1. Polynomial interpolation
2. Linear system solves interpolation
3. Lagrange form
4. Divided differences
5. Interpolation error
6. Rate of convergence
7. Estimating derivatives
8. Quadrature rules & errors
9. Numerically estimating order of convergence
10. Richardson extrapolation
11. MATLAB/octave demos

1 Polynomial interpolation

Given a set of points in the plane such that all the x-values are distinct, how can we find a polynomial of smallest degree that interpolates each of those data points? In other words, given a set of points \((x_i, y_i) \in \mathbb{R}^2\) for \(i = 0, \ldots, n\), how can we find a polynomial \(P : \mathbb{R} \rightarrow \mathbb{R}\) of smallest degree such that \(P(x_i) = y_i\) for each \(i = 0, \ldots, n\)?

2 Solving a linear system to obtain the interpolating polynomial

One way is to use linear algebra: set \(P(x) = a_0 + a_1 x + \cdots + a_n x^n\), then use the interpolation criteria to generate the \((n + 1) \times (n + 1)\) system of equations

\[
\begin{align*}
P(x_0) &= a_0 + x_0 a_1 + \cdots + x_0^n a_n = y_0 \\
P(x_1) &= a_0 + x_1 a_1 + \cdots + x_1^n a_n = y_1 \\
&\vdots \\
P(x_n) &= a_0 + x_n a_1 + \cdots + x_n^n a_n = y_n.
\end{align*}
\]
which is the same as solving the system

\[
\begin{pmatrix}
1 & x_0 & \cdots & x_0^n \\
1 & x_1 & \cdots & x_1^n \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_n & \cdots & x_n^n \\
\end{pmatrix}
\begin{pmatrix}
a_0 \\
a_1 \\
\vdots \\
a_n \\
\end{pmatrix}
= 
\begin{pmatrix}
y_0 \\
y_1 \\
\vdots \\
y_n \\
\end{pmatrix}.
\]

If the matrix on the left were NOT invertible, then there would be \(c_0, c_1, \ldots, c_n \in \mathbb{R}\), not all 0, such that

\[
\begin{pmatrix}
1 & x_0 & \cdots & x_0^n \\
1 & x_1 & \cdots & x_1^n \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_n & \cdots & x_n^n \\
\end{pmatrix}
\begin{pmatrix}
c_0 \\
c_1 \\
\vdots \\
c_n \\
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0 \\
\end{pmatrix}.
\]

Then we could form the polynomial \(g(x) = c_0 + c_1 x + \cdots + c_n x^n\), which has \(x_0, x_1, \ldots, x_n\) as roots. Thus \(g\) has at least \(n + 1\) distinct roots, which implies via Rolle’s theorem that \(g'\) has at least \(n\) distinct roots, \(\ldots, g^{(n)}\) has a root. So \(g^{(k)}\) has at least \(n + 1 - k > 0\) roots for \(k = 0, 1, \ldots, n\). Since \(g^{(n)} = n!c_n\), we get that \(c_n = 0\). Then \(g^{(n-1)} = (n - 1)!c_{n-1}\), so \(c_{n-1}\) must be 0 as well. Continue in this process until we’ve shown that every last \(c_i\) is 0, arriving at a contradiction. (An alternate way to deduce the contradiction would be to use the Euclidean Algorithm to say that \((x - x_i)\) must be a factor of \(g\) for each \(i = 0, 1, \ldots, n\), contradicting that \(g\) is of degree \(\leq n\)).

Thus the matrix in question is invertible, and our polynomial interpolation problem is solved by solving this linear system, which must admit exactly one (unique) solution.

### 3 Lagrange Form of the interpolating polynomial

This analysis reassures us that there exists a unique solution, but it doesn’t offer us a convenient way to compute the coefficients of that interpolating polynomial. In fact, there’s a very nice description of the interpolating polynomial that utilizes so-called “Lagrange interpolants.”

Suppose \(f : [a, b] \to \mathbb{R}\) is some function we are trying to interpolate with a polynomial of degree \(\leq n\) through the \(n + 1\) points \((x_0, f(x_0)), \ldots, (x_n, f(x_n))\) (with distinct \(x_i\’s\)). This is accomplished by the polynomial

\[
p_n(x) = \sum_{i=0}^{n} \left(f(x_i) \prod_{\substack{j=0 \atop j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}\right),
\]

since we may view the product \(\prod_{\substack{j=0 \atop j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}\) as a switch that turns on (equals 1) when \(x = x_i\) and turns off (equals 0) when \(x = x_j, j \neq i\). By our earlier analysis, we know that this is the unique polynomial that interpolates \(f\) at these points. This representation of the interpolating polynomial is the **Lagrange form** of the interpolating polynomial.
4 Divided differences characterization of the interpolating polynomial

So far we have developed two ways to construct interpolating polynomials. It will behoove us to develop one more way, that of divided differences. The benefit of using divided differences to construct interpolating polynomials is threefold: it enables us to quickly compute the interpolating polynomial by hand, it produces an efficient form of the polynomial for a computer to evaluate, and it allows us to derive the error between a function and its interpolating polynomial in a natural way.

