

ON THE SAUER-XU FORMULA FOR THE ERROR IN  
 MULTIVARIATE POLYNOMIAL INTERPOLATION

CARL DE BOOR

ABSTRACT. Use of a new notion of multivariate divided difference leads to a short proof of a formula by Sauer and Xu for the error in interpolation, by polynomials of total degree  $\leq n$  in  $d$  variables, at a ‘correct’ point set.

It is the purpose of this note to give a short proof of a remarkable formula for the error in polynomial interpolation given in [3].

In [3], Sauer and Xu consider interpolation from the space  $\Pi_n(\mathbf{R}^d)$  of  $d$ -variate polynomials of degree  $\leq n$  at a point set  $\mathcal{X}$  which is **correct** for it in the sense that, for an arbitrary  $g$ , there is exactly one  $p \in \Pi_n(\mathbf{R}^d)$ , denoted here by

$$P_n g,$$

which agrees with  $g$  on  $\mathcal{X}$ . Such a correct point set can, as Sauer and Xu point out, be partitioned into subsets

$$x^{(i)} := \{x_r^{(i)} : r = 1, \dots, r_i^d := \dim \Pi_i - \dim \Pi_{i-1} = \binom{i-1+d}{i}\}, \quad i = 0, 1, \dots, n,$$

in such a way that, for each  $j \leq n$ , polynomial interpolation from  $\Pi_j$  at the points in

$$x^{(\leq j)} := x^{(0)} \cup \dots \cup x^{(j)}$$

is uniquely possible. They denote the corresponding Lagrange polynomial in  $\Pi_j$  associated with the point  $x_r^{(j)}$  by

$$p_r^{[j]};$$

i.e.,  $p_r^{[j]}$  is the unique element of  $\Pi_j$  which satisfies

$$p_r^{[j]}(x_s^{(i)}) = \delta_{ji} \delta_{rs}, \quad s = 1, \dots, r_i^d; \quad i = 0, \dots, j,$$

but is called a Newton polynomial in [3], probably because any  $p \in \Pi_j$  which vanishes on  $x^{(<j)}$  can be written in the form

$$(1) \quad p = \sum_{r=1}^{r_j^d} p_r^{[j]} p(x_r^{(j)}).$$

---

Received by the editor May 8, 1995.

1991 *Mathematics Subject Classification*. Primary 41A05, 41A10, 65D05.

*Key words and phrases*. Polynomials, multivariate, interpolation, error, remainder formula, divided difference.

This work was supported by the NSF grant DMS-9224748, by the US-Israel Binational Science Foundation under Grant No. 90-00220, and by ARO under Grant No. DAA H04-95-1-0089.

Let  $E_j g$  denote the error in the polynomial interpolant from  $\Pi_j$  to  $g$ , i.e.,

$$g = P_j g + E_j g.$$

Sauer and Xu prove that (in roughly the same notation, except for a slight reordering and the use of a divided difference here)

$$(2) \quad (E_j g)(x) = \sum_{r=1}^{r_j^d} p_r^{[j]}(x) \sum_{\mu \in \Lambda_r^{(j)}} c_\mu [x_\mu, x; \Delta x_\mu, x - x_r^{(j)}] g,$$

with the following definition of the various quantities appearing here:

$$\Lambda_r^{(j)} := \{ \mu \in \times_{i=0}^j \{1, \dots, r_i^d\} : \mu_j = r \};$$

$$c_\mu := \prod_{i=0}^{j-1} p_{\mu_i}^{[i]}(x_{\mu_{i+1}}^{(i+1)});$$

$$x_\mu := (x_{\mu_i}^{(i)} : i = 0, \dots, j), \quad \Delta x_\mu := (x_{\mu_{i+1}}^{(i+1)} - x_{\mu_i}^{(i)} : i = 0, \dots, j - 1);$$

and, finally,  $[t_0, \dots, t_j; \xi_1, \dots, \xi_j]$  is the  $j$ th **divided difference** introduced in [1], i.e.,

$$(3) \quad [t_0, \dots, t_j; \xi_1, \dots, \xi_j] g := \int_{[t_0, \dots, t_j]} D_{\xi_1} \cdots D_{\xi_j} g,$$

with

$$f \mapsto \int_{[t_0, \dots, t_j]} f := \int_0^1 \int_0^{s_1} \cdots \int_0^{s_{j-1}} f(t_0 + s_1 \nabla t_1 + \cdots + s_j \nabla t_j) ds_j \cdots ds_1$$

the linear functional which was used to such advantage by Micchelli in his analysis of Kergin interpolation and the simplex spline, and which has been dubbed by him in [2] the **divided difference functional for  $\mathbf{R}^d$** . The only facts about the  $j$ th divided difference (3) needed here are that it is symmetric in the *points*  $t_0, \dots, t_j$ , and is symmetric and *linear* in the *directions*  $\xi_1, \dots, \xi_j$  (which is obvious), and that, for an arbitrary point sequence  $T$ , points  $a, b$ , and arbitrary direction sequence  $\Xi$  (with  $\#T = \#\Xi$ ),

$$(4) \quad [T, a, b; \Xi, a - b] = [T, a; \Xi] - [T, b; \Xi],$$

which can be verified directly (see [1]).

