Error formulas for divided difference expansions and numerical differentiation

Michael S. Floater

Abstract: We derive an expression for the remainder in divided difference expansions and use it to give new error bounds for numerical differentiation.

Key words: divided differences, numerical differentiation, boundary-value problems, Lagrange interpolation.

1. Introduction

There are many applications in numerical analysis of divided difference expansions of the form

$$[x_0, \ldots, x_n]f = \sum_{k=n}^{p-1} c_k f^{(k)}(x)/k! + R_p. \tag{1.1}$$

Here, and throughout the paper, we will assume that $x_0 \le x_1 \le \ldots \le x_n$ are arbitrarily spaced real values and x is any real value in the interval $[x_0, x_n]$. We refer the reader to Conte and de Boor [**1**] for basic properties of divided differences. Two things are required: evaluation of the coefficients c_k; and a bound on the remainder term R_p in terms of the maximum grid spacing

$$h := \max_{0 \le i \le n-1} (x_{i+1} - x_i).$$

We take as our canonical example the finite difference expansion

$$\frac{f(x_0) - 2f(x_1) + f(x_2)}{2h^2} = \frac{f''(x_1)}{2!} + h^2 \frac{f^{(4)}(x_1)}{4!} + \ldots + h^{p-2} \frac{f^{(p)}(\xi)}{p!}, \tag{1.2}$$

in which $n = 2$, $x_1 - x_0 = x_2 - x_1 = h$, $x = x_1$, p is even, and ξ is some number in $[x_0, x_2]$. The choice $p = 4$ implies

$$\frac{f(x_0) - 2f(x_1) + f(x_2)}{h^2} - f''(x_1) = \frac{h^2}{12} f^{(4)}(\xi), \tag{1.3}$$

which gives the well known error formula for the approximation of the second derivative of the function f by a second order finite difference, which appears in many books and is used in deriving standard finite difference schemes for second order boundary-value problems with $O(h^2)$ accuracy; see, for example, Varga [**12**] or Keller [**6**]. The more general expansion (1.2) has been used to derive schemes with higher order of approximation. Analogously, the most general expansion (1.1) plays a basic role in deriving difference schemes for both higher order equations and non-uniform grids. Schemes for non-uniform grids have been developed and studied by Osborne [**8**], Doedel [**2**], and Kreiss et al. [**7**].

The usual approach to finding the coefficients c_k is to use the Taylor series

$$f(y) = \sum_{k=0}^{p-1} (y - x)^k f^{(k)}(x)/k! + r_p, \tag{1.4}$$

for some remainder term r_p. Applying the divided difference $[x_0, \dots, x_n]$, gives

$$[x_0, \dots, x_n]f = \sum_{k=n}^{p-1} [x_0, \dots, x_n](\cdot - x)^k f^{(k)}(x)/k! + R_p.$$

This, then, provides the coefficients c_k, which were already found by Steffensen (in Section 76 of [**11**]). He also observed that they can be expressed in the more explicit form,

$$c_k = \sigma_{k-n}(x_0 - x, \dots, x_n - x),$$

where σ_j is the symmetric polynomial of degree j,

$$\sigma_j(\delta_0, \dots, \delta_n) = \sum_{\tau_0 + \dots + \tau_n = j} \delta_0^{\tau_0} \dots \delta_n^{\tau_n}. \tag{1.5}$$

The first examples of c_k are therefore

$$c_n = 1, \quad c_{n+1} = \sum_{0 \le i \le n} (x_i - x), \quad c_{n+2} = \sum_{0 \le i \le j \le n} (x_i - x)(x_j - x). \tag{1.6}$$

As regards a bound on the remainder term R_p, the following theorem is known in many special cases. We will assume that the function f belongs to $C^p[x_0, x_n]$ and we denote by $\|.\|$ the max norm over $[x_0, x_n]$.