Define $f[x_i] = f(x_i)$. Then $p_0(x) = f(x_i) = f[x_i]$ (where the first equality follows from, for example, the Lagrange form of the 0th order interpolating polynomial of $f$ through $x_i$). Let $\phi: \{0, \ldots, n+1\} \to \{0, \ldots, n+1\}$ be a bijection (in other words, a permutation of $\{0, \ldots, n+1\}$).

Then for $n \geq 0$ define $f[x_0, \ldots, x_{n+1}]$ by

$$f[x_0, \ldots, x_{n+1}] = \frac{f(x_{n+1}) - p_n(x_{n+1})}{\prod_{i=0}^{n} (x_{n+1} - x_i)},$$

where $p_n$ is the degree $\leq n$ polynomial that interpolates $f$ at $x_0, \ldots, x_n$.

This definition ensures that

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

$$= \sum_{i=0}^{n} \left( f[x_0, \ldots, x_i] \prod_{j=0}^{i-1} (x-x_j) \right)$$

is in fact another representation of the interpolating polynomial of degree $\leq n$ for $f$ through $x_0, \ldots, x_n$ (valid for $n = 0, 1, 2, \ldots$). This is the divided difference representation of the interpolating polynomial.

Note that the leading coefficient in the above representation of $p_n$ is $f[x_0, \ldots, x_n]$. But by considering the Lagrange form, we know that the leading coefficient can also be expressed as

$$\sum_{i=0}^{n} \left( \frac{f(x_i) \prod_{j=0}^{n} 1}{x_i - x_j} \right),$$

which gives us another way to compute divided differences:

$$f[x_0, \ldots, x_n] = \sum_{i=0}^{n} \left( f(x_i) \prod_{j=0}^{n} \frac{1}{x_i - x_j} \right),$$

which in turn tells us that divided differences are independent of the order of their inputs.

Let $n \geq 0$, and let $r_n$, $q_n$ be the unique polynomials of degree $\leq n$ that interpolate
f at \(x_1, \ldots, x_{n+1}\) and \(x_0, \ldots, x_n\), respectively. Then the degree \(\leq n+1\) polynomial that interpolates \(f\) at \(x_0, \ldots, x_{n+1}\) is given by

\[ p_{n+1}(x) = \frac{(x-x_0)r_n(x)}{x_{n+1}-x_0} + \frac{(x-x_{n+1})q_n(x)}{x_0-x_{n+1}}. \]

Since \(r_n\) has leading coefficient \(f[x_1, \ldots, x_{n+1}]\) and \(q_n\) has leading coefficient \(f[x_0, \ldots, x_n]\), we get that \(p_{n+1}\) has leading coefficient

\[ \frac{f[x_1, \ldots, x_{n+1}] - f[x_0, \ldots, x_n]}{x_{n+1}-x_0}. \]

Also, the leading coefficient of \(p_{n+1}\) is \(f[x_0, \ldots, x_{n+1}]\). This yields a convenient recursive formula for divided differences, namely,

\[ f[x_0, \ldots, x_{n+1}] = \frac{f[x_1, \ldots, x_{n+1}] - f[x_0, \ldots, x_n]}{x_{n+1}-x_0}. \]

Furthermore, since we know that divided differences are independent of the order of their inputs, we get the following formula for \(i \neq j\):

\[ f[x_0, \ldots, x_{n+1}] = \frac{f[x_0, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{n+1}] - f[x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}]}{x_j-x_i}. \]

This last formula allows us to quickly compute the coefficients of interpolating polynomials.

Note that all of the \(x_i\)'s are assumed to be distinct in order for the divided difference formulas to make sense. But if we assume that \(f\) has one derivative (over some interval containing all the interpolating points), then we can extend the definition of divided differences to the case that two of the inputs are identical inputs. This will be useful later on when we derive errors associated with certain quadrature rules.

Suppose \(f\) is differentiable on the interval \([a, b]\) (with one-sided derivatives at the endpoints), and all points of interpolation lie in \([a, b]\). In the case \(n = 1\), we get \(f[x_0, x] = f[x] - f[x_0] \to f'(x_0)\) as \(x \to x_0\). If \(x^* \neq x_0\), then

\[ \lim_{x \to x^*} f[x_0, x] = \frac{f[x^*] - f[x_0]}{x^* - x_0} = f[x_0, x^*] \]

by limit laws and continuity of \(f\). So \(x \mapsto f[x_0, x]\) is continuous on \([a, b]\). Then inductively assuming that \(k \geq 0\) and that the function \(y \mapsto g[y_0, \ldots, y_k, y]\) is continuous on \([a, b]\) for any differentiable \(g : [a, b] \to \mathbb{R}\) and any collection of distinct interpolation nodes \(y_0, \ldots, y_k \in [a, b]\), we get that for all \(x^* \in [a, b]\):

