Lagrange Interpolation on a Lattice: Bounding Derivatives by Divided Differences.

Thomas Kunkle*
Department of Mathematics
University of Wisconsin-Madison
Madison, WI 53706

Abstract

Fix an integer $n > 0$. For a multivariate function defined on a (not necessarily rectangular) lattice, an extension is constructed to have, $\forall k \leq n$, derivatives of total degree k that are bounded by the function's tensor product divided differences of total degree k times a constant independent of the lattice and the function. The extension is locally constructed, can have any prescribed smoothness, and reproduces polynomials of degree $< n$ in each variable.

1. Introduction

We will extend this univariate result of Favard:

Let $M = (m_i)_{-\infty}^{\infty}$ be a strictly increasing sequence in $\mathbf{R}$ having no limit points and n a fixed positive integer. Suppose $f : M \to \mathbf{R}$ is given. By an **extension**, or **interpolant**, of f, we mean a $g \in C^n(\mathbf{R})$ whose **restriction** :=

$$g|_M : M \to \mathbf{R} : m \mapsto g(m)$$

to M is f. A consequence of Rolle's theorem is that the nth derivative of any extension of f must take on the values

$$(n![m_i, \ldots, m_{i+n}]f) \quad , \tag{1.1}$$

where $[m_i, \ldots, m_{i+n}]f$ is an nth degree divided difference of f; locally, then, any extension's nth derivative necessarily has max norm $\geq$ the absolute maximum of (1.1). In [2] and [3], Favard constructs an extension Ef of f whose nth derivative is **not much larger than necessary**, meaning

$$\left|(Ef)^{(n)}(x)\right| \leq C \max\bigl\{|[m_i, \ldots, m_{i+n}]f| : m_i \leq x \leq m_{i+n}\bigr\} \quad , \tag{1.2}$$

with C a constant independent of f and M.

Now suppose that f is a function given on a **lattice** $M \subset \mathbf{R}^d$ (i.e., the image of $\mathbf{Z}^d$ under an invertible $A \in \mathbf{R}^{d\times d}$) and that we extend f via the tensor product of Favard's scheme (i.e., we use Favard's scheme to extend f in each of the directions of the columns of A). On the surface, it would appear that the resulting extension would simply have its $\mathbf{n}$th derivative in the directions of A's columns bounded by the data's $\mathbf{n}$th divided differences in these directions. By $\mathbf{n}$th, we mean of degree n in each direction. However, much more

* This material is based upon work supported in part by the National Science Foundation under Grant No. DMS-8701275

is true: each of the extension's nth total degree derivatives in the standard directions can be bounded by nth total degree divided differences in suitably chosen directions times a constant that depends only on n and d. The precise result is found in the last section of this paper.

Favard's univariate result had been applied in [5] to error analysis of numerical ordinary differential equations. With a multivariate generalization, one should be able to formulate similar applications to numerical pde.

We denote Favard's univariate extension operator by E; we will refer to its d-fold tensor product by F. An outline of their discussion is as follows:

We introduce notation in section 2 and some important identities in section 3. We prove our main result by considering successive increasingly general cases. In section 4 we deal with the special case that $M = \mathbf{Z}^d$. In section 5, we discuss the case $M = A\mathbf{Z}^d$ where A is diagonal. Finally, in section 6, we remove all restrictions on A other than its invertibility. We state there our most general result. Section 7 contains some concluding remarks.

Some important properties of Ff are that it

- depends locally on f,
- can be made smooth to any prescribed order,
- reproduces a certain subspace of polynomials, and most important,
- has, $\forall k \le n$, kth derivatives that are no more than a constant (independent of M) times some kth differences of f.

2. Notation

We denote intervals on the real line by $[a \,.\,.\, b]$, with usual conventions for open and closed intervals. We refer to the ith coordinate of $x \in \mathbf{R}^d$ as x_i. For x, y in $\mathbf{R}^d$, $x \le y$ means that $\forall i$, $x_i \le y_i$.

