Bernoulli **10**(5), 2004, 873–888

Asymptotically exact minimax estimation in sup-norm for anisotropic Hölder classes

KARINE BERTIN

Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 6, 4 place Jussieu, case 188, 75252 Paris Cedex 05, France. E-mail: kbertin@ccr.jussieu.fr

We consider the Gaussian white noise model and study the estimation of a function f in the uniform norm assuming that f belongs to a Hölder anisotropic class. We give the minimax rate of convergence over this class and determine the minimax exact constant and an asymptotically exact estimator.

Keywords: anisotropic Hölder class; minimax exact constant; uniform norm; white noise model

1. Introduction

Let $\{Y_t,\ t \in [0, 1]^d\}$, be a random process defined by the stochastic differential equation

$$\mathrm{d}Y_t = f(t)\mathrm{d}t + \frac{\sigma}{\sqrt{n}}\,\mathrm{d}W_t, \qquad t \in [0, 1]^d, \tag{1}$$

where f is an unknown function, $n > 1$, $\sigma > 0$ is known and W is a standard Brownian sheet in $[0, 1]^d$. We wish to estimate the function f given a realization $y = \{Y_t,\ t \in [0, 1]^d\}$. This is known as the Gaussian white noise problem and has been studied in several papers, starting with Ibragimov and Has'minskii (1981). We suppose that f belongs to a d-dimensional anisotropic Hölder class $\Sigma(\tilde{\beta}, L)$ for $\tilde{\beta} = (\beta_1, \ldots, \beta_d) \in (0, 1]^d$ and $L = (L_1, \ldots, L_d)$ such that $0 < L_i < \infty$. This class is defined by

$$\Sigma(\tilde{\beta}, L) = \\ \left\{f : [0, 1]^d \to \mathbb{R} : |f(x) - f(y)| \leqslant L_1|x_1 - y_1|^{\beta_1} + \ldots + L_d|x_d - y_d|^{\beta_d},\ x, y \in [0, 1]^d\right\},$$

where $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$.

In the following $\mathbb{P}_f$ is the distribution of y under model (1) and $\mathbb{E}_f$ is the corresponding expectation. We denote by β the real number such that $1/\beta = \sum_{i=1}^d (1/\beta_i)$. Let $w(u)$, $u \geqslant 0$, be a continuous non-decreasing function which admits a polynomial majorant $w(u) \leqslant W_0(1 + u^\gamma)$ with some finite positive constants W_0, γ and such that $w(0) = 0$.

Let θ_n be an estimator of f, i.e. a random function on $[0, 1]^d$ with values in $\mathbb{R}$ measurable with respect to $\{Y_t,\ t \in [0, 1]^d\}$. The quality of θ_n is characterized by the maximal risk in sup-norm,

$$R_n(\theta_n) = \sup_{f \in \Sigma(\tilde{\beta}, L)} \mathbb{E}_f w\left(\frac{\|\theta_n - f\|_\infty}{\psi_n}\right),$$

where $\psi_n = ((\log n)/n)^{\beta/(2\beta+1)}$ and $\|g\|_\infty = \sup_{t\in[0,1]^d}|g(t)|$. The normalizing factor ψ_n is used here because it is a minimax rate of convergence. For the one-dimensional case, the fact that ψ_n is the minimax rate for the sup-norm has been proved by Ibragimov and Has'minskii (1981). For the multidimensional case, this fact was shown by Stone (1982) and Nussbaum (1986) for the isotropic setting ($\beta_1 = \cdots = \beta_d$), but it has not been shown for the anisotropic setting considered here. Nevertheless there exist results for estimation in $\mathbb{L}_p$ norm with $p < \infty$ on anisotropic Besov classes (Kerkyacharian *et al.*, 2001) suggesting similar rates but without a logarithmic factor. The case $p = 2$ has been treated by several authors (Neumann and von Sachs 1997; Barron *et al.* 1999).

Our result implies in particular that ψ_n is the minimax rate of convergence for estimation in sup-norm. But we prove a stronger assertion: we find an estimator $\hat{f}_n$ and determine the minimax exact constant $C(\beta, L, \sigma^2)$ such that

$$C(\beta, L, \sigma^2) = \lim_{n\to\infty} \inf_{\theta_n} R_n(\theta_n) = \lim_{n\to\infty} R_n(\hat{f}_n), \tag{2}$$

where $\inf_{\theta_n}$ stands for the infimum over all the estimators. Such an estimator $\hat{f}_n$ will be called asymptotically exact.

The problem of asymptotically exact constants under the sup-norm was first studied in the one-dimensional case by Korostelev (1993) for the regression model with fixed equidistant design. Korostelev found the exact constant and an asymptotically exact estimator for this set-up. Donoho (1994) extended Korostelev's result to the Gaussian white noise model and Hölder classes with $\beta > 1$. However, asymptotically exact estimators are not available in explicit form for $\beta > 1$, except for $\beta = 2$. Korostelev and Nussbaum (1999) found the exact constant and asymptotically exact estimator for the density model. Lepskii (1992) studied the exact constant in the case of adaptation for the white noise model. Bertin (2004) found the exact constant and an asymptotically exact estimator for the regression model with random design.

The estimator $\hat{f}_n$ defined in Section 2 and which will be shown to satisfy (2) is a kernel estimator. For $d = 1$, the kernel used in our estimator (and defined in (3)) is the one derived by Korostelev (1993) and can be viewed as a solution of an optimal recovery problem. This is explained in Donoho (1994) and Lepski and Tsybakov (2000). For our set-up, i.e. the Gaussian white noise model and d-dimensional anisotropic Hölder class $\Sigma(\tilde{\beta}, L)$ for $\tilde{\beta} = (\beta_1, \ldots, \beta_d) \in (0, 1]^d$ and $L \in (0, +\infty)^d$, the choice of optimal parameters of the estimator (i.e. kernel, bandwidth) is also related to a solution of optimal recovery problems. In the same way as in Donoho (1994), the kernel defined in (3) can be expressed, up to a renormalization on the support, as

$$K(t) = \frac{g_{\tilde{\beta}}(t)}{\int_{\mathbb{R}^d} g_{\tilde{\beta}}(s)\mathrm{d}s},$$

where $g_{\tilde{\beta}}$ is the solution of the optimization problem

$$\max g_{\tilde{\beta}}(0), \qquad \|g_{\tilde{\beta}}\|_2 \leqslant 1,\ g_\beta \in \Sigma(\tilde{\beta}, \mathbf{1})$$

where $\|f\|_2 = \left(\int_{\mathbb{R}^d} f^2(t)\mathrm{d}t\right)^{1/2}$ and $\mathbf{1}$ is the vector $(1, \ldots, 1)$ in $\mathbb{R}^d$.

