Good approximation by splines with variable knots+

Carl de Boor*

Consider approximation of a given function f, on $[0,1]$ say, by elements of S_π^k, i.e., by **polynomial splines of order** j (or, degree $< k$) **on some partition** $\pi = (t_i)_0^{N+1}$ of $[0,1]$,

$$0 = t_0 < t_1 \le t_2 \le \cdots \le t_N < t_{N+1} = 1.$$

Here, $t_1, \ldots, t_N$ are the **knots** or **joints** of $s \in S_\pi^k$, and their multiplicity, i.e., equality among two or more of these, indicates reduced smoothness at that knot location in the usual way.

Best approximation to f by elements of S_π^k is quite well understood for a variety of norms since, after all, S_π^k is a *linear* space. It seems practically more important and theoretically more interesting to investigate the approximation to f by **splines of order** k **with** N **knots**, i.e., by elements of

$$S_N^k := \cup_\pi S_\pi^k$$

where the union is taken over all partitions π of $[0,1]$ with N interior knots. For, the approximation power of spline functions seems to lie precisely in the possibility of placing the knots in a usually quite nonuniform way to suit the peculiarities of the given f. Yet the straightforward approach, *viz.* the construction of a best approximation to f in S_N^k, has turned out to be beset with difficulties. It being a somewhat nasty *nonlinear* problem, no satisfactory characterization of a best approximation can be found in general, see e.g., [4] for the case of Chebyshev approximation. Consequently, any computational scheme has to be content to find, by some descent method, a *locally best* approximation, and even that seems to be computationally quite expensive. Also, the function f may be "given" in a way that makes the calculation of best knots impossible simply because $\|f - s\|$ cannot be calculated. E.g., f may be the unique solution of some differential equation

$$D^m f(t) = F(t, f(t), \ldots, D^{m-1} f(t)), \quad \text{for} \quad t \in [0,1]$$

$$\text{with side conditions} \quad \beta_i f = c_i, \quad i = 1, \ldots, m$$

where F, the linear functionals $\beta_1, \ldots, \beta_m$ and the numbers $c_1, \ldots, c_m$ are known, but the value of f at t is not.

For these and other reasons, it becomes important to search for methods which will produce relatively cheaply *good*, if not best, knots for the approximation of a given function from a variety of information about this function. And the literature concerning bounds on

$$\operatorname{dist}_\infty(f, S_N^k) := \inf_{s \in S_N^k} \|f - s\|_\infty$$

seems to be a good place to start such a search.

One approach, taken, e.g., by Freud and Popov [7], [8], and by Sendov and Popov [16], has been to reduce the problem of estimating $\operatorname{dist}(f, S_N^k)$ to the simpler problem of estimating $\operatorname{dist}(f, S_N^2)$ for given $g \in C[0,1]$, making use of a fact such as the following.

+ appeared in ISNM Vol.21, Spline functions and Approximation Theory, Birkhäuser Verlag, Basel, 1973, 57–72, except that, in the present version, certain typos have been corrected and the bibliography has been completed.

* This work was supported in part by NSF Grant GP-07163

LEMMA 1. *For every $f \in C^{(k-2)}[0,1]$, and every partition $\pi = (t_i)_0^{N+1}$ for $[0,1]$,*

$$\operatorname{dist}_\infty(f, S_\pi^k) \leq k! |\pi|^{k-2} \operatorname{dist}_\infty(f^{(k-2)}, S_\pi^2)/2^{k-1} \tag{1}$$

with $|\pi| := \max_i \Delta t_i$.

A simple proof of this lemma goes as follows: On $C[0,1]$, define the linear map P by

$$P : C[0,1] \to S_\pi^k : f \mapsto \sum_j f(\tau_i) N_{i,k}$$

with $(N_{i,k})$ the normalized B-spline basis for S_π^k (see, e.g., [3]). Since the $N_{i,k}$ are nonnegative and add up to 1 at any particular point, in then follows that

$$\begin{aligned} |(f - Pf)(t)| &= |\sum_i (f(t) - f(\tau_i)) N_{i,k}(t)| \\ &\leq \max\{|f(t) - f(\tau_i)| : N_{i,k}(t) \neq 0\}. \end{aligned}$$

On the other hand, since $N_{i,k}$ is nonzero only on (t_i, t_{i+k}), it is possible to choose τ_i in $[0,1]$ so that

$$\text{for all} \quad t \in [0,1], \quad N_{i,k}(t) \neq 0 \quad \Longrightarrow \quad |t - \tau_i| \leq |\pi|/2.$$

With such a choice, one then obtains

$$\operatorname{dist}_\infty(f, S_\pi^k) \leq \|f - Pf\|_\infty \leq (k/2)\omega_f(|\pi|)$$