Elements of $\mathbf{Z}^d$ are called **multiintegers**, or **integers** when the context makes this unambiguous. Elements of $\mathbf{Z}^d_+ := \{\alpha \in \mathbf{Z}^d : \alpha \ge 0\}$ are called **multiindices**, usually denoted α, β, or γ. Define $|\alpha| := \sum \alpha_i$, $\alpha! := \prod \alpha_i!$, and the monomials by

$$()^\alpha : \mathbf{R}^d \to \mathbf{R} : x \mapsto x^\alpha := \prod x_i^{\alpha_i} \quad .$$

Letting $\mathbf{n}$ stand for the multiindex with $\mathbf{n}_i = n$ for all i, the space of polynomials of degree $< n$ in each variable, is denoted

$$\Pi_{<\mathbf{n}} := \Pi_{<\mathbf{n}}(\mathbf{R}^d) := \operatorname{span}\{()^\alpha : \alpha < \mathbf{n}\} \quad .$$

We refer to the jth column of a matrix A as A_j and its typical element as A_{ij}. The jth standard unit vector (i.e., the jth column of the identity matrix) is therefore I_j.

D stands for the gradient. We find it convenient to think of it as a (column) vector in $\mathbf{R}^d$, so that D_i is the derivative in the direction I_i, and the gradient of derivatives in the directions of the columns of A is $D_A := A^T D$. The monomial differential operators are written

$$D^\alpha := \prod D_i^{\alpha_i} \quad .$$

We say that D^α and $()^\alpha$ are of **degree** α, of degree α_i in the ith variable, and of **total degree** $|\alpha|$.

We make frequent use of translations and linear changes of variable; to facilitate this, define for $x \in \mathbf{R}^d$ and for $A \in \mathbf{R}^{d\times d}$ the operators

$$T_x : f \mapsto f(\cdot + x)$$

and

$$S_A : f \mapsto f(A\cdot) \quad .$$

By K or K' we will always mean a compact set in $\mathbf{R}^d$. By $\|\cdot\|_K$ we denote the max norm on K, or **K norm**. If the domain of f does not include all of K, then $\|f\|_K$ is interpreted to mean $\|f\|_{K\cap \operatorname{dom} f}$. If $\{f^j : j \in J\}$ denotes a finite collection of functions, then define

$$\left\|f^j\right\|_K^J := \max_J \left\|f^j\right\|_K \quad .$$

When writing inequalities, we will not distinguish between different constants that are independent of M and f, referring instead to every such constant as C.

For β a multiindex, we denote by $\diamondsuit^\beta$ a βth**-degree tensor product divided difference** having integer support:

$$\diamondsuit^\beta := [0,1,\cdots,\beta_1]\otimes[0,1,\cdots,\beta_2]\otimes\cdots\otimes[0,1,\cdots,\beta_d] \quad .$$

Other divided differences with integer support are obtained by translation:

$$\diamondsuit_z^\beta := \diamondsuit^\beta T_z \quad .$$

In what follows, “divided difference” will sometimes be shortened to “**difference**”. At no time will we mean by this the undivided, “finite differences” used often in approximation theory.

By N^k we denote the (univariate) **Newton polynomial**

$$N^k(x) := x\big(x-1\big)\big(x-2\big)\cdots\big(x-k+1\big) \quad ,$$

and by N_z^k its translate $T_{-z}N^k$. Thus, in our notation, the Newton form of the polynomial of degree $< n$ that agrees with a function g at $\{0,1,\ldots,n-1\}$ is

$$Pg := \sum_0^{n-1} N^k \diamondsuit^k \quad .$$

To interpolate at $\{z,\ldots,z+n-1\}$, one would use the operator $P_z := T_{-z}PT_z$.

3. A Divided Difference Identity

We establish an identity concerning these divided differences. To begin, note that differences compound; that is, if we apply a γth divided difference to the function

$$f^\beta : z \mapsto \diamondsuit_z^\beta f \quad , \tag{3.1}$$

the result is a multiple of a $(\gamma+\beta)$th difference of f. In particular,

$$\Diamond_u^\gamma f^\beta = \binom{\gamma+\beta}{\gamma} \Diamond_u^{\gamma+\beta} f \quad , \tag{3.2}$$

as can be proven by induction. (Here the multinomial coefficient is given by $(\gamma+\beta)!/\gamma!\beta!$.)

Because $\Diamond_u^\gamma$ being a linear combination of point evaluations with coefficients independent of the integer u, there exists C independent of $\gamma \le \mathbf{n}$ and u such that

$$|\Diamond_u^\gamma f| \le C \max\{|f(x)| : x \in \operatorname{supp} \Diamond_u^\gamma\} \quad .$$

We combine this with (3.2) and the finiteness of $K \cap \mathbf{Z}^d$ for every K to yield our

identity 3.3. *For every K there is a K' and a C independent of both $\alpha \le \beta \le \mathbf{n}$ and K so that*

$$\|f^\beta\|_K \le C \|f^\alpha\|_{K'} \quad .$$

4. The Case $M = \mathbf{Z}^d$

In this section we give a construction of Favard's extension Ef in the special case that $M = \mathbf{Z}^d$. It will be assumed throughout that f is a function with this domain.

