Divided differences

Nicholas Hu

Let $\{x_j\}_{j=0}^n \subseteq \mathbb{R}$ be a set of *distinct* abscissae (also called "nodes") and let P_n denote the real vector space of polynomials of degree at most n with real coefficients. If f is a real-valued function defined on $\{x_j\}_{j=0}^n$, then there exists a polynomial $p_n \in P_n$ that *interpolates* f at the abscissae in the sense that $p_n(x_j) = f(x_j)$ for each j: namely, $p_n = \sum_j f(x_j)\ell_j$, where $\ell_j(x) = \prod_{i \neq j} \frac{x - x_i}{x_j - x_i}$. Furthermore, it is the only such polynomial, for if $q_n \in P_n$ also interpolates f at the same abscissae, then $p_n - q_n \in P_n$ has $n+1$ distinct roots and must therefore be the zero polynomial.

The polynomials $\{\ell_j\}_{j=0}^n$ constitute a basis of P_n known as the **Lagrange basis**. Other bases of P_n include the **monomial basis** $\{\varphi_j\}_{j=0}^n$, given by $\varphi_j(x) = x^j$, and the **Newton basis** $\{\omega_j\}_{j=0}^n$, given by $\omega_j(x) = \prod_{i<j}(x - x_i)$.

Divided differences are quantities that can be used (among other things) to compute the coefficients of p_n in the Newton basis. We define the **divided difference** $f[x_0, x_1, \ldots, x_n]$ as the coefficient of φ_n in the monomial basis representation of p_n (which for simplicity we will refer to as the "leading coefficient" of p_n despite the fact that it may be zero). Below we present several properties of divided differences.

> **Coefficients of polynomial interpolant in Newton basis**
>
> $$p_n = \sum_j f[x_0, \ldots, x_j]\omega_j$$

Proof. Write $p_n = \sum_j c_j\omega_j$. For each j, the polynomial $\sum_{i \leq j} c_i\omega_i \in P_j$ interpolates f at $\{x_0, \ldots, x_j\}$ since $\omega_k(x_i) = 0$ for all $i \leq j < k$. Clearly, its leading coefficient is c_j, so by definition, $c_j = f[x_0, \ldots, x_j]$. ∎

Notably, divided differences obey a recurrence relation that allows for their recursive computation.

> **Recurrence relation**
>
> $$\begin{aligned} f[x_0, \ldots, x_n] &= \frac{f[x_1, \ldots, x_n] - f[x_0, \ldots, x_{n-1}]}{x_n - x_0} \\ f[x_0] &= f(x_0) \end{aligned}$$

Proof. Suppose that $n \geq 1$. If $p_+ \in P_{n-1}$ interpolates f at $\{x_1, \ldots, x_n\}$ and $p_- \in P_{n-1}$ interpolates f at $\{x_0, \ldots, x_{n-1}\}$, then the polynomial $p \in P_n$ given by $p(x) = p_-(x) + \frac{x - x_0}{x_n - x_0}(p_+ - p_-)(x)$ interpolates f at $\{x_0, \ldots, x_n\}$. For the base case, observe that the constant polynomial $f(x_0) \in P_0$ interpolates f at $\{x_0\}$. ∎

Interestingly, we observe that $p(x)$ is a convex combination of $p_-(x)$ and $p_+(x)$ for $x_0 \leq x \leq x_n$. The formula above also yields a recursive algorithm for evaluating p known as **Neville's algorithm**. Namely, if $p_{i,j} \in P_{j-i}$ interpolates f at $\{x_i, \ldots, x_j\}$ so that $p = p_{0,n}$, then $p_{i,j}(x) = \frac{(x - x_i)p_{i+1,j}(x) - (x - x_j)p_{i,j-1}(x)}{x_j - x_i}$, where $p_{i,i}(x) = f(x_i)$.

An explicit formula for divided differences can also be deduced by inspecting the Lagrange basis representation of p_n.

$$f[x_0, \ldots, x_n] = \sum_j f(x_j) \prod_{i \neq j} \frac{1}{x_j - x_i}$$

Linearity

If $\alpha, \beta \in \mathbb{R}$, then

$$(\alpha f + \beta g)[x_0, \ldots, x_n] = \alpha(f[x_0, \ldots, x_n]) + \beta(g[x_0, \ldots, x_n]).$$

Proof. If $p_f \in P_n$ and $p_g \in P_n$ interpolate f and g, respectively, at $\{x_0, \ldots, x_n\}$, then $\alpha p_f + \beta p_g \in P_n$ interpolates $\alpha f + \beta g$ at $\{x_0, \ldots, x_n\}$. ∎

Symmetry

If σ is a permutation of $\{0, \ldots, n\}$, then

$$f[x_0, \ldots, x_n] = f[x_{\sigma(0)}, \ldots, x_{\sigma(n)}].$$

Proof. In view of the definition above, this is immediate since $\{x_0, \ldots, x_n\} = \{x_{\sigma(0)}, \ldots, x_{\sigma(n)}\}$. ∎

