JOURNAL OF APPROXIMATION THEORY **62**, 372–380 (1990)

On Piecewise-Polynomial Approximation of Functions with a Bounded Fractional Derivative in an L_p-Norm

G. W. WASILKOWSKI*

Department of Computer Science, University of Kentucky, Lexington, Kentucky 40506, U.S.A.

Communicated by Günther Nürnberger

Received September 26, 1988

We study the error in approximating functions with a bounded $(r+\alpha)$th derivative in an L_p-norm. Here r is a nonnegative integer, $\alpha \in [0, 1)$, and $f^{(r+\alpha)}$ is the classical fractional derivative, i.e., $f^{(r+\alpha)}(y) = \int_0^1 (y-t)_+^{-\alpha}\, d(f^{(r)}(t))$. We prove that, for any such function f, there exists a piecewise-polynomial of degree s that interpolates f at n equally spaced points and that approximates f with an error (in sup-norm) $\|f^{(r+\alpha)}\|_p\, O(n^{-(r+\alpha-1/p)})$. We also prove that no algorithm based on n function and/or derivative values of f has the error equal $\|f^{(r+\alpha)}\|_p\, o(n^{-(r+\alpha-1/p)})$ for any f. This implies the optimality of piecewise-polynomial interpolation. These two results generalize well-known results on approximating functions with bounded rth derivative ($\alpha = 0$). We stress that the piecewise-polynomial approximation does not depend on α nor on p. It does not depend on the exact value of r as well; what matters is an upper bound s on r, $s \geq r$. Hence, even without knowing the actual regularity (r, α, and p) of f, we can approximate the function f with an error equal (modulo a constant) to the minimal worst case error when the regularity were known. © 1990 Academic Press, Inc.

1. INTRODUCTION

For a nonnegative integer r and $p \in [1, +\infty]$, let W_p^r be the Sobolev space, i.e., it consists of all functions $f\colon [0, 1] \to \Re$ with $(r-1)$st derivative absolutely continuous and $f^{(r)} \in L_p$.

It is well known that for any points $0 \leq t_1 < \cdots < t_n \leq 1$ and any function $f \in W_p^r$, the error in approximating f by a piecewise-polynomial $P_{r-1}(f)$ of degree $r-1$ interpolating f at the points t_i is proportional to $\|f^{(r)}\|_p\, \Delta_n^{1/p-r}$, i.e.,

$$\sup_{f \in W_p^r} \frac{\|f - P_{r-1}(f)\|_{\sup}}{\|f^{(r)}\|_p} = O(\Delta_n^{1/p-r}).$$

Here, $\Delta_n = \max_i |t_i - t_{i-1}|$ with $t_0 = 0$ and $t_{n+1} = 1$.

* Partially supported by the National Science Foundation under Grant CCR-86-03674.

0021-9045/90 $3.00

It is also known, see, e.g., [1, 6] or [3, 5] and the papers cited there, that no algorithm that is based only on n function and/or derivative evaluations has the error smaller than $O(n^{1/p-r})$. Hence, modulo a constant, an algorithm provided by piecewise-polynomial interpolation at equally spaced points is optimal.

In this paper we generalize these results by considering functions that have $(r+\alpha)$th derivative bounded in L_p-norm, where α is an arbitrary real number from $[0, 1)$. The class of such functions we denote by $W_p^{r,\alpha}$; see Section 2 for a formal definition.

In particular, in Theorem 1 we show that for any r, α, and p, the piecewise-polynomial interpolation of degree s at n equally spaced points yields an algorithm with the error

$$\sup_{f \in W_p^{r,\alpha}} \frac{\|f - P_s(f)\|_{\sup}}{\|f^{(r+\alpha)}\|_p} = O(n^{-(r+\alpha-1/p)}) \qquad \text{if} \quad s \geqslant r.$$

In Theorem 2 we show that no algorithm based on n function and/or derivative evaluations has the worst case error for the class $W_p^{r,\alpha}$ smaller than $O(n^{1/p-r-\alpha})$. This, with Theorem 1, implies optimality (modulo a constant) of P_s at equally spaced points for $s \geqslant r$.