Let $\psi \in C_c^n(\mathbf{R})$ have support $[-1\,..\,1]$ and satisfy $\sum_{\mathbf{Z}} T_z\psi = 1$. We sometimes call $\psi_z := T_{-z}\psi$ a **weight function**.

We define the extension operator E as a weighted average of the local interpolation operators (P_z):

$$E := \sum_{\mathbf{Z}} \psi_z P_z \quad . \tag{4.1}$$

To interpolate to data f given on $\mathbf{Z}^d$, we use the operator

$$F := \overbrace{E \otimes \cdots \otimes E}^{d} \quad ;$$

Like E, F is the weighted average of local polynomial projectors. To write it as such, we set Ψ equal to the tensor product of d copies of ψ and R equal to the tensor product of the same number of operators P. For z a multiinteger, $\Psi_z := T_{-z}\Psi$, and $R_z := T_{-z}RT_z$, so that R_z is the interpolation projector whose range is $\Pi_{<\mathbf{n}}$ and whose interpolation conditions are point evaluations at $\{z+u : 0 \le u < \mathbf{n}\}$. Then

$$F = \sum_{\mathbf{Z}^{\mathrm{d}}} \Psi_z R_z \quad . \tag{4.2}$$

From (4.2) we can see several properties of F. First, the restriction of F to $\Pi_{<\mathbf{n}}$ is the identity, since the same can be said of each R_z. Second, Ff is as smooth as Ψ, at least in C^n. Third, Ff depends locally on f, since the coefficient of Ψ_z depends only on f's values near z and Ψ has compact support.

Most important for our purposes, we have this

theorem 4.3. *Depending only on our choice of ψ, there is a constant C such that for every $\alpha \le \mathbf{n}$, if $\|f^\alpha\|_\infty < \infty$, then*

$$\|D^\alpha Ff\|_\infty \le C\,\|f^\alpha\|_\infty \quad . \tag{4.3.a}$$

If f^α is not bounded, then it still holds that $\forall K\ \exists K'$ such that

$$\|D^\alpha Ff\|_K \le C\,\|f^\alpha\|_{K'} \quad . \tag{4.3.b}$$

Note that while the computation of Ff uses divided differences of f of total degree $> n$, these differences do not affect the interpolant's nth derivatives.

To begin the proof of (4.3), note that since Ff depends locally on f, its restriction to K depends only on finitely many values of f. Therefore (4.3.a) implies (4.3.b) for some K'. Without loss of generality we can assume that $\|f^\beta\|_\infty < \infty$ for every $|\beta| \le n$.

It will suffice to prove $\exists C \forall z \in \mathbf{Z}^d$

$$\|D^\alpha Ff\|_{z+[0..1]^d} \le C\,\|f^\alpha\|_\infty \quad . \tag{4.4}$$

We show first the existence of such a C for $z = 0$.

For purposes of evaluating E on $[0\,..\,1]$ it can be written

$$E = P + \psi_1(P_1 - P) = \sum_{l\le n} N^l \diamondsuit^l + \psi_1 N_1^{n-1} n \diamondsuit^n \quad . \tag{4.5}$$

Since F is the tensor product of d copies of E, (4.5) allows us to write Ff on $[0\,..\,1]^d$ as a sum of divided differences of f times continuous functions:

$$Ff = \sum_{z,\beta} Q_z^\beta \diamondsuit_z^\beta f \quad .$$

Here z ranges over $\{0,1\}^d$, β over multiindices $\le \mathbf{n}$, and Q_z^β is the product of Newton polynomials and weight functions. When computing the sum's αth derivative, we need sum only over $\beta \ge \alpha$, for if $\beta_i < \alpha_i \le n$, then in the ith variable Q_z^β is a polynomial of degree β_i and therefore has αth derivative equal zero. With this restriction, we see that Ff is a linear combination of smooth functions with coefficients of the form

$$\{\diamondsuit_z^\beta f : \beta \ge \alpha\} \quad .$$

Thus in the $[0\,..\,1]^d$ norm,

$$\begin{aligned} \|D^\alpha Ff\| &\le \left\|f^\beta\right\|_\infty^{\alpha\le\beta\le\mathbf{n}} \sum_{z,\beta} \|D^\alpha Q_z^\beta\| \\ &\le C\,\|f^\alpha\|_\infty \quad . \end{aligned} \tag{4.6}$$

(The last inequality is by (3.3).) Thus (4.4) holds for $z = 0$.

