Numerical Integration (Quadrature)

Another application for our interpolation tools!
Integration: Area under a curve

Curve = data or function

- Integrating data
  - Finite number of data points—spacing specified
  - Data may be noisy
  - Must interpolate or regress a smooth between points

- Integrating functions
  - Non-integrable function (usually)
  - As many points as you need
  - Spacing arbitrary

Today we will discuss methods that work on both of these problems
Numerical Integration: The Big Picture

$$I = \int_{a}^{b} dx \ f(x)$$

- Virtually all numerical integration methods rely on the following procedure:
  - Start from $N+1$ data points $(x_i, f_i)$, $i = 0, \ldots, N$, or sample a specified function $f(x)$ at $N+1$ $x_i$ values to generate the data set
  - Fit the data set to a polynomial, either locally (piecewise) or globally
  - Analytically integrate the polynomial to deduce an integration formula of the general form:

$$I \approx \sum_{i=0}^{N} w_i f(x_i) = \{w\}^T \{f\}$$

  - $w_i$ are the “weights”
  - $x_i$ are the “abscissas”, “points”, or “nodes”

- Numerical integration schemes are further categorized as either:
  - **Closed** – the $x_i$ data points include the end points $a$ and $b$ of the interval
  - **Open** – the $x_i$ data points are interior to the interval
Further classification of numerical integration schemes

- **Newton-Cotes Formulas**
  - Use equally spaced abscissas
  - Fit data to local order N polynomial approximants
  - Examples:
    - Trapezoidal rule, N=1
    - Simpson’s 1/3 rule, N=2
  - Errors are algebraic in the spacing $h$ between points

- **Clenshaw-Curtis Quadrature**
  - Uses the Chebyshev abscissas
  - Fit data to global order N polynomial approximants
  - Errors can be spectral, $\sim \exp(-N) \sim \exp(-1/h)$, for smooth functions

- **Gaussian Quadrature**
  - Unequally spaced abscissas determined optimally
  - Fit data to global order N polynomial approximants
  - Errors can be spectral, and smaller than Clenshaw-Curtis
**Trapezoid rule: 1 interval, 2 points**

Approximate the function by a linear interpolant between the two end points, then integrate that degree-1 polynomial.

\[
Area = \int_{a}^{b} f(x) \, dx
\]

\[
I = \frac{b - a}{2} [f(a) + f(b)] + O((b - a)^3)
\]

Average height:

\[
Area = \frac{1}{2} [f(a) + f(b)] (b - a)
\]
Trapezoid rule: 2 intervals, 3 points

We should be able to improve accuracy by increasing the number of intervals: this leads to “composite” integration formulas.

\[
\text{Area} = \int_a^b f(x) \, dx
\]

\[
h \equiv x_1 - x_0 = x_2 - x_1
\]

\[
\text{Area} = \frac{h}{2} \left[ f(x_0) + f(x_1) \right] + \frac{h}{2} \left[ f(x_1) + f(x_2) \right]
\]

\[
= \frac{h}{2} \left[ f(x_0) + 2f(x_1) + f(x_2) \right]
\]
“Composite” Trapezoid rule: \( 8 \rightarrow n \) intervals, \( n + 1 \) points

\[ \text{Area} = \frac{h}{2} [f(x_0) + f(x_1)] + \frac{h}{2} [f(x_1) + f(x_2)] + \frac{h}{2} [f(x_2) + f(x_3)] + \ldots + \frac{h}{2} [f(x_{n-1}) + f(x_n)] \]

\[ = \frac{h}{2} [f(x_0) + 2f(x_1) + 2f(x_2) + \ldots + 2f(x_{n-1}) + f(x_n)] \]

\[ = \frac{h}{2} \left[ f(x_0) + 2 \sum_{i=1}^{n-1} f(x_i) + f(x_n) \right] \quad \rightarrow \quad \frac{h}{2} \left[ f(x_0) + 2 \sum_{i=1}^{n-1} f(x_i) + f(x_n) \right] \quad ; \quad h = \frac{b-a}{n} \]
Composite Trapezoidal Rule

\[ I = \frac{h}{2} \left[ f(x_0) + 2 \sum_{i=1}^{n-1} f(x_i) + f(x_n) \right] + E_t, \quad h = (b - a)/n \]

