New method of estimation of the smoothness parametr

Bogdan Ćmiel
Karol Dziedziul
Barbara Wolnik

Israeli-Polish Mathematical Meeting
11-15 IX 2011

Smoothness of function in terms of Besov spaces

Let us consider a function $f \in L^2(R)$, $p = 2$. Since $B^s_{2,q} \subset B^s_{2,\infty}$ we consider only $B^s_{2,\infty}$. From the following continuous embedding

$$B^{s_1}_{2,\infty} \subset B^{s_2}_{2,\infty} \text{ for } s_1 > s_2$$

we have: f belongs to each $B^s_{2,\infty}$ or f belongs to none of $B^s_{2,\infty}$ or there is the parameter $s^* = s^*(f)$ such that

$$\text{for all} \quad s < s^*, \quad f \in B^s_{2,\infty}$$

and

$$\text{for all} \quad s > s^*, \quad f \notin B^s_{2,\infty}.$$

In the last case we say that s^* is the smoothness parameter of the function f.

Multiresolution analysis

Let

$$\ldots \subset V_0 \subset V_1 \subset V_2 \subset \ldots$$

be a sequence of nested subspaces of $L^2(\mathbb{R})$ generated by a scaling function ϕ. Let

$$Q_j = P_{j+1} - P_j \quad j \in \mathbb{Z}$$

is the orthogonal projection on W_j - the orthogonal complement to V_j inside V_{j+1}, i.e

$$W_j \oplus V_j = V_{j+1}$$

and

$$Q_j : L^2(R) \to W_j$$

$$P_j : L^2(R) \to V_j.$$

The Besov spaces $B^s_{2,\infty}$ are characterized by operators Q_j only for $0 < s < 1/2$ in Haar case, B.I. Golubov 1972, for full scale of Besov spaces in S. Ropella. 1976.

Determination of smoothness of function

Let ϕ be a scaling function and ψ - the wavelet associated with ϕ. If $r \geq 1$ is a natural number then we will say that ϕ forms a r-regular multiresolution approximation (we will use r-RMA) if

(i) $\phi, \psi \in C^r$ and the support of each of them is compact,

(ii) (zero oscillation condition) there is $d \geq r$

$$\int_{\mathbb{R}} x^k \psi(x) dx = 0 \quad \text{for} \quad 0 \leq k \leq d \quad \text{and} \quad \int_{\mathbb{R}} x^{d+1} \psi(x) dx \neq 0.$$

Note

$$Q_j f = \sum_{k \in Z} < f, \psi_{j,k} > \psi_{j,k}.$$

Theorem (Kucharska,Wolnik,D)

Let $0 < s^* < r$ be such that for all $s < s^*$ $f \in B^s_{2,\infty}$ and for $s > s^*$ $f \notin B^s_{2,\infty}$. Then

$$\liminf_{j \to \infty} \frac{-\log_2^+ \|Q_j f\|_2}{j} = s^*, \tag{1}$$

where $\log_2^+$ means that we take only nonzero arguments.

Density estimator. Histogram

Let $X_1, X_2, \ldots$ be a sequence of iid rvs with unknown density $f \in L^2(R)$. The density estimator is defined by

$$f_{h,n}(x) = \frac{1}{n}\sum_{j=1}^{n} K_h(x, X_j).$$

where the kernel is given by

$$K_h(x,y) = \frac{1}{h}K(x/h, y/h)$$

$$K(x,y) = \sum_{k\in\mathbb{Z}} \phi(x-k)\phi(y-k).$$

Now if $h = 2^{-j}$ then

$$Ef_{h,n} = P_j f = \int_R K_{2^{-j}}(x,y)f(y)dy.$$

In case ψ is Haar function $f_{h,n}$ is the histogram.

Estimation of smoothness of density

Let us denote the density-histogram for parameters $h = 2^{-j}$ and $n = 2^{2(r+1/2)j}$ by f_j, i.e.

$$f_j = f_{2^{-j}, 2^{2(r+1/2)j}}.$$

Theorem (Kucharska,Wolnik,D)

Let be given r-RMA. Let $X_1, X_2, \ldots$ be a sequence of iid random variables with density $f \in L^2(\mathbb{R})$. Moreover let $0 < s^* < r$ be such that for all $s < s^*$ $f \in B^s_{2,\infty}$ and for all $s > s^*$ $f \notin B^s_{2,\infty}$. Then

$$\liminf_{j \to \infty} \frac{-\log_2^+ \|f_j - f_{j-1}\|_2}{j} = s^* \quad a.e. \tag{2}$$

Questions

(i) Indicate a class of function for which the estimator is consistent (Ćmiel)

(ii) Asymptotic distribution of smoothness estimator in Besov spaces(Wolnik)

(iii) Find more efficient new estimators

Answer (i) and (iii)

Assume that there exists $0 < \delta < 1$ such that

$$|\psi(x)| \neq 0, \quad x \in (0, 1+\delta], \tag{3}$$

where ψ is a smooth wavelet that satisfies the conditions: $\psi \in C^r, \mathrm{supp}\ \psi = [0; S(r)]$. This assumption is satisfied by all Daubechies wavelets.