\[
\lim_{x \to x^*} f[x_0, \ldots, x_{k+1}, x] = \frac{\lim_{x \to x^*} f[x_1, \ldots, x_{k+1}, x] - f[x_0, \ldots, x_k, x]}{x_{k+1} - x_0} = \frac{\lim_{x \to x^*} f[x_1, \ldots, x_{k+1}, x] - \lim_{x \to x^*} f[x_0, \ldots, x_k, x]}{x_{k+1} - x_0} = \frac{f[x_1, \ldots, x_{k+1}, x^*] - f[x_0, \ldots, x_k, x^*]}{x_{k+1} - x_0},
\]

4
proving that \( x \mapsto f[x_0, \ldots, x_{k+1}, x] \) can be continuously extended.

Next we’ll demonstrate the simple relationship between divided differences and derivatives. Suppose \( f : [a, b] \to \mathbb{R} \) has \( n \) derivatives \((n \geq 0)\), \( x_0, \ldots, x_n \in [a, b] \) are distinct, and \( p_n \) is the degree \( \leq n \) interpolating polynomial for \( f \) through the \( x_i \)’s. Put

\[
g : [a, b] \to \mathbb{R}, \quad g(x) := f(x) - p_n(x).
\]

Note that \( g(x_i) = 0 \) for \( i = 0, \ldots, x_n \). Then by Rolle’s theorem, \( g' \) has at least \( n \) distinct zeros in \((a, b)\). Then again by Rolle’s theorem, \( g'' \) has at least \( n - 1 \) distinct zeros, \ldots, and \( g^{(n)} \) has at least one zero, say \( \xi \in (a, b) \). Since the leading coefficient of \( p_n \) is \( f[x_0, \ldots, x_n] \), we know that \( p_n^{(n)}(x) \equiv f[x_0, \ldots, x_n]n! \), which then yields

\[
0 = g^{(n)}(\xi) = f^{(n)}(\xi) - f[x_0, \ldots, x_n]n!,
\]

or simply

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

### 5 Interpolation error

Now that we’ve laboriously proved some facts about divided differences, we may eat the fruits of our labor.

Let \( f : [a, b] \to \mathbb{R} \) be \((n+1)\)th-order differentiable, let \( x_0, \ldots, x_n \in [a, b] \) be distinct, and let \( p_n \) be the degree \( \leq n \) interpolating polynomial for \( f \) through the \( x_i \)’s. Let’s temporarily view \( x \in [a, b] \) as fixed. If \( x = x_i \) for some \( i = 0, \ldots, n \), then \( p_n(x) = f(x) \), so the error in approximating \( f(x) \) by \( p_n(x) \) is 0 in this case. Otherwise, let \( p_{n+1} \) be the degree \( \leq n+1 \) polynomial through the distinct points \( x_0, \ldots, x_n, x \). Then

\[
p_{n+1}(t) = p_n(t) + f[x_0, \ldots, x_n, x](t - x_0) \cdots (t - x_n),
\]

and since \( p_{n+1}(x) = f(x) \) (since \( x \) is one of the interpolation nodes), we get that

\[
f(x) = p_n(x) + f[x_0, \ldots, x_n, x](x - x_0) \cdots (x - x_n).
\]

Now using the formula relating divided differences to derivatives that we derived in the previous section, we have the relation

\[
f(x) = p_n(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0) \cdots (x - x_n)
\]

for some \( \xi \in (a, b) \). The last two equations comprise the interpolation error formulas.
In the case of equispaced interpolation nodes, i.e. \( n \geq 1, \quad h = \frac{b - a}{n}, \quad x_i = a + ih \) for \( i = 0, \ldots, n \), we have a nice formula for the error. First note that
\[
|f(x) - p_n(x)| \leq \frac{||f^{(n+1)}||_{[a,b]}}{(n+1)!} \prod_{j=0}^{n} |x - x_j|.
\]

We will try to get a better handle on the last product, since we can’t control \( f^{(n+1)} \).
\( \prod_{j=0}^{n} |x - x_j| \) is largest when \( x \) is in either the leftmost or rightmost interval \([x_j, x_{j+1}]\), since the distances \(|x - x_j|\) only grow as we proceed from the end with \( x \) to the other (a fact which is pictorally obvious, but tedious to prove rigorously ...). This phenomenon (that interpolation errors tend to be largest in the extremal brackets for equispaced interpolation) is encapsulated in Runge’s phenomenon, though that also details how the error actually grows with \( n \). WLOG assume \( x \in [x_0, x_1] \) to obtain an upper bound on \( \prod_{j=0}^{n} |x - x_j| \). Then for \( j > 1, \quad |x - x_j| = x_j - x \leq x_j - x_0 = h_j \), so
\[
\prod_{j=0}^{n} |x - x_j| = |x - x_0||x - x_1| \prod_{j=2}^{n} |x - x_j|
\]
\[
\leq |x - x_0||x - x_1| \prod_{j=2}^{n} h_j
\]
\[
= |x - x_0||x - x_1| \cdot h^{n-1} n!.
\]

