CPSC 303: REMARKS ON DIVIDED DIFFERENCES

JOEL FRIEDMAN

Disclaimer: The material may sketchy and/or contain errors, which I will elaborate upon and/or correct in class. For those not in CPSC 303: use this material at your own risk...

The goal of this note is to fill in some details and give further examples regarding the *Newton polynomial*, also called *Newton's divided difference interpolation polynomial*, used in Sections 10.4–10.7 of the course textbook [A&G] by Ascher and Greif; this refers to the formula

$$
\begin{aligned}
p(x) = f[x_0] + (x-x_0)f[x_0,x_1] + (x-x_0)(x-x_1)f[x_0,x_1,x_2] + \cdots \\
+ (x-x_0)\ldots(x-x_{n-1})f[x_0,x_1,\ldots,x_n]
\end{aligned}
\tag{1}
$$

for the polynomial p that agrees with f at the $n+1$ points $x = x_0, x_1, \ldots, x_n$.

In the textbook we assume that $x_0, \ldots, x_n$ are distinct until Section 10.7—the the last section in Chapter—and Section 10.7 is extremely brief and states results without proof. However, the one of the main "selling points" of (1) is that it still holds (for sufficiently differentiable f) when some of the x_i are allowed to be the same, and that the divided difference $f[x_0, \ldots, x_n]$ remains a well-behaved funtion of $x_0, \ldots, x_n$ even when some x_i are equal or nearly equal.

1. A Proof of Newton's Divided Difference Interpolation Polynomial

The textbook [A&G] does not prove Newton's formula (1): it gives an important first step of the proof, and leaves the second step as an "challenging" exercise (Exercise 7 there) without hints. In CPSC 303 this year we gave the second step.

1.1. **Upper Triangular Systems.**

The first step given in [A&G] is a fundamental observation about "upper triangular change of basis" that occurs in many applications in many disciplines.

In terms of matrices, the point is that matrices that are *upper triangular*, such as

$$
\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}, \quad \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & e \end{bmatrix}
$$

are invertible provided that their diagonal entries are nonzero; furthermore the inverses are also upper triangular, and this can be proven by seeing that all steps in Gauss-Jordan elimination used to compute the inverse are "upper triangular

Research supported in part by an NSERC grant.

operations." [This is taught in CPSC 302, especially in in Chapter 5 of [A&G], which discusses the LU-decomposition.]

In the application in Section 10.4 to Newton polynomials, the story of "upper triangular change of basis" goes like this: if $\phi_0, \ldots, \phi_n$ are polynomials such that for all $i \in [n]$, ϕ_i is *exactly* of degree i, then any polynomial, $p = p(x)$, of degre n over $\mathbb{R}$,

$$p(x) = c_0 + c_1 x + \cdots + c_n x^n$$

can be uniquely expressed as a linear combination

$$p(x) = \alpha_0 \phi_0(x) + \cdots + \alpha_n \phi_n(x), \tag{2}$$

since the α_i's can be written in terms of the c_i's, and vice versa, in terms on an upper triangular matrix. This is an extremely important observation.

Example 1.1. For every c_0, c_1, c_2 there is a unique $\alpha_0, \alpha_1, \alpha_2$ such that

$$c_0 + c_1 x + c_2 x^2 = \alpha_0 + \alpha_1(x-1) + \alpha_2(x-1)^2 \tag{3}$$

(where $=$ means equal as polynomials), since

$$\alpha_0 + \alpha_1(x-1) + \alpha_2(x-1)^2 = x^2\alpha_2 + x\bigl(-2\alpha_2 + \alpha_1\bigr) + \bigl(\alpha_2 - \alpha_1 + \alpha_0\bigr),$$

and therefore (3) is equivalent to

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_2 \\ \alpha_1 \\ \alpha_0 \end{bmatrix} = \begin{bmatrix} c_2 \\ c_1 \\ c_0 \end{bmatrix};$$

we easily see that

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

and hence the linear system above is equivalent to

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_2 \\ c_1 \\ c_0 \end{bmatrix} = \begin{bmatrix} \alpha_2 \\ \alpha_1 \\ \alpha_0 \end{bmatrix}.$$

Example 1.2. Let $p(x) = 3 + 4x + 5x^2$, and consider the task of writing $p(x)$ as

$$\alpha_0(303) + \alpha_1(2020x + 2021) + \alpha_2(13x^2 + 18x + 120),$$

which is (2) in the special case $n = 2$ and $\phi_0 = 303$, $\phi_1 = 2020x + 2021$, $\phi_2 = 13x^2 + 18x + 120$. This gives us the system

$$\begin{bmatrix} 13 & 18 & 120 \\ 0 & 2020 & 2021 \\ 0 & 0 & 303 \end{bmatrix} \begin{bmatrix} \alpha_2 \\ \alpha_1 \\ \alpha_0 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}.$$

The equation

$$c_0 + c_1 x + c_2 x^2 = \alpha_0(303) + \alpha_1(2020x + 2021) + \alpha_2(13x^2 + 18x + 120)$$

is equvalent to writing

$$\begin{bmatrix} 13 & 18 & 120 \\ 0 & 2020 & 2021 \\ 0 & 0 & 303 \end{bmatrix} \begin{bmatrix} \alpha_2 \\ \alpha_1 \\ \alpha_0 \end{bmatrix} = \begin{bmatrix} c_2 \\ c_1 \\ c_0 \end{bmatrix}. \tag{4}$$