The anisotropic class of functions in this paper does not turn into a traditional isotropic Lipschitz class in the case $\beta_1 = \ldots = \beta_d$. For an isotropic class defined as

$$\{f : [0, 1]^d \to \mathbb{R} : |f(x) - f(y)| \leqslant L\|x - y\|^\beta,\ x, y \in [0, 1]^d\},$$

with $\beta \in (0, 1]$, $L > 0$ and $\|\cdot\|$ the Eucleadian norm in $\mathbb{R}^d$, radial symmetric 'cone-type' kernels should be optimal. Such kernels of the form $K(x) = (1 - \|x\|)_+$, for $x \in \mathbb{R}^d$, are studied in Klemelä and Tsybakov (2001). We denote $(t)_+ = \max(0, t)$.

In Section 2, we give an asymptotically exact estimator $\hat{f}_n$ and the exact constant for the Gaussian white noise model. The proofs are given in Sections 3 and 4.

2. The estimator and main result

Consider the kernel K defined for $u = (u_1, \ldots, u_d) \in [-1, 1]^d$ by

$$K(u_1, \ldots, u_d) = \frac{\beta + 1}{\alpha\beta^2}(1 - |u|_\beta)_+, \tag{3}$$

where

$$\alpha = \frac{2^d \prod_{i=1}^{d} \Gamma(1/\beta_i)}{\Gamma(1/\beta) \prod_{i=1}^{d} \beta_i},$$

Γ denotes the gamma function and $|u|_\beta = \sum_{i=1}^d |u_i|^{\beta_i}$.

Lemma 1. *The kernel K satisfies $\int_{[-1,1]^d} K(u)\mathrm{d}u = 1$ and*

$$\int_{[-1,1]^d} K^2(u)\mathrm{d}u = \frac{2(\beta + 1)}{\beta\alpha(2\beta + 1)}.$$

This lemma is a consequence of Lemma 3 in the Appendix.

We consider the bandwidth $\tilde{h} = (h_1, \ldots, h_d)$, where

$$h_i = \left(\frac{C_0}{L_i}\left(\frac{\log n}{n}\right)^{\beta/(2\beta+1)}\right)^{1/\beta_i},$$

with

$$C_0 = \left(\sigma^{2\beta} L_* \left(\frac{\beta + 1}{\alpha\beta^3}\right)^\beta\right)^{1/(2\beta+1)}, \qquad L_* = \left(\prod_{i=1}^{d} L_j^{1/\beta_j}\right)^\beta.$$

Finally, we consider the kernel estimator

$$\hat{f}_n(t) = \frac{1}{h_1 \cdots h_d} \int_{[0,1]^d} K_{\tilde{h}}(u, t) \mathrm{d}Y_u, \tag{4}$$

defined for $t = (t_1, \ldots, t_d) \in [0, 1]^d$, where for $u = (u_1, \ldots, u_d) \in [0, 1]^d$,

$$K_{\tilde{h}}(u, t) = K\left(\frac{u_1 - t_1}{h_1}, \ldots, \frac{u_d - t_d}{h_d}\right) \prod_{i=1}^{d} g(u_i, t_i, h_i),$$

and

$$g(u_i, t_i, h_i) = \begin{cases} 1 & \text{if } t_i \in [h_i, 1 - h_i], \\ 2I_{[0,1]}\left(\dfrac{u_i - t_i}{h_i}\right) & \text{if } t_i \in [0, h_i), \\ 2I_{[-1,0]}\left(\dfrac{u_i - t_i}{h_i}\right) & t_i \in (1 - h_i, 1]. \end{cases}$$

We add the functions $g(u_i, t_i, h_i)$ to account for the boundary effects. Here and later I_A denotes the indicator of the set A. We suppose that n is large enough so that $h_i < \frac{1}{2}$, for $i = 1, \ldots, d$. Using a change of variables and the symmetry of the function K in each of its variables – i.e. for all $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$, $K(u_1, \ldots, u_d) = K(\ldots, u_{i-1}, -u_i, u_{i+1}, \ldots)$ – we obtain that

$$\frac{1}{h_1 \cdots h_d} \int_{[0,1]^d} K_{\tilde{h}}(u, t) \mathrm{d}u = \int_{[-1,1]^d} K(u) \mathrm{d}u = 1. \tag{5}$$

The main result of the paper is given in the following theorem.

Theorem 1. *Under the above assumptions, relation* (2) *holds for the estimator* $\hat{f}_n$ *defined in* (4) *with*

$$C(\beta, L, \sigma^2) = w(C_0).$$

Remark. For $d = 1$ the constant $w(C_0)$ coincides with that of Korostelev (1993).

We will prove this theorem in two stages. Let $0 < \varepsilon < \frac{1}{2}$. In Section 3, we show that $\hat{f}_n$ satisfies the upper bound

$$\limsup_{n\to\infty} \sup_{f \in \Sigma(\tilde{\beta}, L)} \mathbb{E}_f\left[w\left(\|\hat{f}_n - f\|_\infty \psi_n^{-1}\right)\right] \leqslant w(C_0(1 + \varepsilon)). \tag{6}$$

In Section 4, we prove the corresponding lower bound

$$\liminf_{n\to\infty} \inf_{\theta_n} \sup_{f \in \Sigma(\tilde{\beta}, L)} \mathbb{E}_f\left[w\left(\|\theta_n - f\|_\infty \psi_n^{-1}\right)\right] \geqslant w(C_0(1 - \varepsilon)). \tag{7}$$

Since $\varepsilon > 0$ in (6) and (7) can be arbitrarily small and w is a continuous function, this proves Theorem 1.

3. Upper bound

Define, for $t \in [0, 1]^d$ and $f \in \Sigma(\tilde{\beta}, L)$, the bias term,

$$b_n(t, f) = \mathbb{E}_f(\hat{f}_n(t)) - f(t),$$

and the stochastic term,

$$Z_n(t) = \hat{f}_n(t) - \mathbb{E}_f(\hat{f}_n(t)) = \frac{\sigma}{h_1 \cdots h_d \sqrt{n}} \int_{[0,1]^d} K_{\tilde{h}}(u, t) \mathrm{d}W_u.$$

Note that $Z_n(t)$ does not depend on f. Here we prove inequality (6).