Now we can use single-variable calculus methods to find the max that the function \( g(x) = |x - x_0||x - x_1| = (x - x_0)(x_1 - x) \) achieves on \([x_0, x_1]\). After setting the derivative equal to 0 and testing the critical and endpoints of \([x_0, x_1]\), we get that the max \( g \) attains is \( g \left( \frac{x_0 + x_1}{2} \right) = \frac{h^2}{4} \), thus simplifying our bound above to
\[
\prod_{j=0}^{n} |x - x_j| \leq |x - x_0||x - x_1| \cdot h^{n-1} n!
\]
\[
\leq \frac{h^2}{4} \cdot h^{n-1} n!
\]
\[
= \frac{h^{n+1} n!}{4}.
\]

Now plugging this expression into the interpolant error bound, we get that a convenient expression for the error in approximating \( f \) by its degree \( \leq n \) interpolant through \( n + 1 \) equispaced interpolation nodes (with one node at each endpoint) is:
\[
|f(x) - p_n(x)| \leq \frac{||f^{(n+1)}||_{[a,b]}}{(n+1)!} \prod_{j=0}^{n} |x - x_j|
\]
\[
\leq \frac{||f^{(n+1)}||_{[a,b]} \cdot h^{n+1} n!}{4(n+1)}.
\]
6 Rate of Convergence

We already have a notion of order of convergence, which was a measure of how fast a sequence converged to a value. Rate gives us a notion of how fast a function converges to a value. If \( f, g: (a, b) \to \mathbb{R}, \ c \in (a, b), \lim_{x \to c} f(x) = L, \lim_{x \to c} g(x) = 0, \ g(x) \neq 0 \) in a punctured neighborhood around \( c, \) and \( \lim_{x \to c} \frac{f(x) - L}{g(x)} \in (0, \infty), \) we say that “\( f \) converges to \( L \) with rate \( g \)”, and the value \( c \) to which the input tends is usually implied.

We also allow \( c = \pm \infty. \) If \( L = 0, \) we say \( f = O(g). \) We often let \( O(g) \) stand in place of the function \( f, \) even if we never formally defined \( f. \)

The notion of rate is fairly flexible. We also include analogous definitions for the case \( L \pm \infty, \) and also for the case that the ratio \( \frac{f(x) - L}{g(x)} \) fails to converge but whose lim inf and lim sup are still bounded. It can be extended to sequences too. There are a few other cases allowed as well, so to avoid confusion, it is better to gain an understanding of the notion of rate rather than the exact definition.

Relation to numerical analysis: we often want to tell how fast a certain error is decaying to 0 with some input parameter, often “\( h \)” or “\( n \)”.

7 Estimating derivatives

Say we’re trying to estimate \( f'(x), \) where \( f \in C^\infty(\mathbb{R}) \) and \( x \in \mathbb{R} \) fixed, by computing \( d(h) = \frac{f(x + 3h) - f(x + h)}{2h}. \) At what rate does \( d(h) \to f'(x)? \) To answer this, and many other questions of this type, it’s often advantageous to Taylor expand.

\[
\frac{f(x + 3h) - f(x + h)}{2h} = \frac{f(x) + 3hf'(x) + \frac{(3h)^2 f''(\xi_1)}{2} - f(x) + hf'(x) + \frac{h^2 f''(\xi_2)}{2}}{2h}
\]

\[
= f'(x) + \frac{h}{4} (9f''(\xi_1) - f''(\xi_2))
\]

so \( d(h) \to f'(x) \) with a rate of \( h, \) or simply, linearly. Technically, since \( \xi_1 \) and \( \xi_2 \) depend on \( h, \) we should also make the assumption that near \( x, \) \( f'' \) is bounded (away from 0 and \( \pm \infty)).

Note that the error term

\[
\frac{h}{4} (9f''(\xi_1) - f''(\xi_2))
\]

cannot be assumed to be 0, and even if we Taylor expanded to a further term, that wouldn’t have improved our \( O(h) \) rate of convergence (because of the mismatched coefficients of \( f'' \).
in this case).

To make matters terribly confusing, we also call this type of convergence “1st order convergence”, where the “order” here refers to the power of $h$ in the rate function that we are comparing to the error, which is a completely different notion to the “order” that was defined before (for sequences)... Basically, “rate” and “order” are used in a very flexible manner, but their implications should be clear in context.

Here are some other examples of $d(h)$ candidates, along with their rate/order calculations to determine how quickly they approach $f'(x)$.