The short proof of (2) about to be given here is by induction (as is the proof in [3]). For  $j = 0$ , (2) is the special case  $T = () = \Xi, a = x, b = x_1^{(0)}$  of (4). Assuming (2) to hold for  $j = n$ , we observe that

$$(E_n g)(x) = G_n(x, x),$$

with

$$G_n(x, y) := \sum_{r=1}^{r_n^d} p_r^{[n]}(x) \sum_{\mu \in \Lambda_r^{(n)}} c_\mu [x_\mu, y; \Delta x_\mu, x - x_r^{(n)}]g.$$

Now note that, for arbitrary  $y$ ,  $G_n(\cdot, y)$  is a polynomial of degree  $\leq n + 1$  (since the  $p_r^{[n]}$  are in  $\Pi_n$  while  $x \mapsto [T; \Xi, x]g$  is a linear (scalar-valued) function), and  $G_n(\cdot, y)$  vanishes at every  $x_r^{(n)}$ , hence, as in (1),  $G_n(\cdot, y)$  is writeable as

$$G_n(\cdot, y) = \sum_{s=1}^{r_{n+1}^d} p_s^{[n+1]} G_n(x_s^{(n+1)}, y).$$

On the other hand, since  $g = P_n g + E_n g$ , the function

$$P_n g + \sum_{s=1}^{r_{n+1}^d} p_s^{[n+1]} (E_n g)(x_s^{(n+1)})$$

agrees with  $g$  at  $x^{(\leq n+1)}$  and is in  $\Pi_{n+1}$ , hence must equal  $P_{n+1}g$ . Therefore,

$$\begin{aligned} E_{n+1}g &= E_n g - \sum_s p_s^{[n+1]} (E_n g)(x_s^{(n+1)}) \\ &= \sum_s p_s^{[n+1]} (G_n(x_s^{(n+1)}, \cdot) - G_n(x_s^{(n+1)}, x_s^{(n+1)})), \end{aligned}$$

and this implies (2) for  $j = n + 1$  since

$$G_n(x_s^{(n+1)}, \cdot) - G_n(x_s^{(n+1)}, x_s^{(n+1)}) = \sum_{r=1}^{r_n^d} p_r^{[n]}(x_s^{(n+1)}) \sum_{\mu \in \Lambda_r^{(n)}} c_\mu d_{\mu,r}g$$

with

$$\begin{aligned} d_{\mu,r} &:= [x_\mu, \cdot; \Delta x_\mu, x_s^{(n+1)} - x_r^{(n)}] - [x_\mu, x_s^{(n+1)}; \Delta x_\mu, x_s^{(n+1)} - x_r^{(n)}] \\ &= [x_\mu, x_s^{(n+1)}, \cdot; \Delta x_\mu, x_s^{(n+1)} - x_r^{(n)}, \cdot - x_s^{(n+1)}], \end{aligned}$$

by (4). □

It follows that

$$P_n g = \sum_{j=0}^n \sum_{r=1}^{r_j^d} p_r^{[j]} (E_{j-1}g)(x_r^{(j)}),$$

with

$$E_{-1}g := g.$$

Since  $(p_r^{[j]} : r = 1, \dots, r_j^d; j = 0, 1, \dots)$  is linearly independent, it follows that, for any  $j$ ,  $[x;]E_j$  is a linear combination of the linear functionals  $[x;]$  and  $[x_r^{(i)};]$ ,  $r = 1, \dots, r_i^d; i = 0, \dots, j$ . However, its constituents, i.e., the  $j$ th divided differences

$$[x_\mu, x; \Delta x_\mu, x - x_{\mu_j}^{(j)}]$$

by themselves, are not necessarily such linear combinations, as the simple example

$$[0, \mathbf{i}_1, \mathbf{i}_1 + \mathbf{i}_2; \mathbf{i}_1, \mathbf{i}_2],$$

with  $\mathbf{i}_j$  the  $j$ th unit vector, readily shows.

## REFERENCES

1. Carl de Boor, *A multivariate divided difference*, Approximation Theory VIII, Academic Press, New York, 1995, pp. 87–96.
2. C. A. Micchelli, *On a numerically efficient method for computing multivariate B-splines*, Multivariate Approximation Theory, Birkhäuser, Basel, 1979, pp. 211–248. MR **81g**:65017
3. T. Sauer and Yuan Xu, *On multivariate Lagrange interpolation*, Math. Comp. **64** (1995), 1147–1170. MR **95j**:41051

COMPUTER SCIENCES DEPARTMENT, UNIVERSITY OF WISCONSIN-MADISON, 1210 W. DAYTON ST., MADISON, WISCONSIN 53706

*E-mail address:* `deboor@cs.wisc.edu`