- Notice that this composite formula can be written in the generic form:
  \[ I = \sum_{i=0}^{n} w_i f(x_i) + E_t \]
where the weights are given by
  \[ w_i = h, \quad i = 1, \ldots, n - 1 \]
  \[ w_0 = w_n = h/2 \]
- The truncation error of the composite trapezoidal rule is of order
  \[ E_t = O(h^3 \times n) = O((b - a)h^2) \]

Error per interval times number of intervals
Simpson’s 1/3 rule

Basic idea: interpolate between 3 points using a parabola (N=2 polynomial) rather than a straight line (as in trapezoid rule)
Simpson’s 1/3 rule: derivation

Interpolation: find equation for a parabola (N=2 polynomial) passing through 3 points

\[ p(x) \] is a parabola (quadratic in \( x \))
\[ p(x) \] goes through 3 desired points
Use \textbf{Lagrange} form for convenience:

\[
p(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} f(x_0) + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f(x_1) + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f(x_2)
\]

- \( = 0 \) at \( x = x_1 \)
- \( = 0 \) at \( x = x_2 \)
- \( = f(x_0) \) at \( x = x_0 \)
- \( = 0 \) at \( x = x_0 \)
- \( = 0 \) at \( x = x_1 \)
- \( = f(x_1) \) at \( x = x_1 \)
- \( = 0 \) at \( x = x_0 \)
- \( = f(x_2) \) at \( x = x_2 \)
Simpson’s 1/3 rule: derivation for 2 intervals, 3 points

\[ p(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} f(x_0) + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} f(x_1) + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} f(x_2) \]

Area = \( \int_{x_0}^{x_2} p(x) \, dx \)

Remember from calculus:

\[ \int dx \ x^n = \frac{x^{n+1}}{n+1} \]

\[ h = x_2 - x_1 = x_1 - x_0 \]
Composite Simpson’s 1/3 rule: 
\((n\ \text{intervals, } n+1\ \text{points, } n\ \text{even})\)

\[
\text{Area} = \frac{h}{3} \left\{ [f(x_0) + 4f(x_1) + f(x_2)] + [f(x_2) + 4f(x_3) + f(x_4)] + \cdots + [f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)] \right\}
\]

\[
= \frac{h}{3} \left\{ f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n) \right\}
\]

\[
= \frac{h}{3} \left\{ f(x_0) + 4 \sum_{i=1,3,5,\ldots}^{n-1} f(x_i) + 2 \sum_{i=2,4,6,\ldots}^{n-2} f(x_i) + f(x_n) \right\}
\]

The formula is thus:

\[
I = \sum_{i=0}^{n} w_i f(x_i) + E_s, \quad w_i = \begin{cases} 
\frac{h}{3} & i = 0, n \\
\frac{4h}{3} & i = 1, 3, 5, \cdots, n - 1 \\
\frac{2h}{3} & i = 2, 4, 6, \cdots, n - 2 
\end{cases}
\]

The error is:

\[
E_s = O(h^5 \times n) = O((b - a)h^4)
\]

\[
E_t = O(h^3 \times n) = O((b - a)h^2)
\]

compare with trapezoidal rule
Closed Newton-Cotes Formulas

- We have just derived the first two closed Newton-Cotes formulas for equally spaced points that include the end points of the interval \([a,b]\).
- Basic truncation errors
  - Trapezoidal (1 interval, 2 points): \(O(h^3)\)
  - Simpson’s 1/3 (2 intervals, 3 points): \(O(h^5)\)
- Composite formula truncation errors (n intervals, n+1 points)
  - Trapezoidal (“first-order accurate”): \(O(h^2(b-a)) = O(1/n^2)\)
  - Simpson’s 1/3 (“third-order accurate”): \(O(h^4(b-a)) = O(1/n^4)\)
- Higher order formulas can evidently be developed by using local polynomials with \(N>2\) to interpolate with a larger number of points.
- It is also easy to adapt these formulas to composite formulas with unequally spaced intervals, \(h_1 \neq h_2 \neq h_3 \ldots\) if the data is unevenly spaced.
Example: Simpson’s 1/3 Rule