We compute

$$\begin{bmatrix} 13 & 18 & 120 \\ 0 & 2020 & 2021 \\ 0 & 0 & 303 \end{bmatrix}^{-1} = \begin{bmatrix} 1/13 & -9/13130 & -34337/1326130 \\ 0 & 1/2020 & -2021/612060 \\ 0 & 0 & 1/303 \end{bmatrix},$$

and it follows that (4) is equivalent to the "inverse" upper triangular system:

$$\begin{bmatrix} \alpha_2 \\ \alpha_1 \\ \alpha_0 \end{bmatrix} = \begin{bmatrix} 1/13 & -9/13130 & -34337/1326130 \\ 0 & 1/2020 & -2021/612060 \\ 0 & 0 & 1/303 \end{bmatrix} \begin{bmatrix} c_2 \\ c_1 \\ c_0 \end{bmatrix}.$$

Example 1.3. The formulas

$$\cos(2x) = 2\cos^2 x - 1, \quad \cos(4x) = 8\cos^4 x - 8\cos^2 + 1$$

can be written as

$$\begin{bmatrix} 8 & -8 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos^4 x \\ \cos^2 x \\ 1 \end{bmatrix} = \begin{bmatrix} \cos(4x) \\ \cos(2x) \\ 1 \end{bmatrix}.$$

Which is equivalent to writing

$$\begin{bmatrix} \cos^4 x \\ \cos^2 x \\ 1 \end{bmatrix} = \begin{bmatrix} 8 & -8 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \cos(4x) \\ \cos(2x) \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1/8 & 1/2 & 3/8 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(4x) \\ \cos(2x) \\ 1 \end{bmatrix}.$$

This gives rise to the formulas

$$\cos^2 x = (1/2)\cos(2x) + (1/2), \quad \cos^4 x = (1/8)\cos(4x) + (1/2)\cos(2x) + (3/8),$$

useful in integrating $\cos^2 x$ and $\cos^4 x$.

See the exercises for more examples of upper triangular "basis exchange."

1.2. **Divided Differences and the Lagrange Formula.** In class we showed that if $x_0 < x_1 < x_2$ are real, and if $f : \mathbb{R} \to \mathbb{R}$ is any function, then the unique polynomial

$$p(x) = c_0 + c_1 x + c_2 x^2$$

passing through the data points $(x_0, y_0), (x_1, y_1), (x_2, y_2)$ has

$$c_2 = \frac{y_0}{(x_0 - x_1)(x_0 - x_2)} + \frac{y_1}{(x_1 - x_0)(x_1 - x_2)} + \frac{y_2}{(x_2 - x_0)(x_2 - x_1)}. \tag{5}$$

We proved this by considering the Lagrange form of $p(x)$, namely

$$y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_0)},$$

which allows us to easily read off the x^2-coefficient to verify (5). We also showed by explicit computation that

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

$$= \frac{f(x_0)}{(x_0 - x_1)(x_0 - x_2)} + \frac{f(x_1)}{(x_1 - x_0)(x_1 - x_2)} + \frac{f(x_2)}{(x_2 - x_0)(x_2 - x_1)}. \tag{6}$$

Hence (6) and (5) are the same formula when identifying y_i with $f(x_i)$ for $i = 0, 1, 2$, and hence

$$c_2 = f[x_0, x_1, x_2],$$

and so

$$p(x) = c_0 + c_1 x + f[x_0, x_1, x_2]x^2$$

in this case. We can write similar—but more complicated—formulas for c_1 and for c_0 if we like, but we won't need these formulas.

We easily extend this to prove a similar formula for $f[x_0, \ldots, x_n]$ and to show that

$$c_n = f[x_0, \ldots, x_n]$$

where

$$p(x) = c_0 + c_1 x + \cdots + c_n x^n$$

is the unique polynomial such that $p(x_i) = f(x_i)$ for all $i = 0, \ldots, n$; the formula for $c_n = f[x_0, \ldots, x_n]$ is

(7)

$$f[x_0, \ldots, x_n] = \frac{f(x_0)}{(x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n)} + \cdots + \frac{f(x_n)}{(x_n - x_0)(x_n - x_1) \ldots (x_n - x_{n-1})}.$$

1.3. **Invariance Under Order.** The formula (7) for the divided difference $f[x_0, \ldots, x_n]$ shows that the divided difference does not depend on the order of $x_0, \ldots, x_n$.

This independence of the order is clear from the fact that $f[x_0, \ldots, x_n]$ is the x^n-coefficient in the interpolating polynomial. Let us illustrate this in case $n = 2$: if p is the unique polynomial of degree at most two such that

$$p(x_0) = f(x_0),\ p(x_1) = f(x_1),\ p(x_2) = f(x_2),$$

then one gets the same p if one requires that

$$p(x_1) = f(x_1),\ p(x_2) = f(x_2),\ p(x_0) = f(x_0),$$

since these three equalities can be written in any order. So as soon as one knows that $c_2 = f[x_0, x_1, x_2]$ in the expression $p(x) = c_0 + c_1 x + c_2 x^2$, it is clear that $f[x_0, x_1, x_2]$ must be independent of the order of x_0, x_1, x_2.