ω_f being the modulus of continuity of f; hence, for $f \in C^{(1)}[0,1]$, and arbitrary $s \in S_\pi^k$,

$$\begin{aligned} \operatorname{dist}_\infty(f, S_\pi^k) &= \operatorname{dist}_\infty(f - s, S_\pi^k) \leq \frac{k}{2}\omega_{f-s}(|\pi|) \\ &\leq \frac{k}{2}|\pi| \|f^{(1)} - s^{(1)}\|_\infty, \end{aligned}$$

therefore

$$\operatorname{dist}_\infty(f, S_\pi^k) \leq (k/2)|\pi| \operatorname{dist}_\infty(f^{(1)}, S_\pi^{k-1}),$$

making use of the fact that $S_\pi^{k-1} = \{s^{(1)} : s \in S_\pi^k\}$; and repeated application of this last inequality gives the estimate (1).

Choosing, in particular π so that

$$\operatorname{dist}_\infty(f^{(k-2)}, S_\pi^2) = \operatorname{dist}_\infty(f^{(k-2)}, S_N^2)$$

and then augmenting π by at most $N-1$ points to insure that

$$|\pi| \leq 1/N,$$

one obtains from (1) the estimate, valid for $f \in C^{(k-2)}[0,1]$:

$$\operatorname{dist}_\infty(f, S_{2N-1}^k) \leq k!\, N^{-k+2} \operatorname{dist}_\infty(f^{(k-2)}, S_N^2)/2^{k-1}. \tag{2}$$

The simpler problem of best approximation by broken lines, i.e., in S_n^2, is taken care of by a result such as the following

LEMMA 2. *For every $g \in \mathrm{AC}$ with $g' \in \mathrm{BV}$*

$$\operatorname{dist}_\infty(g, S^2_{N-1}) \le N^{-2}\operatorname{Var}(g')/4.$$

This can be found, e.g., in [17] as a special case of a much more general result, but can also be proved directly as follows: If the straight line s interpolates f at the points $a < b$, then

$$(f-s)(t) = f[a,b,t](t-a)(t-b) = \frac{f[b,t]-f[a,t]}{b-a}(t-a)(t-b)$$

with $f[r_0,\ldots,r_k]$ denoting the k-th divided difference of f at $r_0,\ldots,r_k$. It follow that

$$\sup_{a<t<b} |(f-s)(t)| \le (b-a)/4 \;\mathrm{Osc}_{[a,b]} f' \le (b-a)/4 \int_a^b |\,\mathrm{d}f'|$$

if $f \in \mathrm{AC}$ and $f' \in \mathrm{BV}$, where

$$\mathrm{Osc}_{[a,b]} g := \operatorname{ess.sup}_{[a,b]} g \;-\; \operatorname{ess.inf}_{[a,b]} g.$$

Hence, if such f is approximated by the broken line $s \in S^2_{N-1}$ which interpolates f at $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = 1$, and the t_i's are chosen so that

$$(t_{i+1}-t_i)\int_{t_i}^{t_{i+1}} |\,\mathrm{d}f'| = \alpha, \quad \text{all} \;\; i,$$

for some α, then

$$\operatorname{dist}_\infty(f, S^2_{N-1}) \le \|f-s\|_\infty \le \alpha/4$$

while

$$\begin{aligned} \frac{(1-0)}{N^2}\operatorname{Var}(f') &= \Big(\frac{1}{N}\sum_i \Delta t_i\Big)\Big(\frac{1}{N}\sum_i \int_{t_i}^{t_{i+1}} |\,\mathrm{d}f'|\Big) \\ &= \Big(\frac{1}{N}\sum_i \Delta t_i\Big)\Big(\frac{1}{N}\sum_i \frac{\alpha}{\Delta t_i}\Big) \\ &\ge \Big(\frac{1}{N}\sum_i \Delta t_i\Big)\Big(\frac{\alpha}{\frac{1}{N}\sum_i \Delta t_i}\Big) \end{aligned}$$

by Jensen's Inequality, since $1/x$ is convex for $x > 0$. This proves the lemma.

The two lemmata have the desired

COROLLARY. *If $f \in C^{(k-2)}[0,1]$, with $f^{(k-2)} \in \mathrm{AC}$ and $f^{(k-1)} \in \mathrm{BV}$, then*

$$\operatorname{dist}_\infty(f, S^k_N) \le \mathrm{const}_k N^{-k}\operatorname{Var}(f^{(k-1)}). \tag{3}$$

This is to be compared with the customary statement that

$$\operatorname{dist}_\infty(f, S^k_\pi) \le \mathrm{const}_k |\pi|^k \|f^{(k)}\|_\infty \tag{4}$$

in case $f \in C^{(k-1)}[0,1]$ with $f^{(k-1)} \in \mathrm{AC}$ and $f^{(k)} \in L_\infty$.