Proposition 1. *The bias term satisfies*

$$\sup_{f \in \Sigma(\tilde{\beta}, L)} \psi_n^{-1} \|b_n(\cdot, f)\|_\infty \leqslant \frac{C_0}{2\beta + 1}.$$

Proof. Let $f \in \Sigma(\tilde{\beta}, L)$ and $t \in [0, 1]^d$. Suppose n large enough such that (5) is satisfied. Then

$$|\mathbb{E}_f(\hat{f}_n(t)) - f(t)| = \left| \frac{1}{h_1 \cdots h_d} \int_{[0,1]^d} K_{\tilde{h}}(u, t)(f(u) - f(t)) \mathrm{d}u \right|$$

$$\leqslant \frac{\sigma}{h_1 \cdots h_d} \int_{[0,1]^d} K_{\tilde{h}}(u, t) \left(\sum_{i=1}^d L_i |u_i - t_i|^{\beta_i} \right) \mathrm{d}u.$$

Then, using a change of variables and the symmetry of the function K in each of its variables, we have

$$|\mathbb{E}(\hat{f}_n(t)) - f(t)| \leqslant \frac{\beta + 1}{\alpha\beta^2} \sum_{i=1}^d L_i h_i^{\beta_i} B_i,$$

where

$$B_i = \int_{[-1,1]^d} |u_i|^{\beta_i} \left(1 - |u|_\beta\right) \mathrm{d}u = \frac{\alpha\beta^3}{\beta_i(\beta + 1)(2\beta + 1)},$$

the last equality being obtained from Lemma 3. Putting these inequalities together, we obtain, for all $t \in [0, 1]^d$,

$$|b_n(t, f)| \leqslant \frac{C_0}{2\beta + 1} \left(\frac{\log n}{n} \right)^{\beta/(2\beta+1)}.$$

□

Proposition 2. *The stochastic term satisfies, for any $z > 1$ and n large enough,*

$$\sup_{f\in\Sigma(\tilde{\beta},L)} \mathbb{P}_f\left[\psi_n^{-1}\|Z_n\|_\infty \geqslant \frac{2\beta C_0 z}{2\beta+1}\right] \leqslant D_1 n^{-(z^2-1)/(2\beta+1)}(\log n)^{1/2\beta+1},$$

where D_1 is a finite positive constant.

Proof. The stochastic term is a Gaussian process on $[0, 1]^d$. To prove this proposition, we use a more general lemma about the supremum of a Gaussian process (Lemma 4 in the Appendix). We have

$$\mathbb{P}_f\left[\psi_n^{-1}\|Z_n\|_\infty \geqslant \frac{2\beta C_0 z}{2\beta+1}\right] = \mathbb{P}_f\left[\sup_{t\in[0,1]^d}|\xi_t| \geqslant r_0\right],$$

with

$$r_0 = \frac{2\beta C_0 z\psi_n\sqrt{nh_1\cdots h_d}}{\sigma(2\beta+1)}$$

and

$$\xi_t = \frac{1}{\sqrt{h_1\cdots h_d}}\int_{[0,1]^d} K_{\tilde{h}}(u,\, t)\mathrm{d}W_u.$$

We will apply Lemma 4 to the process ξ_t on the sets Δ belonging to

$$S = \left\{\Delta = \prod_{i=1}^d \Delta_i : \Delta_i \in \{[0,\, h_i),\, [h_i,\, 1-h_i],\, (1-h_i,\, 1]\}\right\}.$$

Let $\Delta \in S$. The process ξ_t on Δ has the form

$$\xi_t = \frac{1}{\sqrt{h_1\cdots h_d}}\int_{[0,1]^d} Q\left(\frac{u_1-t_1}{h_1}, \ldots, \frac{u_d-t_d}{h_d}\right)\mathrm{d}W_u,$$

where $Q(u_1, \ldots, u_d) = K(u_1, \ldots, u_d)\prod_{i=1}^d g_i(u_i)$ and

$$g_i(u_i) = \begin{cases} 1 & \text{if } \Delta_i = [h_i,\, 1-h_i], \\ 2I_{[0,1]} & \text{if } \Delta_i = [0,\, h_i), \\ 2I_{[-1,0]} & \text{if } \Delta_i = (1-h_i,\, 1]. \end{cases}$$

The function Q satisfies $\|Q\|_2^2 = \int_{\mathbb{R}^d} Q^2 = \|K\|_2^2$. Moreover, we have the following lemma which will be proved in the Appendix.

Lemma 2. *There exists a constant $D_2 > 0$ such that, for all $t \in [-1, 1]^d$,*

$$\int_{\mathbb{R}^d}(Q(t+u)-Q(u))^2\,\mathrm{d}u \leqslant D_2\left(\sum_{i=1}^d |t_i|^{\min(1/2,\beta_i)}\right)^2. \tag{8}$$

The process ξ_t satisfies the conditions of Lemma 4 and in particular satisfies condition (12) of that lemma with $\alpha_i = \min(\frac{1}{2}, \beta_i)$ in view of Lemma 2. We have, by Lemma 3,

$$h = \prod_{i=1}^{d} h_i = \frac{C_0^{1/\beta}}{L_*^{1/\beta}} \left(\frac{\log n}{n}\right)^{1/(2\beta+1)}, \qquad \frac{r_0^2}{2\|K\|_2^2} = \frac{z^2 \log n}{2\beta+1}.$$

The condition $r_0 > c_2/|\log h|^{1/2}$ is then satisfied for n large enough. We obtain, for n large enough, that the quantity $N(h)$ (cf. Lemma 4) satisfies

$$\begin{aligned} N(h) &\leqslant \frac{D_3}{h}\left(|\log h|^{1/2}\right)^{1/\beta+1/2} \\ &\leqslant D_3 n^{1/(2\beta+1)}(\log n)^{1/2\beta+1}, \end{aligned}$$

where D_3 is a finite positive constant. Moreover the quantity $r_0/|\log h|^{1/2}$ is well defined and bounded independently of n, for n large enough. Then there exists $D_4 > 0$ such that

$$\mathbb{P}_f\left[\sup_{t\in\Delta}|\xi_t| \geqslant r_0\right] \leqslant D_4 n^{-(z^2-1)/(2\beta+1)}(\log n)^{1/2\beta+1}$$

and we obtain Proposition 2 by noting that $\text{card}(S) = 3^d$. □

We can now complete our proof of inequality (6). Let $\Delta_{n,f} = \psi_n^{-1}\|\hat{f}_n - f\|_\infty$ for $f \in \Sigma(\tilde{\beta}, L)$. We have, since w is non-decreasing,

$$\begin{aligned} \mathbb{E}_f\big(w(\Delta_{n,f})\big) &= \mathbb{E}_f\Big(w(\Delta_{n,f})I_{\{\Delta_{n,f}\leqslant(1+\varepsilon)C_0\}}\Big) + \mathbb{E}_f\Big(w(\Delta_{n,f})I_{\{\Delta_{n,f}>(1+\varepsilon)C_0\}}\Big) \\ &\leqslant w((1+\varepsilon)C_0) + \big(\mathbb{E}_f\big(w^2(\Delta_{n,f})\big)\big)^{1/2}\big(\mathbb{P}_f\big[\Delta_{n,f} > (1+\varepsilon)C_0\big]\big)^{1/2}. \end{aligned}$$