For z other than 0, note that $F = T_{-z} F T_z$, and that

$$\|D^\alpha F f\|_{z+[0..1]^{\mathsf{d}}} = \|D^\alpha F T_z f\|_{[0..1]^{\mathsf{d}}} \leq C \,\|(T_z f)^\alpha\|_\infty = C\,\|f^\alpha\|_\infty \quad , \tag{4.7}$$

and for general z, (4.4) holds with the same constant as when $z = 0$. This completes the proof of (4.3).

In (4.6) appears the influence of ψ on (4.3)'s constant. One reduces that constant by choosing ψ so as to reduce the derivatives of products Q_z^β of Newton polynomials and weight functions. [1] discusses this choice for the univariate extension (in the original terms of Favard's construction).

5. The Case $M = \mathrm{diag}(a)\mathbf{Z}^d$

We now examine the more general case that the interpolation points M form the lattice $\mathrm{diag}(a)\mathbf{Z}^d$ for $a \in \mathbf{R}^d$ positive. Throughout this section it will be understood that f is given on this M.

For shorthand, we let

$$S_a := S_{\mathrm{diag}(a)} \quad .$$

We obtain an interpolation scheme for this M by scaling. Specifically, define

$$F^a := S_a^{-1} F S_a \quad ,$$

and define the tensor product divided differences

$$\diamondsuit_{z,a}^\alpha := a^{-\alpha} \diamondsuit_z^\alpha S_a \quad .$$

We again use the notation f^β to denote a function on M:

$$f^\beta : \mathrm{diag}(a)z \mapsto \diamondsuit_{z,a}^\beta f \quad . \tag{5.1}$$

(5.1) is not meant to overwrite the meaning we gave f^β in (3.1), a function on the integers. g^β will be understood to be a function with the same domain as g, whether $\mathbf{Z}^d$ or $\mathrm{diag}(a)\mathbf{Z}^d$, obtained by taking differences in the directions (I_j).

Theorem (4.3) has the following

corollary 5.2. *With C the same constant appearing in (4.3), for every multiindex $\alpha \leq \mathbf{n}$ and for every K we have a K' such that*

$$\|D^\alpha F^a f\|_K \leq C\,\|f^\alpha\|_{K'} \quad .$$

Its proof is trivial; again, it will suffice to assume that f^α is bounded. Then in the sup-norm,

$$\begin{aligned} &\|D^\alpha S_a^{-1} F S_a f\| \\ &= a^{-\alpha}\|D^\alpha F S_a f\| \\ &\leq a^{-\alpha}\|\big(S_a f\big)^\alpha\| \\ &= \|f^\alpha\| \end{aligned}$$

6. The Case $M = A\mathbf{Z}^d$

Finally we address the interpolation problem for $M = A\mathbf{Z}^d$ for an invertible $A \in \mathbf{R}^{d\times d}$. We will assume throughout this section that f has this domain.

To relate the general lattice to those on which we already have a solution, we let U be the matrix obtained from A by dividing each column by its (Euclidean) length, and let a satisfy

$$U \operatorname{diag}(a) = A \quad .$$

Then it is natural to define the interpolant

$$F^A := S_U^{-1} F^a S_U \quad .$$

F^A will have the property that nth derivatives in the directions U will by bounded by some constant times nth divided differences in those directions formed from f's values on $A\mathbf{Z}^d$. To be exact, using $D_A S_B = (BA)^T S_B D$ we have $\forall K \exists K'$

$$\|D_U^\alpha F^A f\|_K = \|D^\alpha F^a S_U f\|_{U^{-1}K} \le C\,\|(S_U f)^\alpha\|_{K'} \quad . \tag{6.1}$$

However, the derivatives D_U^α that appear in (6.1) depend on our mesh; it is possible to improve (6.1) by replacing the directional derivatives by the ordinary ones D^α.