1.4. **End of Proof.** In class and the textbook we easily see that

$$p(x) = f[x_0] + f[x_1, x_0](x - x_0)$$

is the *secant line* of f at $x = x_0$ and $x = x_1$, i.e., $p(x)$ is the unique line $y = p(x)$ that intersects $y = f(x)$ at $x = x_0, x_1$. This is the formula (1) in the case $n = 1$; the case $n = 0$ of (1) is immediate.

Now we explain how to use the $n = 1$ case of (1) to prove the case where $n = 2$.

The upper triangular argument shows that if $q(x)$ is the unique polynomial of degree at most two such that $y = q(x)$ meets $y = f(x)$ at three distinct points $x = x_0, x_1, x_2$ (i.e., $q(x) = f(x)$ for $x = x_0, x_1, x_2$), then

$$q(x) = c_0 + c_1 x + c_2 x^2, \quad \text{with} \quad c_2 = f[x_0, x_1, x_2].$$

It follows that $q(x)$ and $f[x_0, x_1, x_2](x - x_0)(x - x_1)$ have the same x^2-coefficient, hence

$$r(x) = q(x) - f[x_0, x_1, x_2](x - x_0)(x - x_1)$$

is a polynomial of degree at most 1. However, since $(x-x_0)(x-x_1)$ equals 0 at $x=x_0$ and $x=x_1$, we have that $r(x)=q(x)=f(x)$ on both $x=x_0,x_1$. Hence

$$r(x)=f[x_0]+f[x_1,x_0](x-x_0)$$

(by the previous paragraph), and hence

$$q(x)=r(x)+f[x_0,x_1,x_2](x-x_0)(x-x_1)$$

$$=f[x_0]+f[x_1,x_0](x-x_0)+f[x_0,x_1,x_2](x-x_0)(x-x_1).$$

Now we know (1) holds for $n=2$; an argument similar to the previous paragraph shows that if (1) holds for $n=2$ then it also holds for $n=3$. Similarly, the case $n=3$ implies the case $n=4$, etc. More formally we can prove (1) for any n using *induction on* n.

Note that the fact that $r(x)$ above is a polynomial of degree at most 1 is intimately connected to the upper triangular relation between the two "bases"

$$1,x,x^2 \quad \text{and} \quad 1,x-x_0,(x-x_0)(x-x_1).$$

2. Adding One Data Point

Divided differences are useful in adding an additional interpolation point, a feature not found in the other two methods of Chapter 10, namely Section 10.2 (the standard basis) and 10.3 (the Lagrange formula). Let us give a concrete example.

Let $x_0=2$, $x_1=3$, and $x_2=5$ (to be very concrete), and let $p(x)$ be the unique functin c_0+c_1x that agrees with a function, f, at $x=2$ and $x=3$; therefore

$$p(x)=f[2]+f[2,3](x-2).$$

The unique polynomial

$$q(x)=\hat{c}_0+\hat{c}_1x+\hat{c}_2x^2$$

that agrees with f at $x=2,3,5$ is given by

$$q(x)=f[2]+f[2,3](x-2)+f[2,3,5](x-2)(x-3),$$

and hence

$$q(x)=p(x)+f[2,3,5](x-2)(x-3). \tag{8}$$

This means that if we have already computed $p(x)$, we can add the $x=5$ interpolation point and obtain $q(x)$ with the above formula.

Our other methods for finding $q(x)=\hat{c}_0+\hat{c}_1x+\hat{c}_2x^2$ that agrees with f on $x=2,3,5$ are

(1) to solve

$$\begin{bmatrix}1&2&4\\1&3&9\\1&5&25\end{bmatrix}\begin{bmatrix}\hat{c}_0\\\hat{c}_1\\\hat{c}_2\end{bmatrix}=\begin{bmatrix}f(2)\\f(3)\\f(5)\end{bmatrix},$$

or

(2) to write

$$q(x)=f(2)\frac{(x-3)(x-5)}{(2-3)(2-5)}+f(3)\frac{(x-2)(x-5)}{(3-2)(3-5)}+f(5)\frac{(x-2)(x-3)}{(5-2)(5-2)}.$$

Neither of these methods allow us to pass so easily from $p(x)$ to $q(x)$ as in (8). However, note that the third term in the Lagrangian

$$f(5)\frac{(x-2)(x-3)}{(5-2)(5-2)}$$

and the third term in Newton's formula

$$f[2,3,5](x-2)(x-3)$$

are related, as they are both multiples of $(x-2)(x-3)$.

3. The Generalized Mean-Value Theorem

Rolle's Theorem says that if $a<b$ are reals, and f is differentiable on (a,b) and continuous on $[a,b]$, then $f'(\xi)=0$ for some $\xi\in(a,b)$.

In particular, under the same assumptions, since there is a unique line $p(x)=c_0+c_1x$ such that $p(a)=f(a)$ and $p(b)-f(b)$, we may apply Rolle's Theorem to $g(x)=p(x)-f(x)$ and conclude that $0=g'(\xi)=p'(\xi)-f'(\xi)$ for some $\xi\in(a,b)$. For such a ξ we have

$$f'(\xi)=p'(\xi)=c_1,$$

and we easily see that

$$c_1=\frac{f(b)-f(a)}{b-a}=f[a,b].$$

Hence we conclude that

$$\frac{f(b)-f(a)}{b-a}=f[a,b]=f'(\xi)$$

for some $\xi\in(a,b)$, which is the usual Mean-Value Theorem of calculus.

