Lecture 2: Newton–Cotes Quadrature

See Chapter 7 of Süli and Mayers.

Terminology: Quadrature $\equiv$ numerical integration.

Setup: given $f(x_k)$ at $n+1$ equally spaced points $x_k = x_0 + k \cdot h$, $k = 0, 1, ..., n$, where $h = (x_n - x_0)/n$. Suppose that $p_n(x)$ interpolates this data.

Idea: does

$$\int_{x_0}^{x_n} f(x)dx \approx \int_{x_0}^{x_n} p_n(x)dx?$$

We investigate the error in such an approximation below, but note that

$$\int_{x_0}^{x_n} p_n(x)dx = \int_{x_0}^{x_n} \sum_{k=0}^{n} f(x_k) \cdot L_{n,k}(x)dx \tag{1}$$

$$= \sum_{k=0}^{n} f(x_k) \cdot \int_{x_0}^{x_n} L_{n,k}(x)dx \tag{2}$$

$$= \sum_{k=0}^{n} w_k f(x_k), \tag{3}$$

where the coefficients

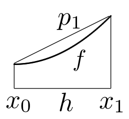
$$w_k = \int_{x_0}^{x_n} L_{n,k}(x)dx$$

$k = 0, 1, ..., n$, are independent of f . A formula

$$\int_a^b f(x)dx \approx \sum_{k=0}^{n} w_k f(x_k)$$

with $x_k \in [a, b]$ and w_k independent of f for $k = 0, 1, ..., n$ is called a quadrature formula; the coefficients w_k are known as weights. The specific form (1)–(3), based on equally spaced points, is called a Newton–Cotes formula of order n.

Examples: Trapezium Rule: $n = 1$ (also known as the trapezoid or trapezoidal rule):

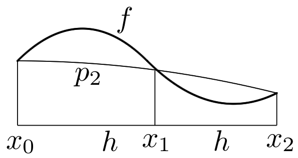


$$\int_{x_0}^{x_1} f(x)dx \approx \frac{h}{2}[f(x_0) + f(x_1)]$$

Proof

$$\begin{aligned}
\int_{x_0}^{x_1} p_1(x)dx &= f(x_0)\int_{x_0}^{x_1} L_{1,0}(x)dx + f(x_1)\int_{x_0}^{x_1} L_{1,1}(x)dx \\
&= f(x_0)\int_{x_0}^{x_1} \frac{x - x_1}{x_0 - x_1}dx + f(x_1)\int_{x_0}^{x_1} \frac{x - x_0}{x_1 - x_0}dx \\
&= f(x_0)\frac{(x_1 - x_0)}{2} + f(x_1)\frac{(x_1 - x_0)}{2}
\end{aligned}$$

Simpson's Rule: $n = 2$:



$$\int_{x_0}^{x_2} f(x)dx \approx \frac{h}{3}[f(x_0) + 4f(x_1) + f(x_2)]$$

Note: The trapezium rule is exact if $f \in \Pi_1$, since if $f \in \Pi_1$, therefore $p_1 = f$. Similarly, Simpson's Rule is exact if $f \in \Pi_2$, since if $f \in \Pi_2$, therefore $p_2 = f$. The highest degree of polynomial exactly integrated by a quadrature rule is called the (polynomial) degree of accuracy (or degree of exactness).

Error: we can use the error in interpolation directly to obtain

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

so that

$$\int_{x_0}^{x_n} [f(x) - p_n(x)]dx \leq \frac{1}{(n+1)!} \max_{\xi \in [x_0, x_n]} |f^{(n+1)}(\xi)| \int_{x_0}^{x_n} |\pi(x)|dx, \tag{4}$$

which, e.g., for the trapezium rule, $n = 1$, gives

$$\left| \int_{x_0}^{x_1} f(x)dx - \frac{(x_1 - x_0)}{2}[f(x_0) + f(x_1)] \right| \leq \frac{(x_1 - x_0)^3}{12} \max_{\xi \in [x_0, x_1]} |f''(\xi)|.$$

In fact, we can prove a tighter result:

Theorem. Error in Trapezoidal Rule:

$$\left| \int_{x_0}^{x_1} f(x)dx - \frac{(x_1 - x_0)}{2}[f(x_0) + f(x_1)] \right| = \frac{(x_1 - x_0)^3}{12} f''(\xi)$$

for some $\xi \in (x_0, x_1)$. (And note equality)

Proof. Omitted (uses Integral Mean-Value Theorem).

For $n > 1$, (4) gives pessimistic bounds. But one can prove better results, e.g., using Taylor Series.

Theorem. Error in Simpson's Rule: if f'''' is continuous on (x_0, x_2), then

$$\left| \int_{x_0}^{x_2} f(x)dx - \frac{x_2 - x_0}{6}[f(x_0) + 4f(x_1) + f(x_2)] \right| = \frac{(x_2 - x_0)^5}{2880} f''''(\xi)$$

for some $\xi \in (x_0, x_2)$.

Proof. See, e.g., Süli and Mayers, Thm. 7.2.

Note: Simpson's Rule is exact if $f \in \Pi_3$ since then $f'''' \equiv 0$. (c.f. earlier statement viz. $f \in \Pi_2$).

Composite Quadrature (optional material)

Motivation: we've seen oscillations in polynomial interpolation—the Runge phenomenon—for high-degree polynomials on equispaced grids.