Theorem 1. *There exists a constant C, depending only on n and p, such that*

$$|R_p| \le C h^{p-n} \|f^{(p)}\|. \tag{1.7}$$

Consider, for example, the special case in which the data points $x_0, \dots, x_n$ are *uniformly spaced*, i.e., such that $x_{i+1} - x_i = h$ (as in equation (1.2)), though $x \in [x_0, x_n]$ may be arbitrary. Then the theorem is well known and easily established using the simple derivative remainder term

$$r_p = (y - x)^p f^{(p)}(\xi_y)/p!,$$

in the Taylor expansion (1.4), where ξ_y is some number between x and y. This is because the application of the divided difference $[x_0, \dots, x_n]$ to (1.4) only involves division by differences of the form $x_j - x_i$. In the uniform case, all such differences are multiples of h, namely $(j - i)h$.

Also in the case $p = n + 1$, with $x_0, \dots, x_n$ arbitrary, Theorem 1 is known, and was proved by Isaacson and Keller [**5**]. Their proof uses the fact that the remainder R_{n+1} is the n-th derivative of the error in interpolating f by a Lagrange (or Hermite) polynomial of degree n at the points $x_0, \dots, x_n$ (up to a factor of $n!$).

The question of whether the theorem holds in general seems less straightforward however. One of the purposes of this paper is to give a simple proof of Theorem 1 in full generality, by deriving a new formula for the remainder R_p. We further show that when $p - n$ is even, the remainder can be expressed in a form similar to that of (1.2).

Theorem 2. *When $p-n$ is even, there exists $\xi \in [x_0, x_n]$ such that*

$$R_p = \sigma_{p-n}(x_0 - x, \dots, x_n - x) f^{(p)}(\xi)/p!.$$

We remark that Steffensen (in Section 76 of [**11**]) proved that the remainder R_p has this form for *all* p in the case that x lies *outside* the interval (x_0, x_n).

Finally, we study the important case $p = n+2$, which occurs frequently in finite difference schemes. The reason is that if x is chosen to be the average,

$$\bar{x} = \frac{x_0 + \ldots + x_n}{n+1}, \tag{1.8}$$

then the coefficient c_{n+1} in (1.6) is zero, so that

$$[x_0, \dots, x_n] f = f^{(n)}(\bar{x})/n! + R_{n+2}.$$

Since Theorem 1 shows that $|R_{n+2}| \le Ch^2 \|f^{(n+2)}\|$, the divided difference $[x_0, \dots, x_n]f$ offers a higher order approximation to $f^{(n)}/n!$ at the point $\bar{x}$. This enables finite difference schemes on non-uniform grids to be designed with $O(h^2)$ truncation error and therefore $O(h^2)$ convergence; see [**2**] and [**7**].

Due to Theorem 2, we prove a more precise result.

Theorem 3. *With $\bar{x}$ as in (1.8),*

$$|n![x_0, \dots, x_n]f - f^{(n)}(\bar{x})| \le \frac{n}{24} h^2 \|f^{(n+2)}\|, \tag{1.9}$$

and if $x_0, \dots, x_n$ are uniformly spaced, the constant $n/24$ is the least possible.

We complete the paper with some examples.

2. New remainder formula

Consider what happens if we use one of the more precise remainder terms in the Taylor series (1.4). For example, if we use the divided difference remainder,

$$r_p = (y-x)^p [y, \underbrace{x, \dots, x}_{p}] f,$$

then, using the Leibniz rule, we get in (1.1) the remainder formula

$$\begin{aligned} R_p &= [x_0, \dots, x_n](\cdot - x)^p [\cdot, \underbrace{x, \dots, x}_{p}] f \\ &= \sum_{i=0}^{n} [x_i, \dots, x_n](\cdot - x)^p [x_0, \dots, x_i, \underbrace{x, \dots, x}_{p}] f \\ &= \sum_{i=0}^{n} \sigma_{p-n+i}(x_i - x, \dots, x_n - x) [x_0, \dots, x_i, \underbrace{x, \dots, x}_{p}] f. \end{aligned} \tag{2.1}$$

However, this remainder formula is not useful for us because it involves divided differences of f of all orders from p to $p+d$, which in general will not be well defined for $f \in C^p[x_0, x_n]$.