We summarize the contents of the paper. Theorems 1 and 2 are presented in Sections 2 and 3, respectively. We conclude the paper, Section 4, with comments concerning applications of Theorems 1 and 2 to other problems.

2. Error of Piecewise-Polynomial Interpolation

For given a nonnegative integer r, $\alpha \in [0, 1)$, and $p \in [1, \infty]$, let

$$W_p^{r,\alpha} = \{f\colon [0, 1] \to \Re : f^{(r)} \text{ is abs. cont. and } \|T_{r,\alpha}(f)\|_p < \infty\}.$$

Here $T_{r,\alpha}(f)$, the fractional derivative $f^{(r+\alpha)}$, is defined by

$$T_{r,\alpha}(f)(y) = \int_0^1 (y-t)_+^{-\alpha}\, d(f^{(r)}(t)).$$

Equivalently,

$$T_{r,\alpha}(f)(y) = P_\alpha(f^{(r+1)})(y)$$

and P_α is a fractional derivative (or integral) operator,

$$P_\alpha(g)(y) = \frac{-1}{1-\alpha} \int_0^1 g(t)\, d((y-t)_+^{1-\alpha}).$$

Observe that for $\alpha = 0$, $T_{r,\alpha}(f)(y) = f^{(r)}(y) - f^{(r)}(0)$. Hence, $W_p^{r,0}$ is a dense subspace of W_p^r.

For given n and $0 \leqslant t_1 < \cdots < t_n \leqslant 1$, let N_n be the information consisting of function values at points t_i,

$$N_n(f) = [f(t_1), \ldots, f(t_n)],$$

and let $A_{s,n}(N(f))$ be a piecewise-polynomial of degree s that interpolates f. We measure the (worst case) error of $A_{s,n}$ in the class $W_p^{r,\alpha}$ by

$$e(A_{s,n}, N_n; r, \alpha, p) = \sup_{f \in W_p^{r,\alpha}} \frac{\|f - A_{s,n}(N_n(f))\|_{\sup}}{\|T_{r,\alpha}(f)\|_p}.$$

THEOREM 1. *Let $r + \alpha - 1/p \geqslant 0$. Then for any n, $0 \leqslant t_1 < \cdots < t_n \leqslant 1$, and $s \geqslant r$,*

$$e(A_{s,n}, N_n; r, \alpha, p) = O(\Delta_n^{s+\alpha-1/p} \delta_n^{r-s}) \qquad as \quad n \to \infty,$$

where $\Delta_n := \max_{1 \leqslant i \leqslant n+1}(t_i - t_{i-1})$ *and* $\delta_n := \min_{2 \leqslant i \leqslant n}(t_i - t_{i-1})$ *with* $t_0 = 0$ *and* $t_{n+1} = 1$.

Proof. We begin with $s = r$.

For an arbitrary $x \in [0, 1]$, let $t_i, t_{i+1}, \ldots, t_{i+r}$ be the interpolation points used by the algorithm $A_{s,n}$ to approximate $f(x)$. That is,

$$A_{s,n}(N_n(f))(x) = p_i(x) = \sum_{j=i}^{i+r} f[t_i, \ldots, t_j] \prod_{k=i}^{j-1} (x - t_k).$$

Obviously, $x \in [t_i - \Delta_n/2, t_{i+r} + \Delta_n/2]$, and for $x \notin \{t_i, \ldots, t_{i+r}\}$ the error at x equals

$$e(x) = f(x) - p_i(x) = a(x) \prod_{j=i}^{r+i} (x - t_j) \tag{2.1}$$

with

$$a(x) = f[t_i, \ldots, t_{r+i}, x] = \frac{f[t_i, \ldots, t_{r+i-1}, x] - f[t_{i+1}, \ldots, t_{r+i-1}, t_{r+i}, x]}{t_i - t_{r+i}}.$$