Similarly, we see that if $x_0,\ldots,x_n$ are distinct real numbers, $f\colon\mathbb{R}\to\mathbb{R}$ is n-times differentiable, and p is the unique polynomial of degree at most n such that $p(x_i)=f(x_i)$ for $i=0,\ldots,n$, then by repeatedly applying Rolle's theorem (to $f(x)-p(x)$), we see that there is a $\xi\in\mathbb{R}$ (in the interval containing $x_0,\ldots,x_n$) at which the n-th derivative of $f(x)-p(x)$ is zero; at any such ξ we have

$$f^{(n)}(\xi)=p^{(n)}(\xi)=n!c_n$$

since $p(x)=c_0+c_1x+\cdots+c_nx^n$. Since $c_n=f[x_0,\ldots,x_n]$, it follows that for any such ξ,

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

This is just the theorem on "Divided Difference and Derivative" at the bottom of [A&G], p.312, Section 10.4. For this reason, this theorem is really a "generalized Mean-Value Theorem."

4. The Remainder Theorem for the Error in Polynomial Interpolation

In this section we use the Generalized Mean-Value Theorem above and one clever idea to prove a Remainder Theorem for the error in polynomial interpolation, given in Section 10.5 in [A&G]. After doing so we summarize Section 10.6 of [A&G].

Given distinct $x_0, \dots, x_n, x_{n+1} \in \mathbb{R}$ in an interval (a, b), and a function $f\colon (a, b) \to \mathbb{R}$ that is $(n+1)$-times differentiable, let $p_n(x)$ be the unique polynomial of at most degree n that agrees with f on $x_0, \dots, x_n$, and p_{n+1} the unique polynomial of degree at most $n+1$ that agrees with f on $x_0, \dots, x_n, x_{n+1}$. Then

$$p_{n+1}(x) - p_n(x) = (x - x_0) \dots (x - x_{n-1})(x - x_n) f[x_0, \dots, x_{n+1}],$$

which by the Generalized Mean-Value Theorem equals

$$(x - x_0) \dots (x - x_{n-1})(x - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

for some $\xi \in (a, b)$. Now take $x = x_{n+1}$ in the above formula (we regard this is a clever trick): we get

$$p_{n+1}(x_{n+1}) = p_n(x_{n+1}) + (x_{n+1} - x_0) \dots (x_{n+1} - x_{n-1})(x_{n+1} - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

But recall that $p_{n+1}(x)$ and $f(x)$ agree on $x = x_{n+1}$. Hence

$$f(x_{n+1}) = p_n(x_{n+1}) + (x_{n+1} - x_0) \dots (x_{n+1} - x_{n-1})(x_{n+1} - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

But since x_{n+1} is any real different from $x_0, \dots, x_n$, one can say that for any $x \in (a, b)$ there is a $\xi \in (a, b)$ such that

$$f(x) = p_n(x) + (x - x_0) \dots (x - x_{n-1})(x - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Of course, if x is not distinct from the $x_0, \dots, x_n$, i.e., for some i we have $x = x_i$, then the above formula holds automatically (for any ξ) since $f(x_i) = p(x_i)$ and the $(x_i - x_0) \dots (x_i - x_n) = 0$.

Section 10.6 of [A&G] makes the following point: imagine that $|f^{(n+1)}(\xi)|$ is bounded on (a, b) by M. Then the error in interpolation, for any $x \in (a, b)$, is bounded by

$$\bigl|f(x) - p_n(x)\bigr| \le \frac{M}{(n+1)!} \max_{x \in (a,b)} |x - x_0| \dots |x - x_n|. \tag{9}$$

Furthermore, by the remainder theorem, this inequality is not far from equality when $|f^{(n+1)}|$ is "close to" M throughout (a, b). So if we are able choose $x_0, \dots, x_n$ as we like, we might choose the $x_0, \dots, x_n$ so that

$$\max_{x \in (a,b)} |x - x_0| \dots |x - x_n|$$

is small as possible; this choice of $x_0, \dots, x_n$ are *Chebyshev points* for the interval (a, b). Section 10.6 explains more about such $x_0, \dots, x_n$.

5. The Newton Polynomial: A "Uniform" Formula in the Presence of Degeneracy

In this section we emphasize some points made in Section 10.7 of [A&G].

The real selling point of the Newton form of interpolation (1) for the unique polynomial p such that p and f agree "on all x_i" is that it is valid for all $x_0, \dots, x_n \in \mathbb{R}$—not merely x_i that are all distinct—provided that f is sufficiently differentiable. Furthermore, $f[x_0, \dots, x_m]$ is continous (differentiable, twice differentiable, etc.) in $x_0, \dots, x_n$ provided that f satisfies certain properties.

In other words, the Newton form of polynomials interpolation holds for any $x_0, \ldots, x_n$, and the divided differences $f[x_0, \ldots, x_m]$ express what happens in "degenerate cases" when some of the x_i are the same (or nearly the same). Let us explain what this means.

5.1. **A Degenerate Case of $x_0 = x_1$.** Let us consider some "degenerate limits" in interpolation from the point of view of Newton's formula.