Idea: split a required integration interval $[a, b] = [x_0, x_n]$ into n equal intervals $[x_{i-1}, x_i]$ for $i = 1, \ldots, n$. Then use a **composite rule**:

$$\int_a^b f(x)\,\mathrm{d}x = \int_{x_0}^{x_n} f(x)\,\mathrm{d}x = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)\,\mathrm{d}x$$

in which each $\int_{x_{i-1}}^{x_i} f(x)\,\mathrm{d}x$ is approximated by quadrature.

Thus rather than increasing the degree of the polynomials to attain high accuracy, instead increase the number of intervals.

Composite Trapezium Rule:

$$\begin{aligned}\int_{x_0}^{x_n} f(x)\,\mathrm{d}x = &\sum_{i=1}^{n}\left[\frac{h}{2}[f(x_{i-1}) + f(x_i)] - \frac{h^3}{12} f''(\xi_i)\right] \\ = &\frac{h}{2}[f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)] + e_h^{\mathrm{T}}\end{aligned}$$

where $\xi_i \in (x_{i-1}, x_i)$ and $h = x_i - x_{i-1} = (x_n - x_0)/n = (b-a)/n$, and the error e_h^{T} is given by

$$e_h^{\mathrm{T}} = -\frac{h^3}{12}\sum_{i=1}^{n} f''(\xi_i) = -\frac{nh^3}{12} f''(\xi) = -(b-a)\frac{h^2}{12} f''(\xi)$$

for some $\xi \in (a, b)$, using the Intermediate-Value Theorem n times. Note that if we halve the stepsize h by introducing a new point halfway between each current pair (x_{i-1}, x_i), the factor h^2 in the error will decrease by four.

Alternatively, divide $[a, b]$ into $2n + 1$ intervals: $[a, b] = [x_0, x_{2n}]$. Then:

Composite Simpson's Rule:

$$\begin{aligned}\int_{x_0}^{x_{2n}} f(x)\,\mathrm{d}x = &\sum_{i=1}^{n}\left[\frac{h}{3}[f(x_{2i-2}) + 4f(x_{2i-1}) + f(x_{2i})] - \frac{(2h)^5}{2880} f''''(\xi_i)\right] \\ = &\frac{h}{3}[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots \\ &+ 2f(x_{2n-2}) + 4f(x_{2n-1}) + f(x_{2n})] + e_h^{\mathrm{S}}\end{aligned}$$

where $\xi_i \in (x_{2i-2}, x_{2i})$ and $h = x_i - x_{i-1} = (x_{2n} - x_0)/2n = (b-a)/2n$, and the error e_h^{S} is given by

$$e_h^{\mathrm{S}} = -\frac{(2h)^5}{2880}\sum_{i=1}^{n} f''''(\xi_i) = -\frac{n(2h)^5}{2880} f''''(\xi) = -(b-a)\frac{h^4}{180} f''''(\xi)$$

for some $\xi \in (a, b)$. Note that if we halve the stepsize h by introducing a new point half way between each current pair (x_{i-1}, x_i), the factor h^4 in the error will decrease by sixteen.

Adaptive (or automatic) procedure: if S_h is the value given by composite Simpson's rule with a stepsize h, then

$$S_h - S_{\frac{1}{2}h} \approx -\frac{15}{16} e_h^{\mathrm{S}}.$$

This suggests that if we wish to compute $\int_a^b f(x)\,\mathrm{d}x$ with an absolute error ε, we should compute the sequence $S_h, S_{\frac{1}{2}h}, S_{\frac{1}{4}h}, \ldots$ and stop when the difference, in absolute value, between two consecutive values is smaller than $\frac{16}{15}\varepsilon$. That will ensure that (approximately) $|e_h^{\mathrm{S}}| \le \varepsilon$.

Sometimes much better accuracy may be obtained: for example, as might happen when computing Fourier coefficients, if f is periodic with period $b-a$ so that $f(a+x) = f(b+x)$ for all x.

Matlab:

```
>> help adaptive_simpson
 ADAPTIVE_SIMPSON  Adaptive (or automatic) Simpson's rule.
    S = ADAPTIVE_SIMPSON(F,A,B,NMAX,TOL) computes an approximation
    to the integral of F on the interval [A,B]. It will take a
    maximum of NMAX steps and will attempt to determine the
    integral to a tolerance of TOL.

    The function uses an adaptive Simpson's rule, as described
    in lectures.
>> f = @(x) sin(x);
>> adaptive_simpson(f, 0, pi, 100, 1e-7);
Step 1 integral is 2.0943951024, with error estimate 2.0944.
Step 2 integral is 2.0045597550, with error estimate 0.089835.
Step 3 integral is 2.0002691699, with error estimate 0.0042906.
Step 4 integral is 2.0000165910, with error estimate 0.00025258.
Step 5 integral is 2.0000010334, with error estimate 1.5558e-05.
Step 6 integral is 2.0000000645, with error estimate 9.6884e-07.
Successful termination at iteration 7:
The integral is 2.0000000040, with error estimate 6.0498e-08.

>> g = @(x) sin(sin(x));
>> fplot(g, [0 pi])
>> adaptive_simpson(g, 0, pi, 100, 1e-7);
Step 1 integral is 1.7623727094, with error estimate 1.7624.
Step 2 integral is 1.8011896009, with error estimate 0.038817.
Step 3 integral is 1.7870879453, with error estimate 0.014102.
Step 4 integral is 1.7865214631, with error estimate 0.00056648.
Step 5 integral is 1.7864895607, with error estimate 3.1902e-05.
Step 6 integral is 1.7864876112, with error estimate 1.9495e-06.
Step 7 integral is 1.7864874900, with error estimate 1.2118e-07.
Successful termination at iteration 8:
The integral is 1.7864874825, with error estimate 7.5634e-09.
```