It is easy to check that for any y and x,

$$f^{(r)}(y) - f^{(r)}(x) = c_\alpha[P_{1-\alpha}(T_{r,\alpha}(f))(y) - P_{1-\alpha}(T_{r,\alpha}(f))(x)] \tag{2.2}$$

with

$$c_\alpha = \int_0^\infty \frac{1}{t^\alpha(1+t)}\, dt.$$

This and the well-known formula for divided differences,

$$f[x_0, ..., x_k] = \int_0^1 \int_0^{\lambda_1} \cdots \int_0^{\lambda_{k-1}} f^{(k)}\left(x_0 + \sum_{j=1}^{k} \lambda_j(x_j - x_{j-1})\right) d\lambda_k \cdots d\lambda_1,$$

yield

$$a(x) = \frac{c_\alpha}{t_{r+i} - t_i} \int_0^1 \int_0^{\lambda_1} \cdots \int_0^{\lambda_{r-1}}$$
$$\times \left[\int_0^1 T_{r,\alpha}(f)(z)((y_1 - z)_+^{\alpha-1} - (y_2 - z)_+^{\alpha-1})\, dz\right] \times d\lambda_r \cdots d\lambda_1,$$

where

$$y_1 = t_{i+1} + \lambda_1(t_{i+2} - t_{i+1}) + \cdots + \lambda_r(x - t_{r+i})$$

and

$$y_2 = t_i + \lambda_1(t_{i+1} - t_i) + \cdots + \lambda_r(x - t_{r+i-1}).$$

Note that the integral over $z \in [0, 1]$ reduces to an integral over $z \in [0, t_{r+i} + \Delta_n/2]$ since y_1 and y_2 are always not greater than $t_{r+i} + \Delta_n/2$. Hence, changing the order of integration, we estimate $|a(x)|$ by

$$|a(x)| \leqslant \frac{c_\alpha}{t_{r+i} - t_i}(I_1 + I_2),$$

where

$$I_1 = \int_0^{t_i - \Delta_n} |T_{r,\alpha}(f)(z)| \left|\int_0^1 \int_0^{\lambda_1} \cdots \int_0^{\lambda_{r-1}} ((y_1 - z)^{\alpha-1} - (y_2 - z)^{\alpha-1})\, d\lambda_r \cdots d\lambda_1\right| dz$$

and

$$I_2 = \int_{t_i - \Delta_n}^{t_{r+i} + \Delta_n/2} |T_{r,\alpha}(f)(z)| \left|\int_0^1 \int_0^{\lambda_1} \cdots \int_0^{\lambda_{r-1}} ((y_1 - z)_+^{\alpha-1} - (y_2 - z)_+^{\alpha-1})\, d\lambda_r \cdots d\lambda_1\right| dz.$$

(Here, we assume that $t_i > \Delta_n$. This is without loss of generality since otherwise $I_1 = 0$ and I_2 is over $z \in [0, t_{r+i} + \Delta_n/2]$.)

To estimate I_1 observe that the mean value theorem yields

$$\int_0^1\int_0^{\lambda_1}\cdots\int_0^{\lambda_{r-1}}((y_1-z)^{\alpha-1}-(y_2-z)^{\alpha-1})\,d\lambda_r\cdots d\lambda_1$$
$$=\frac{(x_1-z)^{\alpha-1}-(x_2-z)^{\alpha-1}}{r!}$$

for $x_j=x_j(z)\in[t_i-\Delta_n/2,\ t_{r+i}+\Delta_n/2]$. Note also that

$$(x_1-z)^{\alpha-1}-(x_2-z)^{\alpha-1}=(x_1-x_2)(\alpha-1)(t_i-\Delta_n/2+\delta-z)^{\alpha-2}$$

for some $\delta\geqslant 0$. Take now q defined by $1/q+1/p=1$. Then Hölder's inequality and the fact that $(\alpha-2)\,q+1<0$ imply

$$I_1\leqslant c_1\,\Delta_n\,\|T_{r,\alpha}(f)\|_p\left[\int_0^{t_i-\Delta_n}(t_i-z-\Delta_n/2)^{(\alpha-2)q}\,dz\right]^{1/q}$$
$$=c_2\,\Delta_n^{\alpha-1+1/q}=c_2\,\Delta_n^{\alpha-1/p}. \tag{2.3}$$

Here c_1 and c_2 are positive constants independent of n.

