APPROXIMATION THEORY:
A volume dedicated to Borislav Bojanov
(D. K. Dimitrov, G. Nikolov, and R. Uluchev, Eds.)
additional information (to be provided by the publisher)

An Efficient Definition of the Divided Difference

Carl de Boor

Dedicated to Borislav Bojanov on the occasion of his sixtieth birthday

A novel definition of the divided difference is shown to lead most readily to the standard basic properties of the divided difference.

1. Introduction

The original, and still standard, definition of divided differences is via their recurrence relation and is, offhand, valid only for divided differences at distinct sites. Divided differences involving repeated sites are then introduced as limiting cases, with the continuity of the divided difference as a function of its sites a conclusion requiring some care.

Kowalewski [?] was apparently the first to define the divided difference from the start for arbitrary sites, including sites with multiplicities, as the leading coefficient of the corresponding interpolating polynomial, and this requires proof of the existence of the interpolating polynomial. To be sure, the definition of the divided difference as the ratio of Vandermonde determinants, though restricted to distinct sites, amounts to the same thing, and was used prior to Kowalewski's book; see, e.g., Hopf [?]. Of course, either way, one then has to derive the recurrence relations from this definition and still has to settle the continuity of the divided difference as a function of its sites.

The definition proposed in the present note goes a step further in the direction taken by Kowalewski, and makes the continuous and even smooth dependence of the divided difference on its sites obvious, yet provides the recurrence relations almost immediately as well and gives, from the start, a unified treatment of divided differences for any sequence of sites regardless of site multiplicities.

To be sure, the definition covers only the divided differences of a polynomial. However, it is one message of this note that the discussion of divided differences simplifies when one considers them only on polynomials, yet such restriction is no loss since their extension to other (sufficiently smooth) functions is immediate in view of the density of polynomials in more general function spaces, and given their ready representation as a weighted sum of values and, perhaps, derivatives at their sites.

2. The Definition

Let $\tau = (\tau_1, \tau_2, \ldots)$ be an arbitrary sequence in $\mathbb{F}$ ($= \mathbb{R}$ or $\mathbb{C}$), and let

$$w_{j,\tau} : \mathbb{F} \to \mathbb{F} : t \mapsto (t-\tau_1)\cdots(t-\tau_j), \qquad j = 0, 1, 2, \ldots,$$

be the corresponding **Newton polynomials**. Then, the map

$$W_\tau : \mathbb{F}_0^{\mathbb{N}} \to \Pi : a \mapsto \sum_{j=1}^{\infty} w_{j-1,\tau}\, a(j),$$

from scalar sequences with finitely many non-zero entries to polynomials, is linear; it is also 1-1 (since $\deg w_{j,\tau} = j$ for all j) and, therefore, is onto (since, for each n, the sequence $(w_{i,\tau} : i < n)$ is linearly independent and in the n-dimensional space $\Pi_{<n}$ of all polynomials of degree $< n$, hence a basis for it).

In other words, each $p \in \Pi$ has a unique **Newton form** with respect to the sequence τ of **centers**:

$$p \;=:\; \sum_{j=1}^{\infty} w_{j-1,\tau} \lambda_{j,\tau} p,$$

with the $\lambda_{j,\tau}$ linear functionals on Π thereby defined. We propose to call $\lambda_{j,\tau} p$ the **divided difference of** p **at the sites** $\tau_1, \ldots, \tau_j$ and, correspondingly, denote it

$$\Delta(\tau_{1:j}) \;:=\; \Delta(\tau_1, \ldots, \tau_j) \;:=\; \lambda_{j,\tau}, \tag{1}$$

although this language and the notation reflecting it requires some explanation.

For this, let

$$p \;=\; \underbrace{\sum_{j \le n} w_{j-1,\tau} \lambda_{j,\tau} p}_{=:\; p_n} \;+\; \underbrace{\sum_{j > n} w_{j-1,\tau} \lambda_{j,\tau} p}_{=:\; w_{n,\tau} q_n}, \tag{2}$$

with q_n a well-defined polynomial since $w_{n,\tau}$ evidently divides each $w_{j-1,\tau}$ for $j > n$. This shows that p_n is the polynomial of degree $< n$, necessarily unique, for which $p - p_n$ is divisible by $w_{n,\tau}$. It follows that $\lambda_{n,\tau} p$ does, indeed, depend only on $\tau_1, \ldots, \tau_n$ (and symmetrically so) and on p, thus justifying (1).