Define the matrix norm

$$\|U\| := \max |U_{ij}| \quad .$$

From $D = (V^{-1})^T D_V$ it follows that for smooth g

$$D_j^k = \sum_{|\gamma|=k} \binom{k}{\gamma} (V^{-1})_j^\gamma D_V^\gamma$$

and

$$|D^\gamma g| \le C \|V^{-1}\|^{|\gamma|} \max_{|\beta|=|\gamma|} |D_V^\beta g| \quad , \tag{6.2}$$

so that to bound D^α in terms of $\{D_U^\beta\}$, with a constant independent of the lattice, we need a bound (in terms of n and d only) on $\|U^{-1}\|$. While this is obviously impossible for arbitrary A, it *is* possible for arbitrary *lattices*; the lattice $A\mathbf{Z}^d$ can also be written $B\mathbf{Z}^d$ for many B other than A. (Any such a B is called a **basis** for the lattice.) A result from the geometry of numbers guarantees us for every A the existence of a basis $B = B(A)$ satisfying

$$\sup_A \left| \frac{\prod \|B_j\|_2}{\det B} \right| < \infty \quad . \tag{6.3}$$

(This quotient can be taken as a measure of the skewness of the columns of B.) One can construct B so that for $d \ge 5$, this supremum is less than

$$\left(\frac{2}{\sqrt{\pi}}\right)^d \Gamma\left(\frac{d+1}{2}\right) \left(\frac{5}{4}\right)^{\frac14 (d-3)(d-4)} .$$

See [4, p. 65] for this advanced result. It is much simpler to arrive at a higher upper bound; i.e., one can construct B so that for $d \geq 2$ the supremum in (6.3) is bounded by

$$2^{-d} \prod_{i=1}^{d} (1+3^i) \quad .$$

For a short proof of this, see [4, pp. 126–128]. A fact of practical interest is that one can program a computer to perform both of these constructions.

Choosing in this way a new basis B for the lattice $A\mathbf{Z}^d$, we reset $a_j := \|B_j\|_2$ and $U := B\,\mathrm{diag}(a)^{-1}$. By (6.3), $|\det U|^{-1}$ is bounded independent of A, and by Cramer's rule so is $\|U^{-1}\|$. Thus for some constant independent of A and for any K

$$\|D^\alpha g\|_K \leq C \left\| D_U^\beta g \right\|_K^{|\beta|=|\alpha|} .$$

Combining this with (6.1) we have this

theorem 6.4. *There is a constant C depending only on our choice of ψ such that for every α with $|\alpha| \leq n$ and every K there is a K' with*

$$\|D^\alpha F^B f\|_K \leq C \left\| (S_U f)^\beta \right\|_{K'}^{|\beta|=|\alpha|} \quad .$$

7. Concluding Remarks

If $\mathrm{supp}\,\psi \subseteq [-1\,..\,n]$, then (4.3), (5.2), and (6.4) would still hold (as long as $\sum_{\mathbf{Z}} \psi_z = 1$), albeit with a more complicated proof. In [1], de Boor takes advantage of this to reduce the constant appearing in (1.2).

Since most lattices of interest to numerical analysts are rectangular, corollary (5.2) may be as important as theorem (6.4). That in this setting one needs only f^α to bound $D^\alpha F^a f$ (and not other differences of total degree $|\alpha|$) and that one can do this for all $\alpha \leq \mathbf{n}$ simultaneously is probably the result of M's great regularity. Such strong results might be too much to ask in a generalization of Favard's result to other interpolation point sets.

It is not difficult to extend this scheme to the case that data are given only on the lattice points within a parallelogram with sides in the directions of the basis B. To extend the scheme to data given only on lattice points within other domains presents a problem.

There exist other generalizations of Favard's scheme of the form (4.2) for Lagrange data on lattice. Those investigated to date fail to provide a (1.2)-like bound without a constant that depends on the aspect ratio of the lattice.

We mention here without proof that $F^a g$ (and more generally $F^B g$) has nth order local accuracy when used to approximate a function g in C^n from $g|_M$. This is a direct consequence of the fact that $F = I$ on $\Pi_{<\mathbf{n}}$. For every $k < n$, the extension's derivatives of total degree k approximate the corresponding derivatives of g with local approximation order $n - k$.

8. References

1. C. de Boor, A Smooth and Local Interpolant with "Small" kth Derivative, *in* "Numerical Solutions of Boundary Value Problems for Ordinary Differential Equations" (A. K. Aziz, Ed.), pp. 177–197, Academic Press, New York, 1974.

2. C. de Boor, How Small Can One Make the Derivatives of an Interpolating Function?, *J. Approx. Theory* **13** (1975), 105–116.

3. J. Favard, Sur l'interpolation, *J. Math. Pures Appl.* **19** (1940), 281–300.

4. C. G. Lekkerkerker, "Geometry of Numbers," John Wiley and Sons, New York, 1969.

5. H. O. Kreiss, Difference Approximations for Singular Perturbation Problems, *in* "Numerical Solutions of Boundary Value Problems for Ordinary Differential Equations" (A. K. Aziz, Ed.), pp. 199–211, Academic Press, New York, 1974.