Therefore to prove inequality (6), it is enough to prove the following two relations:

(i) $\lim_{n\to\infty} \sup_{f\in\Sigma(\tilde{\beta},L)} \mathbb{P}_f\big[\Delta_{n,f} > (1+\varepsilon)C_0\big] = 0$;
(ii) there exists a constant D_5 such that $\limsup_{n\to\infty} \sup_{f\in\Sigma(\tilde{\beta},L)} \mathbb{E}_f\big(w^2(\Delta_{n,f})\big) \leqslant D_5$.

Let $f \in \Sigma(\tilde{\beta}, L)$. To prove (i), note that, for n large enough,

$$\mathbb{P}_f\big[\Delta_{n,f} > (1+\varepsilon)C_0\big] \leqslant \mathbb{P}_f\left[\psi_n^{-1}\|Z_n\|_\infty > \frac{2\beta C_0(1+\varepsilon)}{2\beta+1}\right],$$

which is a consequence of Proposition 1. By Proposition 2 with $z = 1+\varepsilon$, the right-hand side of this inequality tends to 0 as $n \to \infty$.

Let us prove (ii). The assumptions on w imply that there exist constants D_6 and D_7 such that

$$\mathbb{E}_f\big(w^2(\Delta_{n,f})\big) \leqslant D_6 + D_7\Big[\mathbb{E}_f\Big(\big(\psi_n^{-1}\|Z_n\|_\infty\big)^{2\gamma}\Big) + \big(\psi_n^{-1}\|b_n(\cdot, f)\|_\infty\big)^{2\gamma}\Big].$$

Using the fact that

$$\mathbb{E}_f\Big(\big(\psi_n^{-1}\|Z_n\|_\infty\big)^{2\gamma}\Big) = \int_0^{+\infty} \mathbb{P}_f\big[(\psi_n^{-1}\|Z_n\|_\infty)^{2\gamma} > t\big]\mathrm{d}t,$$

and Proposition 2, we prove that $\limsup_{n\to\infty} \mathbb{E}_f\Big[\big(\psi_n^{-1}\|Z_n\|_\infty\big)^{2\gamma}\Big] < \infty$. This and Proposition 1 entail (ii).

4. Lower bound

Before proving inequality (7), we need to introduce some notation and preliminary facts. We write

$$h = C_0^{1/\beta}\left(\frac{\log n}{n}\right)^{1/(2\beta+1)}, \qquad h_i = \left(\frac{C_0}{L_i}\right)^{1/\beta_i}\left(\frac{\log n}{n}\right)^{\beta/\beta_i(2\beta+1)}.$$

Let

$$m_i = \left[\frac{1}{2h_i(2^{1/\beta}+1)} - 1\right], \qquad M = \prod_{i=1}^{d} m_i,$$

where $[x]$ is the integer part of x. Consider the points $a(l_1, \ldots, l_d) \in [0, 1]^d$ for $l_i \in \{1, \ldots, m_i\}$ and $i \in \{1, \ldots, d\}$ such that

$$a(l_1, \ldots, l_d) = 2(2^{1/\beta}+1)(h_1 l_1, \ldots, h_d l_d).$$

To simplify the notation, we denote these points $a_1, \ldots, a_M$, and each a_j takes the form

$$a_j = (a_{j,1}, \ldots, a_{j,d}).$$

Let $\theta = (\theta_1, \ldots, \theta_M) \in [-1, 1]^M$. Denote by $f(\cdot, \theta)$ the function defined for $t \in [0, 1]^d$ by

$$f(t, \theta) = \sum_{j=1}^{M} \theta_j f_j(t),$$

where

$$f_j(t) = h^\beta \left(1 - \sum_{i=1}^{d}\left|\frac{t_i - a_{j,i}}{h_i}\right|^{\beta_i}\right)_+.$$

Define the set

$$\Sigma' = \left\{f_\theta : \theta \in [-1, 1]^M\right\}.$$

For all $\theta \in [-1, 1]^M$, $f(\cdot, \theta) \in \Sigma(\tilde{\beta}, L)$, therefore $\Sigma' \subset \Sigma(\tilde{\beta}, L)$.

Suppose that $f(\cdot) = f(\cdot, \theta)$, with $\theta \in [-1, 1]^M$, in model (1), and denote $\mathbb{P}_{f(\cdot,\theta)} = \mathbb{P}_\theta$. Consider the statistics

$$y_j = \frac{\int_{[0,1]^d} f_j(t)\mathrm{d}Y_t}{\int_{[0,1]^d} f_j^2(t)\mathrm{d}t}, \qquad j \in \{1, \ldots, M\}.$$

Proposition 3. *Let $f = f(\cdot, \theta)$ in model* (1).

(i) *For all $j \in \{1, \ldots, M\}$, y_j is a Gaussian variable with mean θ_j and variance equal to*

$$v_n^2 = \frac{2\beta + 1}{2\log n}.$$

(ii) *Moreover,* $\mathbb{P}_\theta$ *is absolutely continuous with respect to* $\mathbb{P}_0$ *and*

$$\frac{\mathrm{d}\mathbb{P}_\theta}{\mathrm{d}\mathbb{P}_0}(y) = \prod_{j=1}^{M} \frac{\varphi_{v_n}\big(y_j - \theta_j\big)}{\varphi_{v_n}\big(y_j\big)},$$

where φ_{v_n} *is the density of* $\mathcal{N}(0, v_n^2)$ *and* $\mathbb{P}_0 = \mathbb{P}_{(0,\ldots,0)}$.

