CONSTRUCTIVE THEORY OF FUNCTIONS'87

SOFIA, 1988

Local spline approximation and nonparametric density estimation

Z. CIESIELSKI

Abstract. *This note gives the order of approximation of L^p functions on the real line by certain linear local positive spline operators. By means of these operators characterisations of Lipschits classes and of the saturation class are given. The linear method in question is used to construct simple for computation estimators of densities and distributions. Moreover, it follows that this method of estimation is optimal as far as the mean L^p deviation is considered.*

1. Introduction. In this note we consider approximation by splines of order $r \geq 1$ i.e. of degree $r-1$ corresponding to the uniform mesh $\{t_i = ih,\ i \in Z\}$, with the step $h > 0$, where Z is the set of all integers. It is assumed that all the knotes are simple. All the B-splines corresponding to this mesh can be defined by means of the *cardinal* spline of order r

$$N^{(r)}(x) = r[0, \ldots, r; (\cdot - x)_+^{r-1}],$$

where $[s_0, \ldots, s_r; f]$ denotes the devided difference of f taken at the points $s_0, \ldots, s_r$. Namely,

$$N_{i,h}^{(r)}(x) = N^{(r)}(\frac{x - t_i}{h}).$$

It is well known that

$$\sum_{i \in Z} N_{i,h}^{(r)}(x) = 1 \quad for \quad x \in R, \tag{1.1}$$

where $R = (-\infty, \infty)$. For later convenience introduce

$$M_{i,h}^{(r)}(x) = \frac{1}{h} N_{i,h}^{(r)}(x).$$

Now, since

$$\int_R N^{(r)}(x)\, dx = 1,$$

it follows that

$$\int_R M_{i,h}^{(r)}(x)\, dx = 1. \tag{1.2}$$

The approximating operators can now be defined for any $f \in L^1_{loc}(R)$ as follows

$$Q_h^{(r)}(f; x) = \sum_{i \in Z} (f, M_{i,h}^{(r)}) N_{i,h}^{(r)}(x), \tag{1.3}$$

where

$$(f,g)=\int_R f(x)g(x)\,dx.$$

Since the coefficient functional

$$f \mapsto (f, M_{i,h}^{(r)})$$

is local for each admissible pair i, h i.e. its support is contained in the support of $N_{i,h}^{(r)}$, these operators are *local spline operators* in de Boor's terminology [2]. It also follows by (1.1) and (1.2) that these operators have in addition the following properties:$Q_h^{(r)}(f) \geq 0$ for $f \geq 0$ and $Q_h^{(r)}(1) = 1$. It is important to realise that they are not projections except for the case $r = 1$ (see [5]) but their L^p norms are equal to 1,i.e. for all $1 \leq p \leq \infty$

$$\|Q_h^{(r)}(f)\|_p \leq \|f\|_p \qquad \text{for} \quad f \in L^p(R),$$

where

$$\|f\|_p = \left(\int_R |f(x)|^p\,dx\right)^{\frac{1}{p}}.$$

Moreover, for $f \in L^1(R)$

$$\int_R f = 1 \qquad \text{implies} \qquad \int_R M_{i,h}^{(r)} = 1.$$

Consequently, the operator

$$Q_h^{(r)} : L^1(R) \to L^1(R)$$

takes probability densities into probability densities, and it appears to be good for constructing nonparametric density estimators.

We now assume that we are given a probability space $(\Omega, \mathcal{F}, Pr)$ and a simple sample of size n i.e. a sequence $X_1, \ldots, X_n$ of i.i.d. real valued random variables such that their common distribution has a density f. The standard way of producing estimators for f is given by formula

$$f_{h,n}(x) = \frac{1}{n}\sum_{j=1}^{n} Q_h^{(r)}(x, X_j), \tag{1.4}$$

where

$$Q_h^{(r)}(x,y) = \sum_{i\in Z} M_{i,h}^{(r)}(x) N_{i,h}^{(r)}(y). \tag{1.5}$$

Clearly, the kernel $Q_h^{(r)}(x,y)$ is nonnegative and symmetric, and in each variable it is a linear combination of B-splines. Thus, its values can be calculated by the numerically stable algorithm due to C. de Boor, M. G. Cox and L. Mansfield (see e.g.[2]).

Next section contains the main result on order of approximation by the local spline operators in question. Section 3 contains one of the results on estimation of densities. More details on nonparametric density estimation by splines will be published elswhere in our joint work with G. Krzykowski.

2. Order of approximation by the local spline operators. In order to state the results we recall the definition of the modulus of smoothness of order k

$$\omega_{k,p}(f;\delta) = \sup_{|t|<\delta} \|\Delta_t^k f\|_p,$$

where Δ_t^k is the k-th order progressive difference with step t.

To prove the main theorem of this section the following is needed.