The other well known remainder for the Taylor expansion (1.4) is the integral one,

$$r_p = \frac{1}{(p-1)!} \int_x^y (y-s)^{p-1} f^{(p)}(s)\, ds.$$

Applying $[x_0, \ldots, x_n]$ will give an expression for R_p, and by introducing truncated powers, this can be reformulated in terms of a kernel. A kernel approach was used by both Howell [**4**] and Shadrin [**10**] to give a more precise bound than Isaacson and Keller [**5**] on R_{n+1}. However, Theorem 1 can be established using purely elementary properties of divided differences, and without kernels. In Section 5 we show that also Howell and Shadrin's bound on R_{n+1} follows from simple divided difference properties.

In fact we abandon the Taylor series altogether and derive a new formula for R_p, in terms of divided differences, in the spirit of the remainder formulas for Lagrange interpolation derived independently by Dokken and Lyche [**3**] and Wang [**13**].

Lemma 1. *With* $\delta_i = x_i - x$,

$$R_p = \sum_{i=0}^{n} \delta_i \sigma_{p-n-1}(\delta_i, \ldots, \delta_n)[x_0, \ldots, x_i, \underbrace{x, \ldots, x}_{p-i}]f. \tag{2.2}$$

This formula is better than (2.1) because it only involves divided differences of f of the same order p. Note also that though the formula is not symmetric in the points $x_0, \ldots, x_n$, it holds for any permutation of them, an observation we take advantage of when proving Theorem 2.

Proof: The case $p = n+1$ is a special case of the remainder formula of Dokken and Lyche [**3**] and Wang [**13**]. Dokken and Lyche argue that

$$\begin{aligned}
[x_0, \ldots, x_n]f &= [\underbrace{x, \ldots, x}_{n+1}]f + ([x_0, \ldots, x_n]f - [\underbrace{x, \ldots, x}_{n+1}]f) \\
&= \frac{f^{(n)}(x)}{n!} + \sum_{i=0}^{n}([x_0, \ldots, x_i, \underbrace{x, \ldots, x}_{n-i}]f - [x_0, \ldots, x_{i-1}, \underbrace{x, \ldots, x}_{n-i+1}]f) \\
&= \frac{f^{(n)}(x)}{n!} + \sum_{i=0}^{n}(x_i - x)[x_0, \ldots, x_i, \underbrace{x, \ldots, x}_{n-i+1}]f.
\end{aligned}$$

We prove (2.2) in general by induction on p. We assume (2.2) holds for $p > n$ and show that it also holds for $p+1$. Indeed, recalling equation (1.5),

$$
\begin{aligned}
R_p &= \sum_{i=0}^{n} \Big(\sigma_{p-n}(\delta_i,\ldots,\delta_n) - \sigma_{p-n}(\delta_{i+1},\ldots,\delta_n)\Big)[x_0,\ldots,x_i,\underbrace{x,\ldots,x}_{p-i}]f \\
&= \sigma_{p-n}(\delta_0,\ldots,\delta_n)\frac{f^{(p)}(x)}{p!} + \sum_{i=0}^{n}\sigma_{p-n}(\delta_i,\ldots,\delta_n) \\
&\qquad\qquad \Big([x_0,\ldots,x_i,\underbrace{x,\ldots,x}_{p-i}]f - [x_0,\ldots,x_{i-1},\underbrace{x,\ldots,x}_{p+1-i}]f\Big) \\
&= \sigma_{p-n}(\delta_0,\ldots,\delta_n)\frac{f^{(p)}(x)}{p!} + R_{p+1}.
\end{aligned}
$$

■

Interestingly, the above proof derives both the remainder R_p *and* the coefficients c_k of the expansion (1.1), *without* using a Taylor series.