Proof. (i) Let $j \in \{1, \ldots, M\}$. Since the functions f_j have disjoint supports, the statistic y_j is equal to

$$y_j = \theta_j + \frac{\sigma}{\sqrt{n}} \frac{\int_{[0,1]^d} f_j(t)\mathrm{d}W_t}{\int_{[0,1]^d} f_j^2(t)\mathrm{d}t}.$$

Since (W_t) is a standard Brownian sheet, y_j is Gaussian with mean θ_j and variance

$$\mathrm{var}\big(y_j\big) = \frac{\sigma^2}{n\int_{[0,1]^d} f_j^2(t)\mathrm{d}t} = \frac{\sigma^2}{nh^{2\beta}h_1 \cdots h_d I}, \tag{9}$$

where (see Lemma 3)

$$I = \int_{[-1,1]^d} \left(1 - \sum_{i=1}^{d} |t_i|^{\beta_i}\right)_+^2 \mathrm{d}t = \frac{2\alpha\beta^3}{(\beta+1)(2\beta+1)}. \tag{10}$$

Therefore

$$\mathrm{var}\big(y_j\big) = \frac{\sigma^2 L_*^{1/\beta}}{IC_0^{(2\beta+1)/\beta}\log n}.$$

Using the definition of C_0, we obtain (9).

(ii) Using Girsanov's theorem (see Gihman and Skorohod 1974, Chap. VII, Sect. 4), since the functions $f(\cdot, \theta)$ belong to $L^2\big([0, 1]^d\big)$, $\mathbb{P}_\theta$ is absolutely continuous with respect to $\mathbb{P}_0$ and we have

$$\frac{\mathrm{d}\mathbb{P}_\theta}{\mathrm{d}\mathbb{P}_0}(y) = \exp\left\{\frac{\sqrt{n}}{\sigma}\int f(t, \theta)\mathrm{d}W_t - \frac{n}{2\sigma^2}\int f^2(t, \theta)\mathrm{d}t\right\}.$$

Since the functions f_j have disjoint supports,

$$\frac{\mathrm{d}\mathbb{P}_\theta}{\mathrm{d}\mathbb{P}_0}(y) = \exp\left\{\frac{1}{v_n^2}\sum_{j=1}^{M}\theta_j y_j - \frac{1}{2v_n^2}\sum_{j=1}^{M}\theta_j^2\right\} = \prod_{j=1}^{M}\frac{\varphi_{v_n}\big(y_j - \theta_j\big)}{\varphi_{v_n}\big(y_j\big)}.$$

□

With these preliminaries, we can now prove inequality (7). For any $f \in \Sigma(\tilde{\beta}, L)$ and for any estimator θ_n, using the monotonicity of w and the Markov inequality, we obtain that

$$\mathbb{E}_f\left[w\left(\psi_n^{-1}\|\theta_n - f\|_\infty\right)\right] \geqslant w(C_0(1-\varepsilon))\mathbb{P}_f\left[\psi_n^{-1}\|\theta_n - f\|_\infty \geqslant C_0(1-\varepsilon)\right].$$

Since $\Sigma' \subset \Sigma(\tilde{\beta}, L)$, it is enough to prove that $\lim_{n\to\infty}\Lambda_n = 1$, where

$$\Lambda_n = \inf_{\theta_n} \sup_{f\in\Sigma'} \mathbb{P}_f\left[\psi_n^{-1}\|\theta_n - f\|_\infty \geqslant C_0(1-\varepsilon)\right].$$

We have $\max_{j=1,\ldots,M}|\theta_n(a_j) - f(a_j)| \leqslant \|\theta_n - f\|_\infty$. Setting $\hat{\theta}_j = \theta_n(a_j)C_0\psi_n$ and using the fact that $f(a_j, \theta) = C_0\psi_n\theta_j$ for $\theta \in [-1, 1]^M$, we see that

$$\Lambda_n \geqslant \inf_{\hat{\theta}\in\mathbb{R}^M} \sup_{\theta\in[-1,1]^M} \mathbb{P}_\theta(C_n),$$

where $C_n = \left\{\max_{j=1,\ldots,M}|\hat{\theta}_j - \theta_j| \geqslant 1-\varepsilon\right\}$ and $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_M) \in \mathbb{R}^M$ is measurable with respect to $y = \{Y_t,\ t \in [0, 1]^d\}$. We have

$$\Lambda_n \geqslant \inf_{\hat{\theta}\in\mathbb{R}^M} \int_{\{-(1-\varepsilon),1-\varepsilon\}^M} \mathbb{P}_\theta(C_n)\pi(\mathrm{d}\theta),$$

where π is the prior distribution on θ, $\pi(\mathrm{d}\theta) = \prod_{j=1}^M \pi_j(\mathrm{d}\theta_j)$, where π_j is the Bernoulli distribution on $\{-(1-\varepsilon), 1-\varepsilon\}$ that assigns probability $\frac{1}{2}$ to $-(1-\varepsilon)$ and to $(1-\varepsilon)$. Since $\mathbb{P}_\theta$ is absolutely continuous with respect to $\mathbb{P}_0$ (see Proposition 3), we have

$$\begin{aligned}\Lambda_n &\geqslant \inf_{\hat{\theta}\in\mathbb{R}^M} \int \mathbb{E}_0\left(I_{C_n}\frac{\mathrm{d}\mathbb{P}_\theta}{\mathrm{d}\mathbb{P}_0}\right)\pi(\mathrm{d}\theta)\\ &= \inf_{\hat{\theta}\in\mathbb{R}^M} \int \mathbb{E}_0\left(I_{C_n}\prod_{j=1}^M \frac{\varphi_{v_n}(y_j - \theta_j)}{\varphi_{v_n}(y_j)}\right)\pi(\mathrm{d}\theta).\end{aligned}$$

By the Fubini and Fatou theorems, we can write

$$\begin{aligned}\Lambda_n &\geqslant 1 - \sup_{\hat{\theta}\in\mathbb{R}^M} \int \frac{1}{\prod_{j=1}^M \varphi_{v_n}(y_j)} \left(\int \prod_{j=1}^M I_{\{|\theta_j - \hat{\theta}_j| < 1-\varepsilon\}}\varphi_{v_n}(y_j - \theta_j)\pi_j(\mathrm{d}\theta_j)\right)\mathrm{d}\mathbb{P}_0\\ &\geqslant 1 - \int \frac{1}{\prod_{j=1}^M \varphi_{v_n}(y_j)} \left(\prod_{j=1}^M \sup_{\hat{\theta}_j\in\mathbb{R}} \int I_{\{|\theta_j - \hat{\theta}_j| < 1-\varepsilon\}}\varphi_{v_n}(y_j - \theta_j)\pi_j(\mathrm{d}\theta_j)\right)\mathrm{d}\mathbb{P}_0.\end{aligned}$$

