, Preprint of the Strasbourg Louis Pasteur University, IRMA (2005), pp. available online at http://hal.archives–ouvertes.fr/hal–00129707/fr/.
[12] O.V. Lepskiĭ, *A problem of adaptive estimation in Gaussian white noise*, Teor. Veroyatnost. i Primenen. 35(3) (1990), pp. 459–470.
[13] O.V. Lepskiĭ, *Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates*, Teor. Veroyatnost. i Primenen. 36(4) (1991), pp. 645–659.
[14] O.V. Lepskiĭ, *Asymptotically minimax adaptive estimation. II. Schemes without optimal adaptation. Adaptive estimates*, Teor. Veroyatnost. i Primenen. 37(3) (1992), pp. 468–481.
[15] L. Galtchouk and S. Pergamenshchikov, *Sequential nonparametric adaptive estimation of the drift coefficient in diffusion processes*, Math. Methods Statist. 10(3) (2001), pp. 316–330 Meeting on Mathematical Statistics (Marseille, 2000).
[16] L. Galtchouk and S. Pergamenshchikov, *Asymptotically efficient estimates for nonparametric regression models*, Statist. Probab. Lett. 76 (2006), pp. 852–860.
[17] J.Y. Brua, *Asymptotically efficient estimators for nonparametric heteroscedastic regression models*, Statistical Methodology, to appear (2008).
[18] P. Billingsley *Convergence of probability measures*, Second , Wiley Series in Probability and Statistics: Probability and Statistics John Wiley & Sons Inc., New York, 1999 A Wiley-Interscience Publication.
[19] I. Ibragimov and R. Hasminskii *Statistical Estimation: Asymptotic Theory*, Springer, Berlin, New York, 1981.