(a) $d(h) = \frac{f(x + h) - f(x)}{h}$

\[
d(h) = \frac{f(x + h) - f(x)}{h} = \frac{f(x) + hf'(x) + \frac{h^2}{2} f''(\xi_1) - f(x)}{h} = f'(x) + \frac{h}{2} f''(\xi_1) = f'(x) + O(h),\
\]

so this method is a first order (linear) approximation to $f'(x)$.

(b) $d(h) = \frac{f(x + h) - f(x - h)}{2h}$

\[
d(h) = \frac{f(x + h) - f(x - h)}{2h} = \frac{f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f'''(\xi_1) - (f(x) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{6} f'''(\xi_2))}{2h} = f'(x) + \frac{h^2}{12} (f'''(\xi_1) + f'''(\xi_2)) = f'(x) + O(h^2),\
\]

so this method of estimating $f'(x)$ is second order accurate (quadratic convergence).
(c) \[ d(h) = \frac{2f(x + 2h) - 4f(x - h) + 2f(x)}{h^2} \]

\[ d(h) = \frac{2f(x + 2h) - 4f(x - h) + 2f(x)}{h^2} \]
\[ = \frac{2(f(x) + 2hf' (\xi_1)) - 4(f(x) - hf' (\xi_2)) + 2f(x)}{h^2} \]
\[ = \frac{1}{h}(4f' (\xi_1) + 4f' (\xi_2)) \]
\[ = \mathcal{O}(1/h), \]

which blows up, and so \( d(h) \) fails to converge to \( f'(x) \) in this case. Note that we cannot assume that \( f'(x) \) is 0 or that \( f'(\xi_1) \) roughly cancels with \( f'(\xi_2) \) in general.

An important point to make is sometimes you don’t know in advance how many terms to include in the Taylor expansion. A good principle to follow is: stop after the first non-vanishing error term. In practice, you might need to make several attempts before you include the right number of terms. If you include too many terms, then you’re potentially doing unnecessary work; too few, and your method might be higher order than you think. So there’s a tradeoff.

Another type of problem will ask us to find the weights of \( f(x + jh) \) for various \( j \) to estimate some derivative of \( f \) at a given point to some specified degree of accuracy (i.e., at a specified rate).

(d) Suppose we want \( d(h) = \frac{af(x - h) + bf(x) + cf(x + 2h)}{h} \) to approach \( f'(x) \) at a rate of \( h^2 \) (quadratic convergence). Taylor expand to find the constants \( a, b, \) and \( c \). Note that we won’t need to calculate terms past \( \mathcal{O}(h^3) \) in the numerator, since we will already have achieved at least second order accuracy if the terms with fewer \( h \) factors vanish.

\[ d(h) = \frac{a}{h} \left[ f(x) - hf' (x) + \frac{h^2}{2} f'' (x) - \frac{h^3}{6} f''' (\xi_1) \right] + \frac{bf(x)}{h} \]
\[ + \frac{c}{h} \left[ f(x) + 2hf' (x) + \frac{4h^2}{2} f'' (x) + \frac{8h^3}{6} f''' (\xi_2) \right] \]
\[ = \frac{1}{h} \left[ a + b + c \right] f(x) + \left[ -a + 2c \right] f'(x) + h \left[ \frac{a}{2} + 2c \right] f''(x) \]
\[ + h^2 \left[ -\frac{a}{6} f'''(\xi_1) + \frac{4c}{3} f'''(\xi_2) \right], \]

so we should solve the system

\[
\begin{align*}
    a + b + c &= 0 \\
    -a + 2c &= 1 \\
    a/2 + 2c &= 0.
\end{align*}
\]
This system is solved by \(a = \frac{-2}{3},\ b = \frac{1}{2},\) and \(c = \frac{1}{6},\) and this solves the problem. Since 
\[\frac{-a}{6} + \frac{4c}{3} = \frac{1}{3} \neq 0,\]
the method is second order accurate (assuming boundedness of \(f''\)).

A more difficult version of this problem might also ask us to find which derivative we are approximating, to what order we are approximating it, and what constants must be used in the finite difference function \(d(h)\) to approximate this derivative. Let us now solve this problem.

\((e)\) \(d(h) = \frac{af(x + h) + bf(x) + cf(x - h)}{h^2}.\) Immediately we can tell that the only derivative this could approximate is the second, since all other terms in the resulting Taylor expansion will either blow up or shrink to 0. Now we Taylor expand:

\[
d(h) = \frac{af(x + h) + bf(x) + cf(x - h)}{h^2}
= \frac{a}{h^2} \left[ f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f'''(x) + \frac{h^4}{24} f^{(4)}(\xi_1) \right] + \frac{b}{h^2} f(x)
+ \frac{c}{h^2} \left[ f(x) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{6} f'''(\xi_2) + \frac{h^4}{24} f^{(4)}(\xi_2) \right]
= \frac{1}{h^2} \left[ a + b + c \right] f(x) + \frac{1}{h} \left[ a - c \right] f'(x) + \frac{1}{2} \left[ a + c \right] f''(x)
+ \frac{h}{6} \left[ a - c \right] f'''(x) + \frac{h^2}{24} \left[ a f^{(4)}(\xi_1) + c f^{(4)}(\xi_2) \right]
\]