First consider the case $x_0 = 2$, $x_1 = 2 + \epsilon$

$$p(x) = f(2) + (x - 2)f[2, 2 + \epsilon].$$

We have

$$\lim_{\epsilon \to 0} f[2, 2 + \epsilon] = \lim_{\epsilon \to 0} \frac{f(2 + \epsilon) - f(2)}{\epsilon} = f'(2)$$

assuming the derivative $f'(2)$ exists. For this reason it is natural to define

$$f[2, 2] \stackrel{\text{def}}{=} f'(2),$$

when $f'(2)$ exists; then $\epsilon \to 0$ gives the formula

$$p(x) = f(2) + (x - 2)f[2, 2],$$

and the limiting interpolating (linear) polynomial is

$$p(x) = f(2) + (x - 2)f[2, 2] = f(2) + (x - 2)f'(2), \tag{10}$$

which is the familiar tangent line of f at $x = 2$.

5.2. **Agreement to Higher Order.** Note that in (10) we have that $p(x)$ is the tangent line to $f(x)$ at $x = 2$; hence we conclude

$$p(2) = f(2),\ p'(2) = f'(2) = f[2, 2]$$

(which we can also conclude by differentiating $p(x)$), and so we say that p and f *agree to order two* at $x = 2$. More generally, for $k = 1, 2, \ldots$ we say that two functions g, f *agree to order* k at $x = a$ if

$$g(a) = f(a),\ g'(a) = f'(a),\ \ldots, g^{(k-1)}(a) = f^{(k-1)}(a),$$

i.e., if $g - f$ and its first $k - 1$ derivatives vanish at $x = a$ (assuming that all these derivatives exist).

5.3. **Another $x_0 = x_1$ Degenerate Case.** Next consider the case $x_0 = 2$, $x_1 = 2 + \epsilon$, and $x_2 = 3$ in Newton's polynomial, where ϵ is a real number:

$$p(x) = f(2) + (x - 2)f[2, 2 + \epsilon] + (x - 2)\big(x - (2 + \epsilon)\big)f[2, 2 + \epsilon, 5].$$

Taking $\epsilon \to 0$ gives the formula

$$p(x) = f(2) + (x - 2)f[2, 2] + (x - 2)^2 f[2, 2, 5],$$

provided that we define

$$f[2, 2, 5] \stackrel{\text{def}}{=} \lim_{\epsilon \to 0} f[2, 2 + \epsilon, 5]$$

and this limit exists (and that we define $f[2, 2]$ as $f'(2)$). Unlike the situation in the previous subsection with $f[2, 2]$, the question of whether or not the limit

$$\lim_{\epsilon \to 0} f[2, 2 + \epsilon, 5]$$

exists is more subtle; however, since $2, 2+\epsilon, 5$ are distinct reals for small $\epsilon \neq 0$, we have

$$\lim_{\epsilon\to 0} f[2, 2+\epsilon, 5] = \lim_{\epsilon\to 0} \frac{f[2,5] - f[2, 2+\epsilon]}{5 - (2+\epsilon)}$$

(using the symmetry of $f[x_0, x_1, x_2]$ under permuting the x_0, x_1, x_2), so if $f[2,2] = f'(2)$ exists, we have

$$\lim_{\epsilon\to 0} \frac{f[2,5] - f[2, 2+\epsilon]}{5 - (2+\epsilon)} = \frac{f[2,5] - f[2,2]}{3}.$$

More generally, if $x_1 = x_0$ but $x_2 \neq x_0$, and if f is differentiable at $x = x_0$, then $f[x_0, x_0]$ and $f[x_0, x_0, x_2]$ both exist and

$$f[x_0, x_0, x_2] = \frac{f[x_0, x_2] - f[x_0, x_0]}{x_2 - x_0}.$$

Hence the formula

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$$

also holds when $x_0 = x_1$ provided that $f'(x_0)$ exists.

Note that is this case we have a limiting Newton polynomial, $p(x)$, given by

$$p(x) = f[2] + (x-2)f[2,2] + (x-2)^2 f[2,2,5].$$

This shows that

$$p(2) = f(2),\ p'(2) = f'(2),$$

and hence again p, f agree to order two at $x = 2$.

This situation and the one in the previous subsection are "degenerate" cases $x_0 = x_1 = 2$, where the value 2 occurs twice among the $x_0, \ldots, x_n$. This results in p, f agreeing to order two at $x = 2$. This is how we generally interpret degenerate cases of interpolation, where we allow some of the $x_0, \ldots, x_n$ to be the same, and we accordingly get higher order agreement on the x_i that are repeated. This is spelled out at the bottom of page 319 (Section 10.7) of [A&G].

5.4. **Multiple Roots and Multiple Argreement.** Another way to understand agreement to multiple orders is via multiple roots in polynomials, which you have likely already seen somewhere.

The polynomial

$$p(x) = (x-1)(x-5)^2(x-7)^3$$

is said to have a *simple root* at $x = 1$, a *double root* at $x = 5$, and a *triple root* at $x = 7$. Some computation shows that

$$p(x) = 2(x-5)(x-7)^2(3x^2 - 23x + 32),$$

which indicates the general principle that if p has a double root (respectively, triple root, etc.) at $x = a$, then p' has a single root (respectively, double root, etc.) at $x = a$. More genearlly, whenever

$$p(x) = (x-7)^3 q(x)$$

for another polynomial $q(x)$, the product rule shows that

$$p'(x) = 3(x-7)^2 q(x) + (x-7)^3 q'(x),$$

and so $p'(x)$ is necessarily divisible by $(x-7)^2$; hence if $p(x)$ has a root or zero of order 3 at $x = 7$, then $p'(x)$ must have a root or zero of order at least 2 at $x = 7$.