It also follows that p_n agrees with p at $\tau_1, \ldots, \tau_n$.

3. Simple Examples

The **Taylor series**

$$p \;=\; \sum_{j=0}^{\infty} (\cdot - z)^j D^j p(z)/j!$$

is evidently the Newton form for p with respect to the constant center sequence $\tau = (z, z, z, \ldots)$, hence

$$\Delta([z]^{n+1})p := \Delta(\underbrace{z, \ldots, z}_{n+1\text{ terms}})p \;=\; D^n p(z)/n!, \quad j = 1, 2, \ldots. \tag{3}$$

If $\tau_1 \neq \tau_2$, then p_2 is the unique polynomial of degree < 2 that agrees with p at τ_1 and τ_2, hence, necessarily,

$$p_2 = p(\tau_1) + (\cdot - \tau_1)\frac{p(\tau_2) - p(\tau_1)}{\tau_2 - \tau_1}.$$

Thus, altogether,

$$\Delta(\tau_1, \tau_2)p \;=\; \begin{cases} Dp(\tau_1), & \text{if } \tau_1 = \tau_2 \\ \dfrac{p(\tau_2) - p(\tau_1)}{\tau_2 - \tau_1}, & \text{otherwise.} \end{cases}$$

In particular, for $\tau_1 \neq \tau_2$, $\Delta(\tau_1, \tau_2)p$ is, indeed, a divided difference.

4. Continuous Dependence of the Divided Difference on Its Sites

Since the polynomial $W_\tau a$ is evidently a continuous function of the τ_i, the same must be true of the sequence

$$W_\tau^{-1} p = (\Delta(\tau_{1:j})p : j = 1, 2, \ldots).$$

For a (more detailed) proof of this claim, recall the identity

$$A^{-1} - B^{-1} \;=\; A^{-1}(B - A)B^{-1}, \tag{4}$$

valid for any two invertible linear maps A and B (with the same domain and target). Fix some $p \in \Pi_{<n}$, and let

$$W_{n,\tau} : \mathbb{F}^n \to \Pi_{<n} : a \mapsto \sum_{j=1}^{n} w_{j-1,\tau}\, a(j).$$

Then, $W_{n,\tau}$ is an invertible linear map, and is bounded uniformly for $\tau_{1:n}$ in any bounded subset of $\mathbb{F}^n$, hence so is its inverse. Therefore, since $\lim_{\sigma\to\tau} W_{n,\sigma} = W_{n,\tau}$, we know, with (4), that

$$\lim_{\sigma\to\tau}(\Delta(\sigma_{1:j})p : j = 1{:}n) = (\Delta(\tau_{1:j})p : j = 1{:}n),$$

while $\Delta(\sigma_{1:j})p = 0 = \Delta(\tau_{1:j})p$ for all $j > n$.

5. Hermite Interpolation and Its Error

The fact that $w_{n,\tau}$ divides $p - p_n$ implies that p_n agrees with p at $\tau_1, \ldots, \tau_n$ even counting multiplicities, in the sense that

$$D^j p_n(z) = D^j p(z), \quad 0 \le j < \mu_z := \#\{i \in \{1, \ldots n\} : z = \tau_i\}. \tag{5}$$

In other words, p_n is the Hermite interpolant to p at $\tau_1, \ldots, \tau_n$.

By (2),

$$p_{n+1} \;=\; p_n + w_{n,\tau}\Delta(\tau_{1:n+1})p$$

while p_{n+1} agrees with p at τ_{n+1}, hence, again by (2), also

$$p_{n+1} \;=\; p_n + w_{n,\tau} q_n \quad \text{at} \quad \tau_{n+1}.$$

Since τ_{n+1} is arbitrary here, it follows that

$$\Delta(\tau_{1:n}, \cdot)p \;=\; q_n, \tag{6}$$

at least off the set $\{\tau_1, \ldots, \tau_n\}$, but then also on that set since $\Delta(\tau_{1:n}, \cdot)p$ and q_n both are continuous. This, incidentally shows that $\Delta(\tau_{1:n}, \cdot)p$ is a polynomial, of degree

$$\max(-1, \deg p - n).$$

More importantly, it shows that

$$p \;=\; p_n + w_{n,\tau}\Delta(\tau_{1:n}, \cdot)p,$$

the standard error formula for Hermite interpolation.