- Let’s try out the composite Simpson’s 1/3 rule on the integral:

\[ I = \int_{-1}^{1} dx \ x \sin x \]

- Exact result is 2[\sin(1) - \cos(1)] = 0.602337… See SimpsonL11.m

<table>
<thead>
<tr>
<th>n</th>
<th>error</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2.20 x 10^{-3}</td>
</tr>
<tr>
<td>8</td>
<td>1.35 x 10^{-4}</td>
</tr>
<tr>
<td>16</td>
<td>8.34 x 10^{-6}</td>
</tr>
</tbody>
</table>

Errors consistent with \~1/n^4 scaling
Open Newton-Cotes formulas

- Same idea as closed formulas, but end points $x_0 = a$, $x_n = b$ are not included
- These are useful if
  - Data is not supplied at the edges of the interval $[a,b]$
  - The integrand has a singularity at the edges (see later)
- A simple and useful example – the composite midpoint rule:

$$\int_a^b dx f(x) = h[f(x_{1/2}) + f(x_{3/2}) + f(x_{5/2}) + \cdots + f(x_{n-3/2}) + f(x_{n-1/2})] + O(n^{-2})$$

where $x_{1/2}$ is located at the midpoint between $x_0$ and $x_1$, etc.

Same accuracy as composite TR
Options for higher accuracy

If we want highly accurate results for an integral, we have various options:

- Use more points \( n \) in Simpson’s rule (since error \( \sim 1/n^4 \)) – computationally costly
- Derive a higher-order Newton-Cotes formula that fits more points to a polynomial with \( N>2 \) – analytically cumbersome for small gains
- Use repeated Richardson extrapolation to develop higher order approximations \( \rightarrow \) Romberg integration
- Switch to a method based on global approximants:
  - Clenshaw-Curtis Quadrature
  - Gauss Quadrature
Clenshaw-Curtis Quadrature

- Clenshaw-Curtis quadrature is based on integrating a global Chebyshev polynomial interpolant through all \( N+1 \) points

\[
f(x) \approx p(x) = \sum_{i=0}^{N} c_i T_i(x)
\]

- The abscissas are fixed at the Chebyshev points:

\[
x_j = \cos(j\pi/N), \quad j = 0, 1, \ldots, N
\]

- If we integrate \( p(x) \):

\[
I \approx \int_{-1}^{1} dx \ p(x) = \sum_{i=0,2,4,\ldots}^{N} \frac{2c_i}{1 - i^2}
\]

- Finally, the weights are obtained by substituting the solution of the collocation equations \([T] \{c\} = \{f\}\) into this expression for the \( c_i \) values

\[
\{c\} = [T]^{-1} \{f\} \quad \Rightarrow \quad I = \{w\}^T \{f\}
\]

- The M-file `clencurt.m` from L.N. Trefethen, “Spectral Methods in MatLab” (SIAM, 2000) encapsulates these weight calculations
Example: Clenshaw-Curtis

- See `ClenCurtL11.m`
- Compares the error vs. N for integrals of 4 functions:

Spectral accuracy for the two smooth functions:

Error below machine precision for N > 10!

c.f. Simpson’s rule error of $10^{-5}$ at N=16
Gauss Quadrature:
Global like Clenshaw-Curtis, but also optimally choose abscissas

2 point case:

$$I = \int_{-1}^{1} f(x) \, dx$$

poor approximation

$$I \approx \frac{h}{2} [f(a) + f(b)]$$

$$= f(-1) + f(1)$$

much better approximation obtained by adjusting $$x_1$$ & $$x_2$$ values and $$w_1$$ and $$w_2$$ values

$$I \approx w_1 f(x_1) + w_2 f(x_2)$$
Finding optimal choice of parameters

Note that there are 4 unknowns to be determined.