Proof of Theorem 1: This follows from Lemma 1 and the fact that $|\delta_i| \le nh$. In fact the constant C in (1.7) can be taken to be $n^{p-n}/(n!(p-n)!)$, because

$$
\begin{aligned}
|R_p| &\le \sum_{i=0}^{n} |\delta_i|\sigma_{p-n-1}(|\delta_i|,\ldots,|\delta_n|)\|f^{(p)}\|/p! \\
&= \sigma_{p-n}(|\delta_0|,\ldots,|\delta_n|)\|f^{(p)}\|/p! \\
&\le \sigma_{p-n}(nh,\ldots,nh)\|f^{(p)}\|/p! = \binom{p}{n} n^{p-n}h^{p-n}\|f^{(p)}\|/p!.
\end{aligned}
$$

■

We turn next to Theorem 2 and begin with a basic property of the polynomials σ_j.

Lemma 2. *If $j \ge 1$ is odd, any set of real values $\delta_0,\ldots,\delta_n$ can be permuted so that the $n+1$ products*

$$
\delta_0\sigma_j(\delta_0,\ldots,\delta_n), \quad \delta_1\sigma_j(\delta_1,\ldots,\delta_n), \quad \ldots, \quad \delta_n\sigma_j(\delta_n), \tag{2.3}
$$

are simultaneously non-negative.

Proof: We start with the first term and consider two possible cases. If $\sigma_j(\delta_0,\ldots,\delta_n) \ge 0$, then at least one of the δ_i must be non-negative. Indeed, if all the δ_i were negative, then $\sigma_j(\delta_0,\ldots,\delta_n)$ would also be negative, due to j being odd in (1.5). We can therefore permute $\delta_0,\ldots,\delta_n$ so that δ_0 is non-negative, which implies that

$$
\delta_0\sigma_j(\delta_0,\ldots,\delta_n) \ge 0. \tag{2.4}
$$

Similarly, if $\sigma_j(\delta_0,\ldots,\delta_n) \le 0$, then at least one of the δ_i must be non-positive, in which case we choose δ_0 to be non-positive, so that inequality (2.4) holds again.

We continue in this way, next choosing δ_1 from the remaining values $\delta_1, \ldots, \delta_n$ to ensure that the second term in (2.3) is non-negative, and so on. The last term is trivially non-negative because

$$\delta_n \sigma_j(\delta_n) = \sigma_{j+1}(\delta_n) = \delta_n^{j+1} \geq 0.$$

■

Proof of Theorem 2: Since $p - n - 1$ is odd, Lemma 2 implies the existence of a permutation of the points $x_0, \ldots, x_n$ such that the $n+1$ coefficients $\delta_i \sigma_{p-n-1}(\delta_i, \ldots, \delta_n)$ in equation (2.2) are simultaneously non-negative. The result then follows from the Mean Value Theorem and the observation that the coefficients sum to $\sigma_{p-n}(\delta_0, \ldots, \delta_n)$. ■

Note that the above analysis implies that $\sigma_j(\delta_0, \ldots, \delta_n)$ is non-negative for any real values $\delta_0, \ldots, \delta_n$ when j is even, but this is well known and follows from the fact that $\sigma_j(\delta_0, \ldots, \delta_n) = \binom{n+j}{j} \xi^j$ for some point ξ in the interval containing the δ_i (see equation (41) of [**11**]).

3. Optimal error bounds

We next consider Theorem 3. Like Theorem 2, it follows from an elementary property of the symmetric polynomials σ_j in (1.5).

Lemma 3. *If* $\delta_0 + \ldots + \delta_n = 0$, *then*

$$\sigma_2(\delta_0, \ldots, \delta_n) = \frac{1}{2(n+1)} \sum_{0 \leq i < j \leq n} (\delta_j - \delta_i)^2.$$

Proof: This follows from the two identitites

$$\sum_{0 \leq i \leq j \leq n} (\delta_j - \delta_i)^2 = (n+2) \sum_{i=0}^{n} \delta_i^2 - 2 \sum_{0 \leq i \leq j \leq n} \delta_i \delta_j,$$

and

$$0 = \left(\sum_{i=0}^{n} \delta_i \right)^2 = - \sum_{i=0}^{n} \delta_i^2 + 2 \sum_{0 \leq i \leq j \leq n} \delta_i \delta_j.$$