It is not hard to prove that the maximization problem

$$\max_{\hat{\theta}_j\in\mathbb{R}} \int I_{\{|\hat{\theta}_j - \theta_j| < 1-\varepsilon\}}\varphi_{v_n}(y_j - \theta_j)\pi_j(\mathrm{d}\theta_j)$$

admits the solution $\tilde{\theta}_j(y_j) = (1-\varepsilon)I_{\{y_j \geqslant 0\}} - (1-\varepsilon)I_{\{y_j < 0\}}$. By simple algebra, we obtain

$$\int I_{\{|\tilde\theta_j-\theta_j|<1-\varepsilon\}}\varphi_{v_n}(y_j-\theta_j)\pi_j(\mathrm{d}\theta_j)=\frac{1}{2}\Big(\varphi_{v_n}(y_j-(1-\varepsilon))I_{\{y_j\geqslant 0\}}+\varphi_{v_n}(y_j+(1-\varepsilon))I_{\{y_j<0\}}\Big).$$

Under $\mathbb{P}_0$, the random variables y_j are independently and identically Gaussian $\mathcal{N}(0, v_n^2)$. Thus

$$\begin{aligned}\Lambda_n &\geqslant 1-\prod_{j=1}^{M}\frac{1}{2}\int_{\mathbb{R}}\Big(\varphi_{v_n}(y_j-(1-\varepsilon))I_{\{y_j\geqslant 0\}}+\varphi_{v_n}(y_j+(1-\varepsilon))I_{\{y_j<0\}}\Big)\mathrm{d}y_j\\ &\geqslant 1-\left(\int_0^{+\infty}\varphi_{v_n}(y-(1-\varepsilon))\mathrm{d}y\right)^M\\ &\geqslant 1-\left(1-\Phi\left(-\frac{1-\varepsilon}{v_n}\right)\right)^M,\end{aligned}$$

where Φ is the standard normal cdf. Using the relation

$$\Phi(-z)\sim\frac{1}{z\sqrt{2\pi}}\exp\big(-z^2/2\big),\qquad \text{for } z\to+\infty,$$

and the definition of v_n, we have

$$\Phi\left(-\frac{1-\varepsilon}{v_n}\right)=\frac{v_n}{\sqrt{n\pi}(1-\varepsilon)}n^{-(1-\varepsilon)^2/(2\beta+1)}(1+o(1)).$$

Now $M\geqslant C'(n/\log n)^{1/(2\beta+1)}$, for some constant $C'>0$, therefore $\lim_{n\to\infty} M\Phi(-(1-\varepsilon)/v_n)=+\infty$ and

$$\lim_{n\to\infty}\left(1-\Phi\left(-\frac{(1-\varepsilon)}{v_n}\right)\right)^M=0,$$

which completes the proof of the lower bound.

Appendix

Proof of Lemma 2. Let $t\in[-1,1]^d$ and $u\in\mathbb{R}^d$. We have $Q(t+u)-Q(u)=D_8\big(\tilde Q(t+u)-\tilde Q(u)\big)$, where $\tilde Q(u)=(1-|u|_\beta)_+\prod_{i=1}^d I_{G_i}(u_i)$, with $G_i\in\{[0,1],\ [-1,1],\ [-1,0]\}$ and D_8 is a positive constant. We have:

- If $|t+u|_\beta\geqslant 1$ and $|u|_\beta\geqslant 1$, then $\tilde Q(t+u)-\tilde Q(u)=0$.
- If $|t+u|_\beta\leqslant 1$ and $|u|_\beta\geqslant 1$, then

$$0\leqslant\tilde Q(t+u)-\tilde Q(u)=1-|t+u|_\beta\leqslant|u|_\beta-|t+u|_\beta\leqslant|t|_\beta.$$

- If $|t+u|_\beta\geqslant 1$ and $|u|_\beta\leqslant 1$, then for the same reason $|\tilde Q(t+u)-\tilde Q(u)|\leqslant|t|_\beta$.

Thus to prove (8), it is enough to bound from above the integral

$$I(t) = \int (\tilde{Q}(t+u) - \tilde{Q}(u))^2 I_{A_t} \mathrm{d}u,$$

where $A_t = \{u \in \mathbb{R} : |t+u|_\beta \leqslant 1,\ |u|_\beta \leqslant 1\}$. We have $I(t) = B_1(t) + B_2(t)$, where

$$B_1(t) = \int \left(\tilde{Q}(t+u) - \tilde{Q}(u)\right)^2 I_{A_t \cap \tilde{A}_t} \,\mathrm{d}u,$$

$$B_2(t) = \int \left(\tilde{Q}(t+u) - \tilde{Q}(u)\right)^2 I_{A_t \cap \tilde{A}_t^C} \,\mathrm{d}u,$$

$$\tilde{A}_t = \{u \in \mathbb{R} : \tilde{Q}(u) \neq 0,\ \tilde{Q}(t+u) \neq 0\}.$$

We have

$$B_1(t) \leqslant 2^d \left(|t|_\beta\right)^2 \leqslant 2\left(\sum_{i=1}^{d} |t_i|^{\min(\beta_i, 1/2)}\right)^2,$$

since $\operatorname{mes}\{u \in \mathbb{R} : |u|_\beta \leqslant 1\} \leqslant 2$, where $\operatorname{mes}(\cdot)$ denotes the Lebesgue measure. Moreover, we have $\operatorname{mes}(A_t \cap \tilde{A}_t^C) \leqslant 2\sum_{i=1}^d |t_i|$ and then

$$B_2(t) \leqslant 2^d \sum_{i=1}^{d} |t_i| \leqslant D_9 \left(\sum_{i=1}^{d} |t_i|^{\min(\beta_i, 1/2)}\right)^2$$

with D_9 a positive constant. This completes the proof. □

The following lemma (Gradshteyn and Ryzhik 1965, formula 4.635.2) is stated without proof.