And so our equations become
\[
\begin{align*}
a + b + c &= 0 \\
a - c &= 0 \\
a + c &= 2
\end{align*}
\]

Note that two of the equations are identical (thanks to symmetry!), so we only need to include one of them. This of course admits the solution \(a = c = 1,\ b = -2.\) So these are the coefficients that must be used in the finite difference function \(d(h),\) the derivative we are approximating with \(d(h)\) is \(f''(x),\) and this is a second order approximation. The error term approaches \(\frac{h^2}{12} f^{(4)}(x)\) as \(h \to 0,\) assuming continuity of \(f^{(4)}\). So we can say that the rate of convergence of \(d(h)\) to \(f^{(4)}(x)\) is \(h^2 f^{(4)}(x),\) or \(f^{(4)}(x) - d(h) = \mathcal{O}(h^2).\)

\((f)\) Challenge problem: assume \(f\) is twice differentiable (but its 2\textsuperscript{nd} derivative might not be continuous). Show that \(\lim_{h \to 0} \frac{f(x - h) - 2f(x) + f(x + h)}{h^2} = f''(x).\) (Hint: use l’Hospital’s rule). Using the hint, we employ l’Hospital’s rule, and justify later that the resulting expression indeed approaches a limit. Note that the first step is to differentiate both top and
bottom, but with respect to $h$ rather than $x$.

\[
\lim_{h \to 0} \frac{f(x-h) - 2f(x) + f(x+h)}{h^2} = \lim_{h \to 0} \frac{-f'(x-h) + f'(x+h)}{2h} = \frac{1}{2} \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h} + \frac{1}{2} \lim_{h \to 0} \frac{f'(x-h) - f'(x)}{-h} = \frac{1}{2} f''(x) + \frac{1}{2} f''(x) = f''(x).
\]

All right, last round of this game:

(g) Find $a \in \mathbb{R}$ so that $d(h) = \frac{f(x+ah) - f(x-ah)}{2h}$ converges to $f'(x)$ as quickly as possible, and give the first two terms in the asymptotic error expansion.

\[
d(h) = \frac{f(x+ah) - f(x-ah)}{2h} = \left[ f(x) + ahf'(x) + \frac{a^2h^2}{2} f''(x) + \frac{a^3h^3}{6} f'''(x) + \frac{a^4h^4}{24} f^{(4)}(x) + \cdots \right] - \left[ f(x) - ahf'(x) + \frac{a^2h^2}{2} f''(x) - \frac{a^3h^3}{6} f'''(x) + \frac{a^4h^4}{24} f^{(4)}(x) + \cdots \right] \frac{2h}{2h} = af'(x) + \frac{a^3h^2}{6} f''(x) + \frac{a^5h^4}{120} f^{(5)}(x) + \cdots,
\]

which means we must take $a = 1$, the order of convergence is second order, and the first two terms in the asymptotic error expansion are $\frac{h^2}{6} f''(x) + \frac{h^4}{120} f^{(5)}(x)$.

Note that the use of ellipses ($\cdots$) can really only be used if the resulting series converges, which we did not justify. This is often done when computing order though, because we understand that we could make this argument rigorous by simply terminating the sum after a finite number of terms, and evaluating the final derivative of $f$ at an arbitrary $\xi_j$, as prescribed by Taylor’s theorem, just as we did in previous examples. We employ ellipses to make the argument cleaner, so we sacrifice a bit of rigor for clarity.

8 Quadrature rules and errors

Now we try our hand at estimating definite integrals. The technique that we will employ here is approximating the integrand by an interpolating polynomial, and then integrating that polynomial instead of the original integrand. The polynomials will be easy to integrate and will provide us with a simple weighting of evaluations of the integrand.

Let’s start with the simplest of all of these: midpoint rule. We obtain the quadrature
rule by integrating the 0th order interpolating polynomial at the midpoint of the interval of integration. Say we’re trying to estimate \( \int_0^h f(x)dx \). The 0th order interpolating polynomial we will integrate is

\[ p_0(x) = f \left( \frac{h}{2} \right), \]

which integrates to

\[ \int_0^h p_0(x)dx = hf \left( \frac{h}{2} \right). \]

To get a handle on the error involved in estimating the actual integral with this approximation, we can employ integration by parts and some obscure integral manipulations. In this calculation we assume \( f'' \) is continuous (with one-sided derivatives existing at the endpoints).