To estimate I_2, observe that

$$\int_0^{\lambda_{r-1}}(y_1-z)_+^{\alpha-1}\,d\lambda_r\leqslant\frac{(t_{r+i}+\Delta_n/2-z)^\alpha}{|x-t_{r+i}|\,(1-\alpha)}$$

and

$$\int_0^{\lambda_{r-1}}(y_2-z)_+^{\alpha-1}\,d\lambda_r\leqslant\frac{(t_{r+i}+\Delta_n/2-z)^\alpha}{|x-t_{r+i-1}|\,(1-\alpha)}.$$

Hence Hölder's inequality yields

$$I_2\leqslant c_3\,\|T_{r,\alpha}(f)\|_p\,\Delta_n^{\alpha+1/q}\left[\frac{1}{|x-t_{r+i-1}|}+\frac{1}{|x-t_{r+i}|}\right]$$

for a positive constant c_3. This, (2.3), and (2.1) imply that

$$|e(x)|\leqslant c\,\Delta_n^{r+\alpha-1/p},$$

as claimed.

To complete the proof we now consider $s>r$. As before, given x, let p_i interpolate f at $t_i, ..., t_{r+s}$. Then for $x\notin\{t_i, ..., t_{i+s}\}$ the error at x equals

$$e(x)=f(x)-p_i(x)=a(x)\prod_{j=i}^{s+i}(x-t_j)$$

with $a(x)$ being a divided difference of degree $s+1$, $a(x)=f[t_i, ..., t_{s+i}, x]$.

Using the recursive definition for divided differences, we can represent $a(x)$ as a combination of divided differences of degree $r+1$. For those divided differences, we have the same estimate as for $a(x)$ with $s=r$. Since the coefficients in this combination are proportional to $\Omega(\delta_n^{r-s})$, we get that

$$|e(x)| \leqslant c\,\Delta_n^{s+\alpha-1/p}\,\delta_n^{r-s}.$$

This completes the proof. ∎

Remark 1. For equally spaced points, $\Delta_n = \delta_n = 1/n$ and then

$$e(A_{s,n}, N_n; r, \alpha, p) \leqslant c_{s,r,\alpha,p}\, n^{-(r+\alpha-1/p)}$$

whenever $s \geqslant r$. Hence, taking s larger than r does not essentially increase the worst case error (obviously, the constant $c_{s,r,\alpha,p}$ does depend on s and it grows with $s-r$). On the other hand, if $s<r$ (or $s<r-1$ for $\alpha=0$), the worst case error of P_s for the class $W_p^{r,\alpha}$ is unbounded. Hence, taking too small s leads to less efficient algorithms.

3. Minimal Error

We estimate the minimal worst case error of approximating $f \in W_p^{r,\alpha}$. That is, we estimate

$$e(n; r, \alpha, p) := \inf_N \inf_A \sup_{f \in W_p^{r,\alpha}} \frac{\|f - A(N(f))\|_{\sup}}{\|T_{r,\alpha}(f)\|_p},$$

where the first infimum is taken with respect to arbitrary *information* operator N of cardinality n,

$$N(f) = [f^{(i_1)}(t_1), ..., f^{(i_n)}(t_n)],$$

and the second infimum with respect to arbitrary *algorithm* A, i.e., arbitrary mapping from $\mathfrak{R}^n$ to $B[0,1]$ (the space of bounded functions).