More generally, if $f\colon \mathbb{R} \to \mathbb{R}$ is n-times differentiable, and $k < n$ is an integer, we say that $f = f(x)$ has a *root (or zero)* of *order at least* k *at* $x = a$ if

$$f(a) = f'(a) = \cdots = f^{(k-1)}(a) = 0,$$

and of *order exactly* k if, moreover, $f^{(k)}(a) \neq 0$. It follows from this definition that if f has a zero of order at least k (respectively, exactly k) at $x = a$, then f' has a zero of order at least $k-1$ (respectively, exactly $k-1$).

It follows that if two functions g, f agree at $x = a$ to some order k, then it is equivalent to say that $g - f$ has a zero at $x = a$ to order k.

5.5. **Rolle's Theorem for Multiple Agreement.** Rolle's theorem implies that if f has $n+1$ roots on some interval, then f' has n roots on this interval, and f'' has $n-1$ roots on this interval, etc., assuming that f has enough derivatives.

We can also prove a Rolle's theorem for multiple agreement; it is easiest to understand this by an example: if a function f has a zero of order 10 at $x = 1$ and a zero of order 20 at $x = 2$, then "counting mulitiplicites" we say that f has at least $10 + 20 = 30$ zeros. Rolle's theorem implies that $f'(\xi) = 0$ for some ξ with $1 < \xi < 2$; we also know that f' will have a zero of order at least 9 at $x = 1$ and at least 19 at $x = 2$; this which gives $1 + 9 + 19 = 29$ zeros of f. Hence the 30 "zeros counted with multiplicty" of f on $[1, 2]$ implies that f' has at least 29 zeros counted with multiplicity on $[1, 2]$.

In this fashion, counting intermediate zeros of f' along with guaranteed zeros of f due to multiplicity, we can prove that if f is differentiable on some interval and has $n+1$ zeros there (counted with mulitplicity), then f' has at least n zero there (counted with multiplicity), and f'' at least $n-1$, etc.

5.6. **General Interpolation.** Say that f is a differentiable function, and we seek a polynomial

$$p(x) = c_0 + c_1 x + c_2 x^2$$

such that p that agrees with f on the points x_0, x_1, x_2 where $x_0 = x_1 = 5$ and $x_2 = 8$: we interpret this problem is that we want

$$p(5) = f(5),\ p'(5) = f'(5),\ p(8) = f(8),$$

since the value 5 occurs twice among the x_0, x_1, x_2. Since

$$p'(x) = c_1 + 2c_2 x,$$

and hence

$$p'(5) = c_1 + 10c_2,$$

the above problem amounts to solving the system

$$\begin{bmatrix} 1 & 5 & 25 \\ 0 & 1 & 10 \\ 1 & 8 & 64 \end{bmatrix} \begin{bmatrix} c_2 \\ c_1 \\ c_0 \end{bmatrix} = \begin{bmatrix} f(5) \\ f'(5) \\ f(8) \end{bmatrix}.$$

We can prove that this system has a unique solution by modifying the proof that the interpolation problem with x_0, x_1, x_2 distinct has a unique solution: namely, the homogeneous system is

$$\begin{bmatrix} 1 & 5 & 25 \\ 0 & 1 & 10 \\ 1 & 8 & 64 \end{bmatrix} \begin{bmatrix} c_2 \\ c_1 \\ c_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

and any solution c_2, c_1, c_0 yields a polynomial $p(x) = c_0 + c_1 x + c^2 x^2$ that has a double zero at $x = 5$ and a single zero at $x = 8$. This implies that $p(x)$, if nonzero, must be divisible by $(x-5)^2(x-8)$, which is impossible since p is of degree at most 2. [One could also prove that $p(x)$ must be zero using our generalized Rolle's Theorem.] Hence the only solution to the homogeneous system is $c_0 = c_1 = c_2 = 0$. Hence any non-homogeneous form of this system has a unique solution.

5.7. **Taylor Series.** If $f\colon \mathbb{R} \to \mathbb{R}$ is n-times differentiable near a point $x = a$, then as $x_0, \ldots, x_n$ all tend to a, the Mean-Value theorem implies that

$$f\big[\underbrace{a, \ldots, a}_{k \text{ times}}\big] \stackrel{\text{def}}{=} \lim_{x_0,\ldots,x_n \to a} f[x_0, \ldots, x_n] = \frac{f^{(k)}(a)}{k!},$$

provided that f is k-times differentiable near $x = a$ and its k-derivative is continuous at $x = a$. In this way (1), in the case

$$x_0 = x_1 = \ldots = x_n = a$$

becomes the polynomial

$$p(x) = f(a) + (x-a)f'(a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

which is Taylor's theorem, we know agrees with f "up to order $n+1$" by Taylor's theorem. Furthermore, the error in the Taylor expansion is given by the "Remainder Term" in Taylor's theorem,

$$p(x) - f(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

for some ξ between a and x, which is a special case of the "Error in Polynomial Interpolation" formula.

6. Divided Differences: What We Can Prove and What We "Sweep Under the Rug"

Let us briefly comment on what we are "sweeping under the rug" (i.e., avoiding) in CPSC 303 regarding divided differences; we will complement this by stating some theorems. The real question is how does $f[x_0, \ldots, x_n]$ behave as a function of $(x_0, \ldots, x_n) \in \mathbb{R}$ (including those points where some of the x_i are equal): this is both a "selling point" of divided differences, but also a subtle issue.