Lemma 3. *For a continuous function $f : \Delta_0 \to \mathbb{R}$, we have*

$$\int_{\Delta_0} f\left(x_1^{\beta_1} + \ldots + x_d^{\beta_d}\right) x_1^{p_1 - 1} \cdots x_d^{p_d - 1} \,\mathrm{d}x_1 \cdots \mathrm{d}x_d$$

$$= \frac{1}{\beta_1 \cdots \beta_d} \frac{\Gamma(p_1/\beta_1) \cdots \Gamma(p_d/\beta_d)}{\Gamma(p_1/\beta_1 + \ldots + p_d/\beta_d)} \int_0^1 f(x) x^{p_1/\beta_1 + \ldots + p_d/\beta_d - 1} \,\mathrm{d}x,$$

where

$$\Delta_0 = \left\{(x_1, \ldots, x_d) \in [0, 1]^d : x_1^{\beta_1} + \ldots + x_d^{\beta_d} \leqslant 1\right\}$$

and the β_i and p_i are positive numbers.

Lemma 4. *Let $Q : \mathbb{R}^d \to \mathbb{R}$ be a function such that $\|Q\|_2^2 = \int_{\mathbb{R}^d} Q^2 < \infty$, Δ be a compact set $\Delta = \prod_{i=1}^d \Delta_i$ with Δ_i intervals of $[0, +\infty)$ of length $T_i > 0$, and W be the standard Brownian sheet on Δ. Let $h_1, \ldots, h_d$ be arbitrary positive numbers and write $h = \prod_{i=1}^d h_i$. We consider the Gaussian process defined for $t = (t_1, \ldots, t_d) \in \Delta$:*

$$X_t = \frac{1}{\sqrt{h_1 \cdots h_d}} \int_{\mathbb{R}^d} Q\left(\frac{u_1 - t_1}{h_1}, \ldots, \frac{u_d - t_d}{h_d}\right) \mathrm{d}W_u, \tag{11}$$

with $u = (u_1, \ldots, u_d)$. *Let* $(\alpha_1, \ldots, \alpha_d) \in (0, \infty)^d$ *and let* α *be the number such that* $1/\alpha = \sum_{i=1}^d 1/\alpha_i$. *Let* $T = \prod_{i=1}^d T_i$. *We suppose that there exists* $0 < c_1 < \infty$ *such that, for* $t \in [-1, 1]^d$,

$$\int_{\mathbb{R}^d} (Q(t+u) - Q(u))^2 \,\mathrm{d}u \leqslant \left(c_1 \sum_{i=1}^d |t_i|^{\alpha_i}\right)^2. \tag{12}$$

Then there exists a constant $c_2 > 0$, *such that, for* $b \geqslant c_2/|\log h|^{1/2}$ *and* h *small enough,*

$$\mathbb{P}\left[\sup_{t \in \Delta} |X_t| \geqslant b\right] \leqslant N(h) \exp\left(-\frac{b^2}{2\|Q\|_2^2}\right) \exp\left(\frac{c_2 b}{\|Q\|_2^2 |\log h|^{1/2}}\right), \tag{13}$$

where $c_2 = c_3(c_4 + 1/\sqrt{\alpha})$, c_3 *and* c_4 *do not depend on* $h_1, \ldots, h_d$, T *and* α, $\mathbb{P}$ *denotes the distribution of* $\{X_t, t \in \Delta\}$, *and*

$$N(h) = 2 \prod_{i=1}^d \left(\frac{T_i}{h_i}\left(c_1 d |\log h|^{1/2}\right)^{1/\alpha_i} + 1\right).$$

Note that if the $h_i/T_i \to 0$, *then, for* h_i/T_i *small enough,*

$$N(h) \leqslant 2^{d+1} \frac{T}{h}\left(c_1 d |\log h|^{1/2}\right)^{1/\alpha}.$$

This lemma is close to various results on the supremum of Gaussian processes (see Adler 1990; Lifshits 1995; Piterbarg 1996). The closest result is Theorem 8-1 of Piterbarg (1996) which, however, cannot be used directly since there is no explicit expression for the constants that in our case depend on h and T and may tend to 0 or ∞. Also the explicit dependence of the constants on α is given here. This can be useful for the purpose of adaptive estimation.

Proof. Let $\lambda > 0$ and $N_1(\lambda, S)$ be the minimal number of hyperrectangles with edges of length $h_1(\lambda/c_1 d)^{1/\alpha_1}, \ldots, h_d(\lambda/c_1 d)^{1/\alpha_d}$ that cover a set $S \subset \Delta$. We have

$$N_1(\lambda, \Delta) \leqslant \prod_{i=1}^d \left(\frac{T_i}{h_i}\left(\frac{c_1 d}{\lambda}\right)^{1/\alpha_i} + 1\right).$$

Denote by $B_1, \ldots, B_{N_1(\lambda, \Delta)}$ such hyperrectangles that cover Δ and choose $\lambda = |\log h|^{-1/2}$, well defined for $h < 1$. We have, for $b \geqslant 0$,

$$\mathbb{P}\left[\sup_{t \in \Delta} |X_t| \geqslant b\right] \leqslant \sum_{j=1}^{N_1(\lambda, \Delta)} \mathbb{P}\left[\sup_{t \in B_j} |X_t| \geqslant b\right]. \tag{14}$$

Let $j \in \{1, \ldots, N_1(\lambda, \Delta)\}$. Using Theorem 12.2 and Lemma 12.2 of Liftshits (1995), we obtain, for $b \geqslant \mu$,

$$\mathbb{P}\left[\sup_{t\in B_j}|X_t| \geqslant b\right] \leqslant 2\mathbb{P}\left[\sup_{t\in B_j}X_t \geqslant b\right] \leqslant 2\exp\left(-\frac{1}{2\sigma^2}(b-\mu)^2\right), \tag{15}$$

where $\sigma^2 = \sup_{t\in\Delta}\mathbb{E}(X_t^2)$ and $\mu = \sup_j \mathbb{E}\big(\sup_{t\in B_j}X_t\big)$. Let us evaluate σ^2. We have, by a change of variables,

$$\sigma^2 = \sup_{t\in\Delta}\frac{1}{h_1\cdots h_d}\int_\Delta Q^2\left(\frac{u_1-t_1}{h_1},\ldots,\frac{u_d-t_d}{h_d}\right)\mathrm{d}u \leqslant \|Q\|_2^2. \tag{16}$$

Before evaluating μ, we need to introduce a semi-metric ρ on Δ defined by

$$\rho(s,\,t) = \left(\mathbb{E}\left[(X_s - X_t)^2\right]\right)^{1/2}, \qquad s,\,t\in\Delta,$$

where $\mathbb{E}$ is the expectation with respect to $\mathbb{P}$. Let $s,\,t\in B_j$. For h small enough, we have $|(s_i-t_i)/h_i|<1$ and, using (12) and a change of variables, we obtain

$$\rho(s,\,t) \leqslant c_1\sum_{i=1}^d\left|\frac{s_i-t_i}{h_i}\right|^{\alpha_i}. \tag{17}$$