Theorem 2.

$$e(n; r, \alpha, p) = \begin{cases} \Theta(n^{-(r+\alpha-1/p)}) & \textit{if} \quad r+\alpha-1/p \geqslant 0 \\ +\infty & \textit{otherwise}. \end{cases}$$

Proof. Due to Theorem 1, we need only to prove that for any $N(f) = [f^{(i_1)}(t_1), ..., f^{(i_n)}(t_n)]$ and any A, the error

$$e(A, N; r, \alpha, p) = \sup_{f \in W_p^{r,\alpha}} \frac{\|f - A(N(f))\|_{\sup}}{\|T_{r,\alpha}(f)\|_p}$$

is bounded by

$$e(A, N; r, \alpha, p) = \begin{cases} \Omega(n^{-(r+\alpha-1/p)}) & \text{if} \quad r+\alpha-1/p \geqslant 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{3.1}$$

As a matter of fact, we prove a stronger result by showing that (3.1) holds even if N consists of the following $n(r+1)$ evaluations

$$N(f) = [f(t_1), ..., f^{(r)}(t_1), ..., f(t_n), ..., f^{(r)}(t_n)].$$

Since $T_{r,0}(f)(y) = f^{(r)}(y) - f^{(r)}(0)$, (3.1) with $\alpha = 0$ follows from the well-known results, see, e.g., [6]. Hence we can assume that $\alpha > 0$.

We begin with $r+\alpha-1/p \geqslant 0$. Let $0 = t_0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant t_{n+1} = 1$ be given. Let i be an index such that $h := t_{i+1} - t_i \geqslant 1/(2n)$ and $1 - t_{i+1} \geqslant 1/2$. Let f be a function defined by

$$f(x) = (x-t_i)_+^{r+1} (t_{i+1}-x)_+^{r+1}.$$

Then $\pm f \in W_p^{r,\alpha}$ and $N(\pm f) = [0, ..., 0]$. Hence

$$e(A, N; r, \alpha, p) \geqslant \max\left\{ \frac{\|f - A(0)\|_{\sup}}{\|T_{r,\alpha}(f)\|_p}, \frac{\|-f - A(0)\|_{\sup}}{\|T_{r,\alpha}(f)\|_p} \right\} \geqslant \frac{\|f\|_{\sup}}{\|T_{r,\alpha}(f)\|_p}$$

for every algorithm A. Since $\|f\|_{\sup} = (h/2)^{2r+2}$, to prove (3.1) we need only to show that

$$\|T_{r,\alpha}(f)\|_p \leqslant c h^{r+2-\alpha+1/p} \tag{3.2}$$

for some positive constant c independent of n. (As in Section 2, we shall use $c_1, c_2, ..., c_6$ to denote positive constants independent of n.)

Note that $T_{r,\alpha}(f)(y) = 0$ for $y \leqslant t_i$, and $T_{r,\alpha}(f)(y) = \int_{t_i}^{t_{i+1}} (y-t)_+^{-\alpha} f^{(r+1)}(t)\, dt$ for $y > t_i$. Since $|f^{(r+1)}(t)| \leqslant c_1 h^{r+1}$, we thus have

$$\begin{aligned} \|T_{r,\alpha}(f)\|_p^p &\leqslant c_1^p h^{(r+1)p} \int_{t_i}^{t_{i+1}+h} \left[\int_{t_i}^{t_{i+1}} (y-t)_+^{-\alpha}\, dt \right]^p dy + I \\ &= \left(\frac{c_1}{1-\alpha} \right)^p h^{(r+1)p} \int_{t_i}^{t_{i+1}+h} [(y-t_i)^{1-\alpha} - (y-t_{i+1})_+^{1-\alpha}]^p\, dy + I \\ &\leqslant c_2 h^{(r+1)p} \int_{t_i}^{t_{i+1}+h} (y-t_i)^{(1-\alpha)p}\, dy + I \\ &\leqslant c_3 h^{(r+2)p} h^{1-\alpha p} + I, \end{aligned}$$

where

$$I = \int_{t_{i+1}+h}^{1} \left| \int_{t_i}^{t_{i+1}} (y-t)^{-\alpha} f^{(r+1)}(t)\, dt \right|^p dy.$$

Integration by parts and the fact that $f^{(r)}$ vanishes at t_i and t_{i+1} yield

$$I=\int_{t_{i+1}+h}^{1}\left|\alpha\int_{t_i}^{t_{i+1}}(y-t)^{-\alpha-1}f^{(r)}(t)\,dt\right|^p dy.$$