But, although this improvement of (3) over (4) was achieved by a particular choice of knots, the argument has to be suspect since it relies on choosing the knots so as to produce a good approximation to $f^{(k-2)}$ rather than to f. Even the more sophisticated argument of Subbotin and Chernykh [17] (who obtain (3) by constructing an approximation to f in the spirit of Birkhoff's local spline approximation by moments [1], [2], followed by an appropriate choice of knots so as to make the error small) excludes consideration of such practically interesting functions as

$$f(t) = t^{\alpha}, \quad \text{some} \quad \alpha \in (0,1)$$

and therefore does not give, e.g., Rice's startling result [10] that

$$\text{for} \quad f(t) = t^{\alpha} \quad \text{with} \quad 0 < \alpha, \quad \operatorname{dist}_{\infty}(f, S_N^k) \le \text{const}_{\alpha,k} N^{-k}$$

Rice's argument is a direct verification that for a certain set of knots selected according to a rule depending on α everything works out. In an attempt to generalize Rice's result, H. Burchard [5] proved the following intriguing

THEOREM 1. *For $f \in C^{(k)}[0,1]$, and $N \ge N_f$, and for $1 \le p \le \infty$,*

$$\operatorname{dist}_p(f, S_N^k) \le \text{const}_k N^{-k} \|f^{(k)}\|_{\sigma}$$

where

$$\sigma = \sigma_{p,k} := 1/(k + 1/p).$$

Similar results have been obtained, for the special case $p = 2$, by Sacks and Ylvisaker [12-15], and more or less by McClure [9], again dealing only with $f \in C^{(k)}$ or even $f \in C^{(k+1)}$, and therefore not giving Rice's result (5). Nevertheless, these considerations bring out the importance of the σ-norm of $f^{(k)}$ for $\alpha < 1$ in the discussion of $\operatorname{dist}_p(f, S_N^k)$ and suggest that, e.g., (5) holds because, for $f(t) = t^{\alpha}$,

$$\|f^{(k)}\|_{1/k} \sim \left(\int_0^1 t^{(\alpha-k)/k}\,\mathrm{d}t\right)^k$$

is finite. This is confirmed by the following

THEOREM 2. *If $f \in C[0,1] \cap C^{(k)}(0,1]$, and $|f^{(k)}(t)|$ is monotone decreasing, then*

$$\operatorname{dist}_{\infty}(f, S_N^k) \le \text{const}_k N^{-k}\, \|f^{(k)}\|_{1/k}.$$

This can be proved as follows: Consider approximation to f in S_{π}^k, where π has $N-1$ distinct points in $(0,1)$,

$$t_1 < t_2 < \cdots < t_{N-1}$$

say, but each repeated k times. Then

$$\operatorname{dist}_{\infty}(f, S_{k(N-1)}^k) \le \operatorname{dist}_{\infty}(f, S_{\pi}^k) \le \|f - Tf\|_{\infty}$$

with Tf the piecewise polynomial function of order k which, on (t_i, t_{i+1}), agrees with the Taylor series for f around t_{i+1} up to terms of order $< k$, hence, for $t \in [t_i, t_{i+1}]$,

$$\begin{aligned}
|f(t) - Tf(t)| &= \frac{1}{(k-1)!}\left|\int_{t_{i+1}}^{t} f^{(k)}(r)(t-r)^{k-1}\,\mathrm{d}r\right| \\
&\le \frac{1}{(k-1)!}\int_{t_i}^{t_{i+1}} |f^{(k)}(r)|(t-r)^{k-1}\,\mathrm{d}r \\
&\le \frac{1}{k!}\left(\int_{t_i}^{t_{i+1}} |f^{(k)}(r)|^{1/k}\,\mathrm{d}r\right)^k \;=:\; \frac{1}{k!}\beta_i.
\end{aligned}$$

The last inequality is easy to prove if, as we assume, $|f^{(k)}|$ decreases monotonely – consider both sides as a function of t_{i+1} and differentiate – but impossible to find in the literature. In any event, choose now the t_i's so that the β_i's defined above are all constant, $\beta_i = \beta$ for all i. Then

$$N\beta^{1/k} = \int_0^1 |f^{(k)}(r)|^{1/k}\,\mathrm{d}r,$$

hence

$$\|f - Tf\|_\infty \le \frac{1}{k!}\,N^{-k}\|f^{(k)}\|_{1/k}$$

which proves the theorem in view of the fact that $\mathrm{dist}_\infty(f, S_N^k)$ decreases with increasing N.