6.1. **The Divided Difference $f[x_0, x_1]$.** We have already mentioned that

$$\lim_{\epsilon \to 0} f[2, 2+\epsilon]$$

exists and equals $f'(2)$. However, it is not generally true that

$$\lim_{x_0, x_1 \to 2} f[x_0, x_1]$$

exists even if $f'(2)$ exists: indeed, $f[x_0, x_1]$ is the slope of the secant line of f at $x = x_0$ and $x = x_1$; it is not hard to see that if f' is discontinuous at $x = 2$,[1] then

[1] Here is a standard example of a function, f, whose derivative exists everywhere but is discontinuous at $x = 2$: if we define $f(2) = 0$, and for $x \neq 2$ we define

$$f(x) = (x-2)^2 \sin\big(1/(x-2)^2\big),$$

the limit

$$\lim_{x_0,x_1\to 2} f[x_0,x_1]$$

does not exist. In this case it impossible to define $f[x_0,x_1]$ in a way that makes it a continuous function at $x_0=x_1=2$ (although we generally define $f[2,2]=f'(2)$ for reasons mentioned before).

However, the optimistic side to this secant line consideration is the following easy result.

Theorem 6.1. *If $f\colon\mathbb{R}\to\mathbb{R}$ is differentiable (on all of $\mathbb{R}$), then define $f[x_0,x_0]$ to be $f'(x_0)$ for all $x_0\in\mathbb{R}$. If f' is continuous (on all of $\mathbb{R}$), then $f[x_0,x_1]$ is continuous (at all $(x_0,x_1)\in\mathbb{R}^2$).*

(This theorem speaks of *continuity* of functions on $\mathbb{R}^2$; this knowledge is not a prerequisite for CPSC 303, and hence I will briefly explain this concept when/if we cover in in CPSC 303.)

Proof. It suffices to fix $(a_0,a_1)\in\mathbb{R}^2$ and to show that

$$\lim_{(x_0,x_1)\to(a_0,a_1)} f[x_0,x_1]=f[a_0,a_1].$$

If $a_0\neq a_1$, then for (x_0,x_1) sufficiently close to (a_0,a_1) we have $x_0\neq x_1$, and hence

$$\lim_{(x_0,x_1)\to(a_0,a_1)} f[x_0,x_1]=\lim_{(x_0,x_1)\to(a_0,a_1)}\frac{f(x_1)-f(x_0)}{x_1-x_0}=\frac{f(a_1)-f(a_0)}{a_1-a_0}=f[a_0,a_1].$$

Otherwise $a_0=a_1$; for $x_0=x_1$ we have

$$f[x_0,x_1]=f[x_0,x_0]=f'(x_0),$$

and for $x_0\neq x_1$, the Mean-Value theorem implies

$$f[x_0,x_1]=f'(\xi)$$

for some ξ between x_0 and x_1. It follows that

$$\bigl|f[x_0,x_1]-f[a_0,a_0]\bigr|\le\bigl|f'(\xi)-f'(a_0)\bigr|,$$

and so for $|x_0-a_0|\le\epsilon$ and $|x_1-a_0|\le\epsilon$ we have

$$\bigl|f[x_0,x_1]-f[a_0,a_0]\bigr|\le\max_{|\xi-a_0|\le\epsilon}\bigl|f'(\xi)-f'(a_0)\bigr|.$$

Since f' is continuous, it follows that

$$\lim_{(x_0,x_1)\to(a_0,a_1)} f[x_0,x_1]=f[a_0,a_1].$$

□

---

then we have $f'(2)=0$ (essentially because f is bounded above by $(x-2)^2$ and below by $-(x-2)^2$, and these two functions osculate at $x=2$ (i.e., $\pm(x-2)^2$ agree to order two at $x=2$). On the other hand, for $x\neq 2$ we have

$$f'(x)=2(x-2)\sin\bigl(1/(x-2)^2\bigr)+\frac{-2}{x-2}\cos\bigl(1/(x-2)^2\bigr),$$

which is not even bounded as $x\to 2$.

Ideally one wants that $f[x_0,\ldots,x_n]$ not merely be continuous in $x_0,\ldots,x_n$, but also differentiable, or infinitly-differentiable, etc. At present I don't know a reference where such issues are studied in a simple fashion. Such issues are discussed starting in Section 7 of de Boor's survey, "Divided Differences" (available at https://arxiv.org/abs/math/0502036) which I recommend. This survey is more technical than [A&G] and requires some math on the level of UBC's Math 320: for example, you need to know that if $M\colon X\to\mathbb{R}^{n\times n}$ is a continuous map from a topological space, X (e.g., $X=\mathbb{R}^m$ for some m), to the space of real $n\times n$ matrices, then if $M(x)$ is invertible for all $x\in X$, then the map

$$x\mapsto \bigl(M(x)\bigr)^{-1}\in\mathbb{R}^{n\times n}$$

is also continuous for all x (in view of the formula $M^{-1}(I-A)^{-1}=M^{-1}(I+A+A^2+\cdots)$ for $\|A\|<1$ in any matrix norm).

Exercises

(1) Describe does the following MATLAB code does:

```
clear
i = -10:1:15
x = i/10
y = x.*x
z = x.^3
xpi = x * pi
f = sin(x * pi)
t = -1 : 0.1 : 1
```

and explain or summarize the error message(s) that you get when you type

```
x*x
x^3
```

(i.e., don't just copy the error message down word for word.)