\[
\begin{align*}
\int_0^h f(x)dx - hf(h/2) &= \int_0^h f(x) - f(h/2)dx \\
&= \int_0^{h/2} f(x) - f(h/2)dx + \int_{h/2}^h f(x) - f(h/2)dx \\
&= [(f(x) - f(h/2))x]_0^{h/2} - \int_0^{h/2} f'(x)dx \\
&+ [(f(x) - f(h/2))(x-h)]_0^{h/2} - \int_{h/2}^h f'(x)(x-h)dx \\
&= - \left[ f'(x)\frac{x^2}{2} \right]_0^{h/2} + \int_0^{h/2} f''(x)\frac{x^2}{2}dx \\
&- \left[ f'(x)\frac{(x-h)^2}{2} \right]_{h/2}^h + \int_{h/2}^h f''(x)\frac{(x-h)^2}{2}dx \\
&= f''(\xi_1) \int_0^{h/2} \frac{x^2}{2}dx + f''(\xi_2) \int_{h/2}^h \frac{(x-h)^2}{2}dx \\
&= f''(\xi_2) \frac{h^3}{48} + f''(\xi_2) \frac{h^3}{48} \\
&= \left( f''(\xi_1) + f''(\xi_2) \right) \frac{h^3}{24} \\
&= \frac{f''(\xi)h^3}{24}.
\end{align*}
\]

The final step is justified by the IVT. Thus the midpoint rule is 3rd order accurate over one panel. Let’s refer to this as the mini midpoint rule.

In the above calculation, we made use of the integral mean value theorem (IMVT): if \( f, g : [a, b] \to \mathbb{R} \) are continuous, \( g \geq 0 \) on \([a, b]\), and \( g(x^*) > 0 \) for some \( x^* \in [a, b] \), then there exists \( \xi \in [a, b] \) such that

\[
\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx.
\]
This is true by the IVT: since $f$ is continuous on the compact interval $[a, b]$, $f$ achieves an absolute min and max over this interval. So say $m, M \in [a, b]$ are such that $f(m) \leq f(x) \leq f(M) \forall x \in [a, b]$. Then
\[
 f(m) \int_a^b g(x) dx = \int_a^b f(m) g(x) dx \leq \int_a^b f(x) g(x) dx \leq \int_a^b f(M) g(x) dx = f(M) \int_a^b g(x) dx.
\]
Then since $g$ is not identically 0 and it's continuous, we know $\int_a^b g(x) dx > 0$, so we get that
\[
 f(m) \int_a^b g(x) dx \leq \int_a^b f(x) g(x) dx \leq f(M) \int_a^b g(x) dx.
\]
and so then by the IVT, $\exists \xi \in [a, b]$ such that
\[
 f(\xi) = \int_a^b f(x) g(x) dx,
\]
thereby proving the IMVT. Of course, a similar argument holds in the case that $g \leq 0$ and $g(x^*) < 0$ for some $x^* \in [a, b]$, and also for more general cases where $f$ merely satisfies the IVT and is Riemann integrable (but not necessarily continuous), and $g$ is integrable with nonzero integral and never crosses the x-axis.

This is great, but what if we want to integrate over some interval $[a, b]$ instead of $[0, h]$? Simple! We use the composite midpoint rule, which just means chopping $[a, b]$ up into many (evenly-spaced, in this class) mini intervals and using the mini midpoint quadrature rule on each subinterval.

By a change of variables, we get that the midpoint rule applied to an interval $[x_i, x_{i+1}]$, where $h = x_{i+1} - x_i$ is:
\[
 \int_{x_i}^{x_{i+1}} f(x) dx = hf \left( \frac{x_i + x_{i+1}}{2} \right) + \frac{f''(\xi)h^3}{24}.
\]
Now proceeding with the error derivation for composite midpoint rule, with panels of uniform width ($n \in \mathbb{N}$, $h = \frac{b - a}{n}$, $x_i = a + ih$ for $i = 0, 1, \ldots, n$, and $f : [a, b] \to \mathbb{R}$ is twice continuously differentiable with one-sided derivatives at the endpoints):
\[
 \int_a^b f(x) dx = \sum_{i=0}^{n-1} \left( \int_{x_i}^{x_{i+1}} f(x) dx \right) \\
 = \sum_{i=0}^{n-1} \left( hf \left( \frac{x_i + x_{i+1}}{2} \right) + \frac{f''(\xi_i)h^3}{24} \right) \\
 = \left( h \sum_{i=0}^{n-1} f \left( \frac{x_i + x_{i+1}}{2} \right) \right) + \left( \frac{1}{n} \sum_{i=0}^{n-1} f(\xi_i) \right) \frac{(nh)h^2}{24} \\
 = \left( h \sum_{i=0}^{n-1} f \left( \frac{x_i + x_{i+1}}{2} \right) \right) + \frac{f''(\xi)(b - a)h^2}{24},
\]
where again the last step was due to the IVT, and also we made use of the fact the $nh = b - a$. 13
So the composite midpoint rule converges of order 2 to the true integral. This demonstrates the general phenomenon of the composite rule order always being one less than the mini (or, single-panel) rule order. This happens because of the simplification $nh = b - a$.