A similar result holds for $\mathrm{dist}_p(f, S_N^k)$ with $p < \infty$, as proved by D.S. Dodson in [6], where one can also find the following

THEOREM 3. *If, for some converging net π of partitions of $[0,1]$, the corresponding lower Riemann sums for*

$$\|f^{(k)}\|_\sigma^\sigma = \int_0^1 |f^{(k)}(r)|^\sigma\,\mathrm{d}r \quad \text{with} \quad \sigma := 1/(k + 1/p)$$

converge to A, then

$$\liminf_N \mathrm{dist}_p(f, S_N^k)N^k \ge \mathrm{const}_k A^{1/\sigma}$$

for some positive constant const_k independent of (π) and f.

These facts and arguments suggest that in approximating f by elements of S_N^k, one should choose the N knots $t_1, \ldots, t_N$ so as to make

$$\int_{t_i}^{t_{i+1}} |f^{(k)}(r)|^{1/k}\,\mathrm{d}r$$

approximately constant as a function of i. This has been tried by Dodson [6] in a scheme for the adaptive solution of an ordinary differential equation. From a current piecewise polynomial approximation of order k to the solution f, he guesses a piecewise constant approximation g to $f^{(k)}$, and then selects a new knot set so as to equalize $\int |g(r)|^{1/k}\,\mathrm{d}r$ over subintervals. To give an example, Russell and Shampine [11] solve the problem

$$\varepsilon f''(t) - (2 - t^2)f(t) = -1 \quad \text{on} \quad [-1,1] \quad \text{with} \quad f(-1) = f(1) = 0$$

with $\varepsilon = 10^{-8}$ by collocation, using splines of order 6 with 47 distinct knots, each of multiplicity 3. The knots are placed in an *ad hoc* basis so as to pile up near ± 1. They obtain an approximation error of $5 \cdot 10^{-4}$ near the boundary. Dodson obtains the same accuracy with 19 distinct knots, and obtains, with 47 knots, an accuracy of $2 \cdot 10^{-6}$ even near the boundary (and an 10^{-8} error in the middle of the interval).

REFERENCES

1. G. Birkhoff (1967), "Local spline approximation by moments", *J. Math. Mech.* **16**, 987–990.
2. C. de Boor (1968), "On local spline approximation by moments", *J. Math. Mech.* **17**, 729–735.
3. C. de Boor (1972), "On calculating with B-splines", *J. Approx. Theory* **6**, 50–62.
4. D. Braess (1971), "Chebyshev approximation by spline functions with free knots", *Numer. Math.* **17**, 357–366.
5. H. G. Burchard (1974), "Splines (with optimal knots) are better", *Appl. Anal.* **3**, 309–319.

6. D. S. Dodson (1972), "Optimal order approximation by polynomial spline functions", dissertation, Purdue Univ..
7. G. Freud and V. Popov (1970), "Some questions connected with spline functions (Russian)", *Studia Sci. Math. Hung.* **5**, 161–171.
8. G. Freud and V. Popov (19xx), "On approximation by splines functions", in *Proceedings of the Conference on Constructive Theory of Functions* xxx (xxx), 163–172.
9. D. E. McClure (1970), "Feature selection for the analysis of line patters", dissertation, Div. Appl. Math, Brown U. (Providence RI).
10. J. R. Rice (1969), "On the degree of convergence of nonlinear spline approximation", in *Approximation with Special Emphasis on Spline Functions* (I. J. Schoenberg Ed.), Academic Press (New York), 349–365.
11. R. D. Russell and L. F. Shampine (1972), "A collocation method for boundary value problems. (I. and II.)", *Numer. Math.* **19**, 1–28.
12. J. Sachs and D. Ylvisaker (1966), "Designs for regression problems with correlated errors", *Annals of Mathematical Statistics* **37(1)**, 66–89.
13. J. Sachs and D. Ylvisaker (1968), "Designs for regression problems with correlated errors: Many parameters", *Annals of Mathematical Statistics* **39(1)**, 46–69.
14. J. Sachs and D. Ylvisaker (1970), "Designs for regression problems with correlated errors III", *Annals of Mathematical Statistics* **41(6)**, 2057–2074.
15. J. Sachs and D. Ylvisaker (1970), "Statistical design and integral approximation", in *Proc. 12th biennial Seminar of the Canadian Mathematical Congress* xxx (xxx), 115–136.
16. B. Sendov and V. A. Popov (1970), "Approximation of curves by piecewise polynomial curves", *C. R. Acad. Bulg. Sci.* **23**, 639–642.
17. Y. N. Subbotin and N. I. Chernykh (1970), "Order of the best spline approximations of some classes of functions", *Math. Notes* **7**, 20–26.