Since $|f^{(r)}(t)|\leqslant c_4h^{r+2}$, we get

$$I\leqslant c_5h^{(r+2)p}\int_{t_{i+1}+h}^{1}[(y-t_i)^{-\alpha}-(y-t_{i+1})^{-\alpha}]^p\,dy.$$

Since $(y-t_i)^{-\alpha}-(y-t_{i+1})^{-\alpha}=-\alpha h(y-t_{i+1}+\delta)^{-\alpha-1}$ for every $y\geqslant t_{i+1}+h$ with some $\delta\geqslant 0$,

$$I\leqslant c_6h^{(r+2)p}h^p[h^{1-(1+\alpha)p}-(1-t_{i+1})^{1-(1+\alpha)p}]\leqslant c_6h^{(r+2-\alpha+1/p)p}.$$

Thus,

$$\|T_{r,\alpha}(f)\|_p\leqslant ch^{r+2-\alpha+1/p},$$

as claimed in the first part of (3.2). This completes the proof of (3.1) for $r+\alpha-1/p\geqslant 0$.

For $r+\alpha-1/p<0$, i.e., $r=0$ and $\alpha<1/p$, take a suitably large integer m, and define $f_m(x)=x(1/m-x)_+$. Again, $N(f_m)=0$ and $\|f_m\|_{\sup}=(2m)^{-2}$. Then (3.2) yields $\|T_{0,\alpha}(f_m)\|_p\leqslant cm^{-2+\alpha-1/p}$ and therefore

$$\frac{\|f_m\|_{\sup}}{\|T_{0,\alpha}(f_m)\|_p}\geqslant\frac{1}{4c}m^{1/p-\alpha}\to\infty.$$

This completes the proof of (3.1). ∎

4. Concluding Remarks

In this section we briefly discuss some generalizations and applications of the reported results.

4.1. *Adaptation*

It follows from general results, see [1, 2, 7], that Theorem 2 is valid for a more general class of information that includes adaptive information, that is, information consisting of function and/or derivative values taken at adaptively selected points. Thus, function values at equally spaced points are optimal (modulo a constant) in the class of adaptive information.

4.2. *Randomization*

Theorem 2 combined with results of [4] and/or [8] implies optimality of a piecewise polynomial interpolation even among all random methods. Here by a random method we mean an arbitrary algorithm that uses function and/or derivative values of f at adaptively and randomly selected points; there is no restriction on employed randomization and in particular the randomization may depend on f via already computed function and/or derivative values. The error of such a method is defined as the worst case (with respect to f) expected (with respect to the randomization) error.

REFERENCES

1. N. S. BAKHVALOV, On approximate computation of integrals, *Vestnik Moscow Univ. Ser. Math. Mek. Astronom. Fiz. Chem.* **4** (1959), 3–18. [In Russian]
2. S. GAL AND C. A. MICCHELLI, Optimal sequential and non-sequential procedures for evaluating a functional, *Appl. Anal.* **10** (1980), 105–120.
3. E. NOVAK, "Deterministic, Stochastic, and Average Error Bounds in Optimal Recovery," *in* "Lecture Notes in Mathematics," Vol. 1349, Springer-Verlag, New York/Berlin, 1988.
4. E. NOVAK, "Stochastic Properties of Quadrature Formulas," Dept. Math. Report, Erlangen University, 1987.
5. A. PINKUS, "n-Widths in Approximation Theory," Springer-Verlag, Berlin, 1985.
6. V. M. TIKHOMIROV, Best methods of approximation and interpolation of differentiable functions in the space $C[0, 1]$, *Math. USSR-Sb.* **9**, No. 2 (1969), 275–289.
7. J. F. TRAUB AND H. WOŹNIAKOWSKI, "A General Theory of Optimal Algorithms," Academic Press, New York, 1980.
8. G. W. WASILKOWSKI, Randomization for continuous problems, *J. Complexity* **5** (1989), 195–218.