■

Proof of Theorem 3: Putting $x = \bar{x}$ and $p = n+2$ in the expansion (1.1), Theorem 2 implies

$$[x_0, \ldots, x_n] f = f^{(n)}(\bar{x})/n! + \sigma_2(\delta_0, \ldots, \delta_n) f^{(n+2)}(\xi)/(n+2)!.$$

So Lemma 3 implies that

$$n![x_0, \ldots, x_n] f - f^{(n)}(\bar{x}) = \frac{1}{2(n+1)^2(n+2)} \sum_{0 \leq i < j \leq n} (x_j - x_i)^2 f^{(n+2)}(\xi). \qquad (3.1)$$

The inequality (1.9) now results from the observation that

$$\sum_{0 \leq i < j \leq n} (x_j - x_i)^2 \leq \sum_{0 \leq i < j \leq n} (j-i)^2 h^2 = \frac{n(n+1)^2(n+2)}{12} h^2. \qquad (3.2)$$

In the uniform case, $x_{i+1} - x_i = h$, inequality (3.2) becomes an equality, so that equation (3.1) reduces to

$$\frac{\Delta^n f(x_0)}{h^n} - f^{(n)}(\bar{x}) = \frac{n}{24} h^2 f^{(n+2)}(\xi), \tag{3.3}$$

where $\bar{x} = (x_0 + x_n)/2$ and $\Delta^n f(x_0)$ denotes the n-th order finite difference,

$$\Delta^n f(x_0) = \sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} f(x_i),$$

So if we set $f(x) = x^{n+2}$, the error bound (1.9) becomes an equality. ■

4. Examples

Though Theorem 3 gives a simple error bound for non-uniformly spaced points, better bounds can be derived for specific configurations of $x_0, \ldots, x_n$ by going back to the exact equation (3.1). For example, in the simplest non-uniform case, namely $n = 2$, equation (3.1) reduces to

$$2[x_0, x_1, x_2]f - f''(\bar{x}) = \frac{1}{36}(h_0^2 + h_0 h_1 + h_1^2) f^{(4)}(\xi),$$

where $h_0 = x_1 - x_0$ and $h_1 = x_2 - x_1$. Various approaches have been used to show that the approximation $2[x_0, x_1, x_2]f$ to $f''(\bar{x})$ is of order $O(h^2)$, see, for example, Samarskii, Vabishchevich, and Matus [**9**], but the above exact formula appears to be new. Taking for example the Hermite case $x_2 = x_1$, so that $h = h_0$ and $h_1 = 0$, then with $\bar{x} = (x_0 + 2x_1)/3$, we get the optimal error bound

$$|2[x_0, x_1, x_1]f - f''(\bar{x})| \le \frac{h^2}{36} \|f^{(4)}\|.$$

Another example could be the case $h = h_0 = 2h_1$, giving the optimal bound

$$|2[x_0, x_1, x_2]f - f''(\bar{x})| \le \frac{7}{144} h^2 \|f^{(4)}\|.$$

When the points $x_0, \ldots, x_n$ are uniformly spaced, the error formula (3.3) is known for $n = 1, 2$, but appears to be new for $n \ge 3$. The case $n = 2$ reduces to equation (1.3). The case $n = 3$ gives a new error formula for a well known approximation,

$$\frac{-f(x_0) + 3f(x_1) - 3f(x_2) + f(x_3)}{h^3} - f^{(3)}(\bar{x}) = \frac{h^2}{8} f^{(5)}(\xi)$$

with $\bar{x} = (x_1 + x_2)/2$. The case $n = 4$ gives the new error formula

$$\frac{f(x_0) - 4f(x_1) + 6f(x_2) - 4f(x_3) + f(x_4)}{h^4} - f^{(4)}(x_2) = \frac{h^2}{6} f^{(6)}(\xi).$$