(2) In this exercise we consider

$$f(x)=\sin(x)$$

Taylor's theorem with remainder implies that for every $x\in[-1,1]$ we have (the Taylor expansion)

$$f(x)=\sin(x)=x-x^3/3!+x^5/5!-R_7(x),$$

where $R_7(x)$ is a function of x such that for every $x\in[-1,1]$ there is a $\xi\in[-1,1]$ such that

$$R_7(x)=x^7\cos(\xi)/7!$$

(a) Explain why for any $x\in[-1,1]$ we have

$$|R_7(x)|\le 1/7!=.00019841\ldots$$

(b) What is the largest value of $|R_7(x)|$ where $R_7(x)$ is given as above, with

$$R_7(x) = \sin(x) - x + x^3/3! - x^5/5!$$

for $x = i/1000$ and $i = -1000, \ldots, 1000$? Check this by running the MATLAB code:

```
clear
x = ( -1000:1000 ) / 1000
abs_r7=abs( sin(x)-x+x.^3/6-x.^5/120)
max( abs_r7 )
```

How close is this maximum absolute value of the error to the upper bound on $|R_7(x)|$ in part (a)?

(3) The *Chebyshev points* on $(-1,1)$ are defined (see Section 10.6 of [A&G]) for each positive integer n as the $n+1$ points $x_0, \ldots, x_n$.

$$x_i = \cos\left(\frac{(2i+1)\pi}{2(n+1)}\right), \quad i = 0, 1, \ldots, n$$

(a) Generate the $n = 5$ values of $x_0, \ldots, x_5$ using the code:

```
clear
vi = 0:5
cheb = cos( ( 2 * vi + 1 ) * pi / 12 )
```

(b) For $x_0, \ldots, x_5$ being the Chebyshev points above (with $n = 5$), find the maximum absolute value of

$$v(x) = (x - x_0)(x - x_1) \ldots (x - x_5)$$

for $x = i/1000$ with $i = -1000, -999, \ldots, 999, 1000$ using MATLAB. You could do this by adding the code

```
v = 1:2001
for i= 1 : 2001 ,  v(i) = prod( (i-1001)/1000 -cheb); end
max(abs(v))
```

(c) Based on (1), if we interpolate $f(x) = \sin(x)$ at $x_0, \ldots, x_5$, show that the error in this interpolation at any $x \in \mathbb{R}$ is at most

$$\left|\frac{(x - x_0)(x - x_1) \ldots (x - x_5)}{6!}\right|$$

in absolute value. [Hint: $f^{(6)}(x) = -\sin(x)$.]

(d) Use the bound in part (c) and the experiment in part (b), give an upper bound on the largest error in interpolation for $x = i/1000$ with $i = -1000, -999, \ldots, 1000$.

(e) Then interpoloate $\sin(x)$ at the Chebyshev point $x_0, \ldots, x_5$, and find the error in interpolation over all $x = i/1000$ with $i = -1000, -999, \ldots, 1000$.

```
sin_cheb = sin(cheb)
p = polyfit( cheb, sin_cheb, 5)    % this returns the coefficients c0,...,c5

x = -1 : 0.001 : 1
y = polyval(p,x)

sinx = sin(x)
max( abs( y - sinx ) )
```

(4) In this exercise we will specify 6 real numbers $x_0, \ldots, x_5$ and consider the largest absolute value of the polynomial

$$v(x) = (x - x_0)(x - x_1) \ldots (x - x_5)$$

over all $x \in [-1, 1]$ (i.e., all $x \in \mathbb{R}$ with $-1 \le x \le 1$).

(a) If $x_0 = x_1 = \ldots = x_5 = 0$, at which $x \in [-1, 1]$ does $v(x)$ attain its maximum value, and what is this value? **Just give the answer; it should be clear once you compute** $v(x)$**.** [Hint: In this case, $v(x) = x^6$.]

(b) Let $x_0 = x_1 = -1$, $x_2 = x_3 = 0$, and $x_4 = x_5 = 1$. Using calculus, find the (exact) value(s) of $x \in [-1, 1]$ at which $v(x)$ attains its maximum absolute value. [Hint: $v(x) = (x^3 - x)^2$; you need to check $v(x)$ at the endpoints ± 1, and then check the value of $v(x)$ for the values of x where $v'(x) = 0$.] What is the maximum absolute value of $v(x)$, both exactly and as a decimal to 4 digits?

(c) If $x_0, \ldots, x_5$ are the $n = 5$ Chebyshev points of Problem 3, approximate the maximum absolute value of $v(x)$ by checking the values $x = i/1000$ and $i = -1000, -999, \ldots, 1000$. [If you have done Problem 3 above, then you have already found this value.]

(d) How close is your value in part (c) to the $1/32$? Type `max(abs(v)) - 1/32` to find out. [In Section 10.6 we will learn that the maximum absolute value of $v(x)$ over all $x = [-1, 1]$ is (in exact arithmetic) $1/32$.]

(e) By what factor is the maximum absolute value in part (b) an improvement over part (a)? And, similarly, for part (c) over part (b)?

(5) More exercises to follow.

Department of Computer Science, University of British Columbia, Vancouver, BC V6T 1Z4, CANADA.

E-mail address: jf@cs.ubc.ca

URL: http://www.cs.ubc.ca/~jf