6. The Recurrence Relation

By (2),

$$q_n \;=\; \sum_{j>n} \frac{w_{j-1,\tau}}{w_{n,\tau}}\Delta(\tau_{1:j})p,$$

with the right side evidently a Newton form with respect to the centers $(\tau_{n+1}, \tau_{n+2}, \ldots)$. Therefore, with (6), we get the **basic divided difference identity**

$$\Delta(\tau_{n+1:j})\Delta(\tau_{1:n}, \cdot) \;=\; \Delta(\tau_{1:j}), \tag{7}$$

and it includes the recurrrence relation as the special case $j = n+2$, but, in view of (3), also provides a formula for the derivatives of $\Delta(\tau_{1:n}, \cdot)$.

7. Kowalewski's Definition

Since the divided difference is a symmetric function of its sites, we may always assume, without loss of generality, that all multiplicities in $\tau_{1:n}$ are **clustered**, i.e., $i < k < j$ and $\tau_i = \tau_j$ implies $\tau_i = \tau_k$. With that assumption, the following consequence of (3) and (7),

$$\Delta(\tau_{i:j})p = \left\{\begin{array}{ll} D^{j-i}p(\tau_i)/(j-i)!, & \text{if } \tau_i = \cdots = \tau_j \\ \dfrac{\Delta(\tau_{i+1:j})p - \Delta(\tau_{i:j-1})p}{\tau_j - \tau_i}, & \text{if } \tau_i \neq \tau_j \end{array}\right\}, \quad i \le j,$$

provides a means for computing recursively all the divided differences $\Delta(\tau_{i:j})p$, $1 \le i < j \le n$, from the n numbers $D^j p(z)$ mentioned in (5), i.e., the numbers

$$\nu_{j,\tau}p \;:=\; D^{\#\{i<j:\tau_i=\tau_j\}}p(\tau_j), \quad j = 1{:}n.$$

These calculations are customarily viewed as filling in the corresponding **divided difference table**.

The goal of these calculations is the computation of the sequence $(\Delta(\tau_{1:j})p : j = 1{:}n)$ of the coefficients in the Newton form for the Hermite interpolant, p_n, of p. But we could start out with the numbers $(\nu_{j,\tau}f : j = 1{:}n)$ for any smooth enough function f and obtain, in this way, the Hermite interpolant to f at $\tau_{1:n}$. Its **leading coefficient**, i.e., the coefficient of the $n{-}1$st power, is, by Kowalewski's definition, $\Delta(\tau_{1:n})f$. This makes $\Delta(\tau_{1:n})f$ the divided difference at $\tau_{1:n}$ of any polynomial that agrees with f at $\tau_{1:n}$ (counting multiplicities). To be sure, there are such polynomials since the sequence $(\nu_{j,\tau} : j = 1{:}n)$ is linearly independent even over $\Pi_{<n}$.

This extension of $\Delta(\tau_{1:n})$, to functions other than polynomials, maintains its description, provided implicitly by the above calculations, as an element of the span of $(\nu_{j,\tau} : j = 1{:}n)$.

8. Notation

The nonstandard notation, $\Delta(\tau_1, \ldots, \tau_n)$, used in this note, I learned from Kahan [**?**] some time ago. It cannot be confused, as can the standard notation $[\tau_1, \ldots, \tau_n]$, with standard notations for intervals or matrices, and has the advantage of being literal (given that Δ is standard notation for a difference). Here is a TEX macro for it:

```
\def\divdif{\mathord\kern.43em\vrule width.6pt height5.6pt
              depth.-28pt \kern-.43em\Delta}
```

References

[1] Eberhard Hopf, "Über die Zusammenhänge zwischen gewissen höheren Differenzenquotienten reeller Funktionen einer reellen Variablen und deren Differenzierbarkeitseigenschaften", dissertation, Universität Berlin (30 pp), 1926.

[2] W. Kahan, "Divided differences of algebraic functions (Notes for Math 228A, Fall 1974)", Department of Mathematics, UC Berkeley, 1974.

[3] G. Kowalewski, "Interpolation und genäherte Quadratur", B. G. Teubner, Berlin, 1932.

Carl de Boor
P.O. Box 1076
Eastsound WA 98245
USA
E-mail: deboor@cs.wisc.edu