5. Howell and Shadrin's error bound

Shadrin [**10**] has shown that if p_n denotes the polynomial of degree n interpolating f at the points $x_0, \ldots, x_n$, then for $k = 0, 1, \ldots, n$,

$$|p_n^{(k)}(x) - f^{(k)}(x)| \le \|\psi_n^{(k)}\| \, \frac{\|f^{(n+1)}\|}{(n+1)!},$$

where

$$\psi_n(x) = (x - x_0)(x - x_1)\ldots(x - x_n).$$

This bound was earlier conjectured by Howell [**4**] who also proved it for the highest derivative $k = n$. Both Howell and Shadrin used kernels and properties of B-splines to establish the case $k = n$. We now offer an elementary proof using the simple remainder formula of Dokken, Lyche, and Wang,

$$p_n^{(n)}(x) - f^{(n)}(x) = n! \sum_{i=0}^{n} (x_i - x)[x_0, \ldots, x_i, \underbrace{x, \ldots, x}_{n-i+1}]f.$$

Note that since $\sum_{i=0}^n |x_i - x|$ is a convex function of x, its maximum value in the interval $[x_0, x_n]$ is attained at one of the end points, where it agrees with $|\psi^{(n)}(x)|/n!$. Therefore,

$$\begin{aligned} |p_n^{(n)}(x) - f^{(n)}(x)| &\le n! \sum_{i=0}^{n} |x_i - x| \, \frac{\|f^{(n+1)}\|}{(n+1)!} \\ &\le n! \max\left\{ \sum_{i=0}^{n} |x_i - x_0|, \sum_{i=0}^{n} |x_i - x_n| \right\} \frac{\|f^{(n+1)}\|}{(n+1)!} \\ &= \max\{|\psi_n^{(n)}(x_0)|, |\psi_n^{(n)}(x_n)|\} \, \frac{\|f^{(n+1)}\|}{(n+1)!} \\ &\le \|\psi_n^{(n)}\| \, \frac{\|f^{(n+1)}\|}{(n+1)!}. \end{aligned}$$

Acknowledgement. I wish to thank Carl de Boor, Tom Lyche, and the referee for valuable comments which helped in the revised version of this paper.

References

1. S. D. Conte and C. de Boor, *Elementary numerical analysis*, McGraw-Hill, 1980.
2. E. J. Doedel, The construction of finite difference approximations to ordinary differential equations, SIAM J. Numer. Anal. **15** (1978), 450–465.
3. T. Dokken and T. Lyche, A divided difference formula for the error in Hermite interpolation, BIT **19** (1979), 539–540.
4. G. W. Howell, Derivative error bounds for Lagrange interpolation, J. Approx. Th. **67** (1991), 164–173.

5. E. Isaacson and H. B. Keller, *Analysis of numerical methods*, Wiley, 1966.
6. H. B. Keller, *Numerical methods for two-point boundary-value problems*, Dover, 1992.
7. H.-O. Kreiss, T. A. Manteuffel, B. Swartz, B. Wendroff, and A. B. White, Jr., Supraconvergent schemes on irregular grids, Math. Comp. **47** (1986), 537–554.
8. M. R. Osborne, Minimizing truncation error in finite difference approximations to ordinary differential equations, Math. Comp **21** (1967), 133–145.
9. A. A. Samarskii, P. N. Vabishchevich, and P. P. Matus, Finite-difference approximations of higher accuracy order on non-uniform grids, Differents. Urav. **32** (1996), 265–274.
10. A. Shadrin, Error bounds for Lagrange interpolation, J. Approx. Th. **80** (1995), 25–49.
11. J. F. Steffensen, *Interpolation*, Baltimore, 1927.
12. R. S. Varga, *Matrix iterative analysis*, Prentice-Hall, NJ, 1962.
13. X. Wang, On remainders of numerical differentiation formulas, Ke Xue Tong Bao (in Chinese) **24(19)** (1979), 869–872.

Michael S. Floater
SINTEF
Postbox 124, Blindern
0314 Oslo
NORWAY
`mif@sintef.no`