Gauss procedure: we choose the 4 unknowns so that the integral gives an exact answer whenever \( f(x) \) is a polynomial of order 3 or less

\[
I = \int_{-1}^{1} dx \ f(x) \approx w_1 f(x_1) + w_2 f(x_2)
\]

\( \Rightarrow \) it would have to give the exact answer for the following four cases:

\[
f(x) = 1, \quad f(x) = x, \quad f(x) = x^2, \quad f(x) = x^3.
\]
Determining the unknowns: Gauss-Legendre Quadrature (N=2 case)

- More generally, for an N point formula, the **abscissas** are the N **roots** of the Legendre Polynomial \( P_N(x) \). The **weights** can be obtained by solving a linear system with a tridiagonal matrix.
- Notice that Gauss-Legendre is an **open** formula, unlike Clenshaw-Curtis.
Example: Gauss-Legendre

- See `gauss.m` (for abscissas and weights) and `GaussL11.m`
- Compares the error vs. N for integrals of same 4 functions:

Spectral accuracy for the two smooth functions:

Error below machine precision for N > 6!

Faster convergence than even Clenshaw-Curtis for smooth functions

c.f. Simpson's rule error of $10^{-5}$ at N=16

Gauss exactly integrates a polynomial of degree $2N-1$!
Adjusting the limits of integration

- The spectral accuracy of the Gauss-Legendre and Clenshaw-Curtis methods can be traced to the fact that they employ global polynomial interpolation and cluster their abscissas at the edges of the interval.

- Remember that the Clenshaw-Curtis and Gauss-Legendre formulas apply only for integrals with limits of -1 and +1!

- The limits need to be rescaled for other integrals:

\[
J = \int_a^b dx \ f(x) \\
\text{Let } \tilde{x} = \left( \frac{2}{b-a} \right) x - \frac{a+b}{b-a} \in [-1, 1] \\
\text{Then } d\tilde{x} = \left( \frac{2}{b-a} \right) dx \\
\]

\[
J = \left( \frac{b-a}{2} \right) \int_{-1}^{1} d\tilde{x} \ f \left( \left[ \frac{b-a}{2} \right] \left[ \tilde{x} + \frac{a+b}{b-a} \right] \right)
\]
Integrals with Singularities

- Occasionally you will encounter integrals with weak (integrable) singularities, e.g.

\[ I = \int_{0}^{1} dx \, x^{-1/2} \exp(x) \]

- In this case there is an inverse square root singularity at the endpoint \( x=0 \). We obviously cannot apply a closed integration formula in this case!

- Alternatives are:
  - Apply an open formula, like composite midpoint rule or Gauss-Legendre quadrature
  - Use an open formula that explicitly includes the \( x^{1/2} \) factor (better option)
  - Rescale \( x \) variable as \( x = t^2 \), so the singularity is removed (best option, if available):

\[ I = 2 \int_{0}^{1} dt \, \exp(t^2) \]
Indefinite (improper) integrals

- Often you will encounter integrals with infinite limits, e.g.

\[ I = \int_{0}^{\infty} dx \ f(x) \]

- There are several alternatives:
  - Apply a Newton-Cotes formula to a similar integral, but with \( \infty \) replaced with a large number \( R \)
  - Rescale \( x \) variable as \( x = -\ln t \), (assuming resulting integral not singular):

\[ I = \int_{0}^{1} dt \ t^{-1} f(-\ln t) \]

- A better approach is to use a Gaussian quadrature formula appropriate for the interval \([0, \infty)\), such as the Gauss-Laguerre formula:

\[ \int_{0}^{\infty} dx \ x^\alpha \exp(-x) f(x) = \sum_{j=1}^{N} w_j f(x_j) \]

Don’t forget to factor the \( x^\alpha \exp(-x) \) out of your function!