LEMMA 2.1. *Let f satisfy the Lipschitz condition*

$$\omega_{1,\infty}(f;\delta) \leq M\cdot\delta \qquad \textit{for } \delta > 0.$$

Then

$$\|f - Q_h^{(r)}(f)\|_\infty \leq M\cdot r\cdot h \qquad \textit{for } h > 0.$$

PROOF: We may assume without loss of generality that for given $x \in R$ there is $i \in Z$ such that $t_i < x < t_{i+1}$. Now,

$$f(x) - Q_h^{(r)}(f)(x) = \sum_{i-r<j\leq i} N_{j,h}^{(r)}(x)\int_{t_j}^{t_{j+r}} M_{j,h}^{(r)}(y)(f(x)-f(y))\,dy,$$

$$|f(x) - Q_h^{(r)}(f)(x)| \leq M \sum_{i-r<j\leq i} N_{j,h}^{(r)}(x)\int_{t_j}^{t_{j+r}} M_{j,h}^{(r)}(y)|x-y|\,dy,$$

$$\leq M\cdot r\cdot h\sum_{i-r<j\leq i} N_{j,h}^{(r)}(x)\int_{t_j}^{t_{j+r}} M_{j,h}^{(r)}(y)\,dy = M\cdot r\cdot h.$$

In what follows $W_p^k(R)$ denotes the Sobolev space over R with the index of smoothness k and the exponent of integration p.

LEMMA 2.2. *Let $r \geq 2$ and let f_0 be a linear function on R. Then*

$$Q_h^{(r)}(f_0) = f_0.$$

PROOF: We know already that the operator $Q_h^{(r)}$ reproduces constants. It is therefore sufficient to prove the Lemma for $f_0(x) = x$. However, by known formula (see e.g.[8])

$$f_0(x) = \sum_{i\in Z} h(i+\frac{r}{2})N_{i,h}^{(r)}(x).$$

On the other hand,Peano's formula for divided difference gives

$$\int_R yM_{i,h}^{(r)}(y)\,dy = \frac{1}{1+r}[ih,\ldots,(i+r)h;s^{r+1}] = (i+\frac{r}{2})h,$$

whence by the previous identity and by the definition of $Q_h^{(r)}$ we infer the desired equality.

Next lemma was earlier applied in [7] (see also [1]) in the case of Bernstein polynomials and as it appears it is natural to use the same idea in the case of our local spline operators.

LEMMA 2.3. *Let $f \in W_p^2(R)$ and let support $f \subset < a, b >$ for some finite $a, b \in R$. Then for all $x \in R$ we have*

$$f(x) = -\int_R K_{a,b}(x,y) D^2 f(y)\,dy,$$

where $D = \frac{d}{dx}$ and for $y \in < a, b >$

$$K_{a,b}(x.y) = \frac{1}{b-a} min(x-a, y-a) min(b-x, b-y).$$

PROOF: Using Peano's formula for divided differences

$$[a, x, b; f] = \frac{1}{2}\int_R M(a, x, b; y) D^2 f(y)\,dy$$

we find

$$f(x) = \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) - \int_a^b K_{a,b}(x,y) D^2 f(y)\,dy,$$

and this completes the proof.

THEOREM 2.4. *Let $r \geq 2$, p, $1 \leq p \leq \infty$, be given and let $f \in L^p$. Then*

$$\|f - Q_h^{(r)}(f)\|_p \leq 8(2r^2 + 1)\omega_{2,p}(f; h) \qquad \textit{for} \quad 0 < h < 1.$$

PROOF: Since the L^p-norms of the operators $Q_h^{(r)}$ are equal to 1 and $\omega_{2,p}(f;h) \leq 4\|f\|_p$ it is sufficient to prove the result for $f \in L^p(R)$ with compact support. Using Steklov means we define in the standard way the smoothing of f of order k,i.e.

$$g(x) = -\sum_{j=1}^{k} (-1)^{j+k} \binom{k}{j} \int_0^1 \cdots \int_0^1 f(x + jh(s_1 + \cdots + s_k))\,ds_1 \ldots ds_k.$$

The function g is in $W_p^k(R)$. If the support of f is contained in $< a', b' >$, then the support of g is contained in $< a, b >$ with $b' = b$ and $a = a' - k^2 h$. Moreover,

$$f(x) - g(x) = \int_0^1 \cdots \int_0^1 \Delta^k_{h(s_1+\cdots+s_k)} f(x)\,ds_1 \ldots ds_k,$$

whence

$$\|f - g\|_p \leq k^k \omega_{k,p}(f; h). \tag{2.5}$$

It also follows that

$$D^k g(x) = -\sum_{j=1}^{k} (-1)^{j+k} \frac{1}{(jh)^k} \binom{k}{j} \Delta^k_{jh} f(x).$$

Thus,

$$h^k \|D^k g\|_p \le 2^k \omega_{k,p}(f;h). \tag{2.6}$$

In what follows it is enough to take $k = 2$. The next step is to prove

$$\|g - Q_h^{(r)}(g)\|_p \le (2r^2)h^2 \|D^2 g\|_p. \tag{2.7}$$

The locallity of the operators $Q_h^{(r)}$ and Lemmas 2.2 and 2.3 imply

$$Q_h^{(r)}(g)(x) - g(x) = \int_{|x-y|<rh} \left(K_{a,b}(x,y) - Q_h^{(r)}(K_{a,b}(\cdot,y)(x)\right) D^2 g(y)\, dy\,. \tag{2.8}$$