Theorem 14.1 of Lifshits (1995) implies

$$\mathbb{E}\left(\sup_{t\in B_j}X_t\right) \leqslant 4\sqrt{2}\int_0^{\sigma/2}(\log N_{B_j}(u))^{1/2}\,\mathrm{d}u,$$

where $N_{B_j}(u)$ is the minimal number of ρ-balls of radius u necessary to cover B_j. In view of (17), we have a rough bound, for h small enough,

$$N_{B_j}(u) \leqslant N_1(u,\,B_j) \leqslant \prod_{i=1}^d\left(1+\left(\frac{\lambda}{u}\right)^{1/\alpha_i}\right).$$

Thus, for h small enough,

$$\mu = \mathbb{E}\left(\sup_{t\in B_j}X_t\right) \leqslant 4\sqrt{2}\int_0^\lambda[\log(N_1(u,\,B_j))]^{1/2}\,\mathrm{d}u \leqslant 4\lambda\sqrt{2}\int_0^1\left[\sum_{i=1}^d\log(1+u^{-1/\alpha_i})\right]^{1/2}\mathrm{d}u$$

$$\leqslant 4\lambda\sqrt{2}\sum_{i=1}^d\int_0^1[\log(1+u^{-1/\alpha_i})]^{1/2}\,\mathrm{d}u.$$

Here

$$\int_0^1[\log(1+u^{-1/\alpha_i})]^{1/2}\,\mathrm{d}u = \int_0^1\left[\log\left(1+u^{1/\alpha_i}\right)-\frac{1}{\alpha_i}\log u\right]^{1/2}\mathrm{d}u$$

$$\leqslant \sqrt{\log 2}+\frac{1}{\sqrt{\alpha_i}}\int_0^1|\log x|^{1/2}\,\mathrm{d}x.$$

Then we have

$$\mu \leqslant \lambda c_3(c_4 + 1/\sqrt{\alpha}) = c_2\lambda, \tag{18}$$

where c_3 and c_4 are positive constants independent of j, T, h and α. Substituting (15), (16) and (18) into inequality (14), we obtain, for $b \geqslant c_2\lambda$ and for h small enough,

$$\begin{aligned} \mathbb{P}\left[\sup_{t\in\Delta}|X_t| \geqslant b\right] &\leqslant 2N_1(\lambda, \Delta)\exp\left(-\frac{1}{2\|Q\|_2^2}(b-\mu)^2\right), \\ &\leqslant N(h)\exp\left(-\frac{b^2}{2\|Q\|_2^2}\right)\exp\left(\frac{c_2\lambda b}{\|Q\|_2^2}\right). \end{aligned}$$

Then for $b \geqslant c_2/|\log h|^{1/2}$ and for h small enough, we obtain (13). □

Ackowledgements

I would like to thank Professor Alexandre Tsybakov for a wealth of advice and encouragement.

References

Adler, R.J. (1990) *An Introduction to Continuity, Extrema, and Related Topics for general Gaussian Processes*, IMS Lecture Notes – Monogr. Ser. 12. Hayward, CA: Institute of Mathematical Statistics.

Barron, A., Birgé, L. and Massart, P. (1999) Risk bounds for model selection via penalization. *Probab. Theory Related Fields*, **113**(3), 301–413.

Bertin, K. (2004) Minimax exact constant in sup-norm for nonparametric regression with random design. *J. Statist. Plann. Inference*. To appear.

Donoho, D.L. (1994) Asymptotic minimax risk for sup-norm loss: solution via optimal recovery. *Probab. Theory Related Fields*, **99**(2), 145–170.

Gihman, I.I. and Skorohod, A.V. (1974) *The Theory of Stochastic Processes. I.* New York: Springer-Verlag.

Gradshteyn, I.S. and Ryzhik, I.M. (1965) *Table of Integrals, Series, and Products*, Fourth edition prepared by Ju. V. Geronimus and M. Ju Ceîvtlin. Translated from the Russian by Scripta Technica, Inc. Translation edited by Alan Jeffrey, Academic Press, New York.

Ibragimov, I.A. and Has'minskii, R.Z. (1981) *Statistical Estimation. Asymptotic Theory*. New York: Springer-Verlag.

Kerkyacharian, G., Kepski, O. and Picard, D. (2001) Nonlinear estimation in anisotropic multi-index denoising. *Probab. Theory Related Fields*, **121**(2), 137–170.

Klemelä, J. and Tsybakov, A.B. (2001) Sharp adaptive estimation of linear functionals. *Ann. Statist.*, **29**(6), 1567–1600.

Korostelev, A.P. (1993) An asymptotically minimax regression estimator in the uniform norm up to a constant. *Teor. Veroyatnost. i. Primenen.*, **38**(4), 857–882.

Korostelev, A. and Nussbaum, M. (1999) The asyptotic minimax constant for sup-norm loss in nonparametric density estimation. *Bernoulli*, **5**(6), 1099–1118.

Lepski, O.V. and Tsybakov, A.B. (2000) Asymptotically exact nonparametric hypothesis testing in sup-norm and at a fixed point. *Probab. Theory Related Fields*, **117**(1), 17–48.

Lepskii, O.V. (1992) On problems of adpative estimation in white Gaussian noise. In R.Z. Khasminskij (ed.), *Topics in Nonparametric Estimation*, Adv. Soviet Math. 12, pp. 87–106. Providence, RI: American Mathematical Society.

Lifshits, M.A. (1995) *Gaussian Random Functions*. Dordrecht: Kluwer Academic.

Neumann, M.H. and von Sachs, R. (1997) Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra. *Ann. Statist.*, **25**(1), 38–76.

Nussbaum, M. (1986) On the nonparametric estimation of regression functions that are smooth in a domain in $\mathbf{R}^k$. *Teor. Veroyatnost. i Primenen.*, **31**(1), 118–125.

Piterbarg, V.I. (1996) *Asymptotic Methods in the Theory of Gaussian Processes and Fields*, Transl. Math. Monogr. 148. Providence, RI: American Mathematical Society.

Stone, C.J. (1982) Optimal global rates of convergence for nonparametric regression. *Ann. Statist.*, **10**(4), 1040–1053.

Received March 2003 and revised February 2004