Factor property

If $g(x) = (x - x_0)f(x)$ and $n \geq 1$, then

$$g[x_0, x_1, \ldots, x_n] = f[x_1, \ldots, x_n].$$

Proof. If $p_f \in P_{n-1}$ interpolates f at $\{x_1, \ldots, x_n\}$, then the polynomial $p_g \in P_n$ given by $p_g(x) = (x - x_0)p_f(x)$ interpolates g at $\{x_0, \ldots, x_n\}$. ∎

In fact, the recurrence relation (excluding the base case) can be derived solely from linearity, symmetry, and the factor property:

$$\begin{aligned}(x_n - x_0)(f[x_0, \ldots, x_n]) &= ([(x - x_0) - (x - x_n)]f)[x_0, \ldots, x_n] \\ &= f[x_1, \ldots, x_n] - f[x_0, \ldots, x_{n-1}]\end{aligned}$$

Thus, these three properties along with the property $f[x_0] = f(x_0)$ determine the values of all divided differences.

The identity $f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$ suggests a relationship between divided differences and derivatives: if, say, $x_0 < x_1$, $f \in C([x_0, x_1])$, and f' exists on (x_0, x_1), then the mean value theorem amounts to the assertion that $f[x_0, x_1] = f'(\xi)$ for some $\xi \in (x_0, x_1)$. This generalizes readily to divided differences and derivatives of higher order.

Mean value theorem

Let $a = \min \{x_j\}_{j=0}^n$ and $b = \max \{x_j\}_{j=0}^n$. If $f \in C([a, b])$ and $f^{(n)}$ exists on (a, b), then

$$f[x_0, \ldots, x_n] = \frac{f^{(n)}(\xi)}{n!}$$

for some $\xi \in (a, b)$.

Proof. By symmetry, we may assume that $a = x_0 < \cdots < x_n = b$. If $p \in P_n$ interpolates f at $\{x_0, \ldots, x_n\}$, then $f - p$ has $n+1$ distinct zeroes in $[x_0, x_n]$. By (repeated applications of) Rolle's theorem, $(f-p)^{(n)} = f^{(n)} - f[x_0, \ldots, x_n]n!$ has a zero $\xi \in (x_0, x_n)$. ∎

> **Polynomial interpolation error**
>
> Let $a = \min\{x_j\}_{j=0}^n$ and $b = \max\{x_j\}_{j=0}^n$. If $f \in C([a,b])$ and $f^{(n+1)}$ exists on (a,b), then for all $x \in [a,b]$ there exists a $\xi \in (a,b)$ for which
>
> $$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\prod_j (x - x_j).$$

Proof. If $x \in \{x_j\}_{j=0}^n$, the conclusion is trivial; otherwise, the polynomial $p_n + f[x_0, \ldots, x_n, x]\omega_{n+1}$ interpolates f at $\{x_0, \ldots, x_n, x\}$ and the conclusion follows from the mean value theorem. ∎

In the same vein, $f[x_0, x_0 + h] = \frac{f(x_0+h) - f(x_0)}{h} \to f'(x_0)$ as $h \to 0$ if f is differentiable at x_0, the geometric interpretation being that the slope of the secant line through $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$ tends to that of the tangent line through $(x_0, f(x_0))$. This observation along with the identity $(fg)[x_0, x_1] = f[x_0]g[x_0, x_1] + f[x_0, x_1]g[x_1]$ allows us to recover the product rule for derivatives. More generally, we have the following identity for divided differences.

> **Product rule**
>
> $$(fg)[x_0, \ldots, x_n] = \sum_j f[x_0, \ldots, x_j]g[x_j, \ldots, x_n]$$

Proof. If $p \in P_n$ interpolates f at $\{x_0, \ldots, x_n\}$, then $(fg)[x_0, \ldots, x_n] = (pg)[x_0, \ldots, x_n]$ since fg agrees with pg on $\{x_0, \ldots, x_n\}$. By linearity and the factor property, we have

$$\begin{aligned}(pg)[x_0, \ldots, x_n] &= \left(\sum_j f[x_0, \ldots, x_j]\omega_j g\right)[x_0, \ldots, x_n] \\ &= \sum_j f[x_0, \ldots, x_j](\omega_j g)[x_0, \ldots, x_n] \\ &= \sum_j f[x_0, \ldots, x_j]g[x_j, \ldots, x_n]. \quad \blacksquare\end{aligned}$$

By Taylor's theorem, $(\Delta_h^k f)(x_0 + jh) = f^{(k)}(x_0)h^k + o(h^k)$ if f is k times differentiable at x_0, where Δ_h denotes the forward difference operator with step h (that is, $(\Delta_h f)(x) = f(x+h) - f(x)$). As a result,

$$f[x_0 + jh, \ldots, x_0 + nh] = \frac{(\Delta_h^{n-j} f)(x_0 + jh)}{(n-j)!h^{n-j}} \to \frac{f^{(n-j)}(x_0)}{(n-j)!}$$

and from the above we can derive the generalized product rule for derivatives, $(fg)^{(n)} = \sum_j \binom{n}{j} f^{(j)} g^{(n-j)}$.