Theorem

If f is B-Spline of degree $S < r$ and $\psi(x)$ satisfies above assumption, then

$$\lim_{j \to \infty} \frac{-\log_2^+ \|\beta_{j\cdot}\|_2}{j} = S + \frac{1}{2}.$$

where $\beta_{jk} := \int_{\mathbb{R}} \psi_{jk}(x) f(x) dx.$

Answer (i) and (iii) SLLN

Let $X_1, X_2, \ldots$ be a sequence of i.i.d random variables with density f.

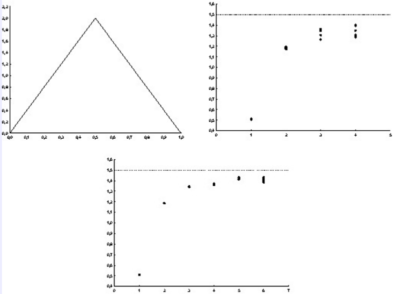
Theorem

Let a density f be B-spline of degree $S < r$. Then

$$\lim_{j\to\infty} \frac{-\log_2^+ \|\hat{\beta}_{j\cdot}\|_2}{j} = S + \frac{1}{2} = s^*(f) \quad a.e.\ ,$$

where $\hat{\beta}_{jk} = \frac{1}{n}\sum_{i=1}^{n} \psi_{jk}(X_i)$, *and* $n \asymp 2^{2j(r+1/2)}$.

Numerical Experiment



For an estimation the Daubechies wavelets DB12 with the support $[0; S(2)] = [0; 23]$ were used. From top: Density function, 6 estimators of the smoothness parameter for $n = 2^{16}$, 6 estimators of the smoothness parameter for $\mathrm{n} = 2^{24}$.

Answer (ii)

Note

$$\|f_{2^{-(j+1)},n} - f_{2^{-j},n}\|_2^2 == \frac{2}{n^2}\sum_{i<s}^{n} G_j(X_i, X_s) + \sum_{i=1}^{n} G_j(X_i, X_i),$$

where

$$G_j(x,y) = \sum_k \psi_{j,k}(x)\psi_{j,k}(y).$$

Moreover

$$EG_j(X_1, X_2) = \|Q_j f\|_2^2.$$

So to improve estimation take

$$U_{n,j} = \frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{l=i+1}^{n}\left(G_j(X_i, X_l) - \|Q_j f\|_2^2\right).$$

U-statitsics CLT

Berry -Esseen inequality

Theorem (Berry -Esseen inequality)

Let Φ denote the Normal Standard distribution function. Let $X_1, X_2, \ldots, X_n$ be a sequence of iid random variables and let U_n be given by

$$U_n = \binom{n}{2}^{-1} \sum_{1 \le l < s \le n} h(X_l, X_s),$$

where $h(x, y)$ is a symmetric, real-valued function. Let $Eh(X_1, X_2) = 0$, $\sigma^2 = Eh^2(X_1, X_2) < \infty$ and $\tilde{\sigma}^2 = Eg^2(X_1) > 0$ where $g(x) = E(h(X_1, X_2)|X_1 = x)$. If $(E|g(X_1)|^3 < \infty$ then

$$\sup_{z \in \mathbb{R}} \left| P\left(\frac{\sqrt{n}}{2\tilde{\sigma}} U_n \le z\right) - \Phi(z) \right| \le \frac{6.1 E|g(X_1)|^3}{\sqrt{n}\tilde{\sigma}^3} + \frac{(1+\sqrt{2})\sigma}{\sqrt{2(n-1)}\tilde{\sigma}}$$

Asymptotic formulas

Theorem

Let r-RMA be given. Let d be vanishing number of the wavelets moments. If $f \in C^{d+1}$ is nonnegative and has compact support, then there exist a natural number N and constants $C_1, C_2 > 0$ such that for all $j > N$

$$C_1 \int_{\mathbb{R}} |Q_j f(x)|^2 dx \leq \int_{\mathbb{R}} |Q_j f(x)|^2 f(x) dx \leq C_2 \int_{\mathbb{R}} |Q_j f(x)|^2 dx. \quad (4)$$