Now, the function $K_{(a,b)}(\cdot, y)$ for fixed $y \in <a,b>$ satisfies Lipschits condition in the L^∞ norm with the constant $M = 1$. Lemma 2.1 and (2.8) then give

$$|Q_h^{(r)} g(x) - g(x)| \le 2r^2 h^2 \frac{1}{2rh} \int_{|x-y|<rh} |D^2 g(y)|\, dy,$$

and this implies (2.7). Applying now the triangle inequality to the identity

$$f - Q_h^{(r)}(f) = (f - g) + (g - Q_h^{(r)}(g)) + Q_h^{(r)}(g - f)$$

we obtain

$$\|f - Q_h^{(r)}(f)\|_p \le 2\|f - g\|_p + \|g - Q_h^{(r)}(g)\|_p. \tag{2.10}$$

Now, (2.10), (2.5), (2.6) and (2.7) give the inequality stated in the theorem.

Using the same methods as in [3], [4] we are able to establish the following

THEOREM 2.11. *Let us introduce for given natural r and k the equivalence*

$$\|f - Q_h^{(r)}\|_p = O(h^\alpha) \quad as \quad h \to 0_+ \iff \omega_{k,p}(f;h) = (h^\alpha) \quad as \quad h \to 0_+. \tag{2.12}$$

Then, condition (2.12) holds in the following domains:
(i) $r \ge 2,\ k = 2,\ 0 < \alpha < 1 + \frac{1}{p},\ 1 \le p < \infty.$
(ii) $r \ge 3,\ k = 2,\ 0 < \alpha < 2,\ 1 \le p \le \infty.$

3. Density estimation. To formulate the last result we introduce for $0 < \alpha \le 2$ and $1 \le p \le \infty$

$$\psi(\alpha, p) = \begin{cases} \frac{1}{2\alpha+1}, & \text{if } 1 \le p < 2; \\ \frac{1}{p\alpha+p-1}, & \text{if } 2 \le p < 2 + \frac{1}{\alpha+1}; \\ \frac{1}{2\alpha+2}, & \text{if } p \ge 2 + \frac{1}{\alpha+1}. \end{cases}$$

Theorem 3.1. *Let $1 \le p \le \infty$ and let f be a density on R. Let α, $0 < \alpha < 2$, and β, $0 < \beta \le \psi(\alpha, p)$ be given. Moreover, let $h = n^{-\beta}$. Then the following conditions are equivalent:*

(i) $$\omega_{3,p}(f;\delta) = O(\delta^{\alpha}) \quad as \quad \delta \to 0_+,$$

(ii) $$\left(E\|f - f_{m,n}\|_p^p\right)^{\frac{1}{p}} = O\left(\frac{1}{n^{\alpha\beta}}\right) \quad as \quad n \to \infty,$$

where E is the expectation operator with respect to the given probability space.

Example: Direct compution shows that for the arcsin law density, i.e. for

$$f_0(x) = \frac{1}{\pi}\frac{1}{\sqrt{x(1-x)}}, \qquad x \in <0,1>,$$

and f_0 being zero elswhere, we have for $1 \le p < 2$ and for each $k \ge 1$

$$\omega_{k,p}(f_0;\delta) \sim \delta^{\frac{1}{p}-\frac{1}{2}} \quad as \quad \delta \to 0_+.$$

Thus, for this density $\alpha = \frac{1}{p} - \frac{1}{2}$ and the optimal choice for β is $\beta = \frac{p}{2}$.

The most interesting is the case $p = 1$. The optimal choice for β, given α in $(0, 2>$, is $\beta = \frac{1}{2\alpha+1}$. In particular, for $\alpha = 2$ we obtain the celebrated exponent $\alpha\beta = \frac{2}{5}$. For more detailed discussian of these questions we refer e.g. to [6].

Aknowledgement. In this place the author would like to thank to J.Koronacki for enlighting discusions concernig the estimation questions.

References

[1] H.Berens and G.Lorentz, *Inverse theorems for Bernstein polynomials*, Ind.Univ.M.J. **21** (1973), 693-708.

[2] C. de Boor, *Linear combinations of B-splines*. Approximation Theory II. Acad.Press 1976, 1-47.

[3] Z.Ciesielski, *Properties of the orthonormal Franklin system*, II, Studia Math. **27** (1966), 290-323.

[4] Z.Ciesielski, *Constructive function theory and spline systems*, Studia Math. **53** (1975), 277-302.

[5] Z.Ciesielksi, *Haar System and Nonparametric Estimation in Several Variables*, Prob. and Math.Stat. (to appear).

[6] L.Devroye and L.Györfi, L_1 theory of the nonparametric density estimation, 1986.

[7] K.Ivanov, *Approximation by Bernstein polynomials in L_p metric*. Constructive Theory of Functions. Sofia 1984, 421-429.

[8] L.L.Schumaker, Spline functions: basic theory. John Wiley 1981.

Instytut Matematyczny PAN
Oddział w Gdańsku
ul. Abrahama 18
81-825 Sopot, POLAND