Next, we derive the mini trapezoidal quadrature rule by integrating the linear interpolating polynomial interpolating $f$ at the endpoints of $[0, h]$. This polynomial is

$$p_1(x) = f(0) \frac{x - h}{-h} + f(h) \frac{x}{h},$$

and it makes the quadrature rule

$$\int_0^h p_1(x)dx = f(h) \frac{h^2}{2h} - f(0) \frac{-h^2}{2h} = h \left( \frac{f(0) + f(h)}{2} \right),$$

and we can employ the same integration by parts and integral MVT trick as before to derive the error in mini trapezoidal method:

$$\int_0^h f(x)dx - h \left( \frac{f(0) + f(h)}{2} \right) = \frac{1}{2} \int_0^h f(x)dx - f(0)dx + \frac{1}{2} \int_0^h f(x) - f(h)dx$$

$$= \frac{1}{2} \left[ (f(x) - f(0))(x - h) \right]_0^h - \frac{1}{2} \int_0^h f'(x)(x - h)dx$$

$$+ \frac{1}{2} \left[ (f(x) - f(h))x \right]_0^h - \frac{1}{2} \int_0^h f'(x)dx$$

$$= -\frac{1}{2} \int_0^h f'(x) [(x - h) + (x)] dx$$

$$= -\frac{1}{2} \left[ f'(x)(x - h)(x) \right]_0^h + \frac{1}{2} \int_0^h f''(x)(x - h)(x)df$$

$$= \frac{f''(\xi)}{2} \int_0^h x^2 - xhdx$$

$$= \frac{f''(\xi)}{2} \left( h^3 \frac{3}{3} - h^3 \frac{3}{2} \right)$$

$$= -\frac{f(\xi)h^3}{12},$$

and then by a similar calculation to the one above for composite midpoint, we get that the composite trapezoid error is:

$$\int_a^b f(x)dx = h \left( \sum_{i=0}^{n-1} \frac{f(x_i) + f(x_{i+1})}{2} \right) - \frac{f(\xi)(b - a)h^2}{12}.$$
Now let’s derive a quadrature rule by integrating the quadratic polynomial through 
\((0, f(0)), (h/2, f(h/2)), \) and \((h, f(h))\). That polynomial is

\[
p_2(x) = f(0) \frac{x - h/2}{h^2/2} + f(h/2) \frac{x(x-h)}{-h^2/4} + f(h) \frac{x(x-h/2)}{h^2/2}.
\]

So integrating (over one panel) gives

\[
\int_0^h p_2(x) \, dx = \frac{2f(0)}{h^2} \int_0^h ((x - h) + (h - h/2)(x - h)) \, dx
\]
\[- \frac{4f(h/2)}{h^2} \int_0^h x^2 - hx \, dx
\]
\[+ \frac{2f(h)}{h^2} \int_0^h x^2 - hx/2 \, dx
\]
\[= \frac{2f(0)}{h^2} \frac{h^3}{3} - \frac{h^3}{4} - \frac{4f(h/2)}{h^2} \frac{h^3}{3} - \frac{h^3}{2} + \frac{2f(h)}{h^2} \frac{h^3}{3} - \frac{h^3}{4}
\]
\[= \frac{h}{6} (f(0) + 4f(h/2) + f(h))
\]

The error in using Simpson’s rule over one panel is derived by Taylor expansion. Note that the last term from Taylor expansion (with error term) is continuous in \(x\) even if we have no control over the behavior of \(\xi(x)\), since we can express that last term as a difference of \(f\) and a polynomial.

---

Note: the discussion notes will end here! The material skipped in this set of notes, that was covered either in lecture or discussion, is:

derivation of the asymptotic error expansion for Midpoint, Trapezoid, and Simpson (3/8ths rule not really important)

Taylor expand to get asymptotic error for midpoint

can use add & subtract midpoint from trap error equation and then taylor expand to get asymptotic error expansion of trap from midpoint

can use that Simpson=(2/3)Mid + (1/3)Trap to get asymptotic error expansion for Simpson Gaussian quadrature...
*open* Newton-Cotes quadrature formulas not so important for this class

using u-substitution to get new, translated/scaled quadrature rules from existing quadrature rules

LU matrix factorization

Cholesky matrix factorization for symmetric positive definite matrices

Richardson extrapolation to knock off first term in the asymptotic error expansion

Romberg’s method (just means multiple applications of Richardson extrapolation)

Aitken extrapolation (note: completely different from Aitken acceleration...), a way to estimate order of convergence WITHOUT knowing the true quantity being estimated (!!) by computing (e.g., for a quadrature problem where \( I_E \) represents the exact integral, \( I_h \) represents the approximate integral obtained using parameter \( h \), and we expect \( I_E - I_h \approx c_p h^p \), \( c_p \) independent of \( h \))

\[
\log_2 \left| \frac{I_{h/2} - I_h}{I_{h/4} - I_{h/2}} \right|
\]

splines... cubic (linear and quadratic not as important) and the constraints they must satisfy

Good luck on the final and may our paths cross again!