Lecture 3: Simple Integrals

Slide Type Slide

Simple Numerical Integration Techniques

Slide Type

Basic Idea of Numerical integration

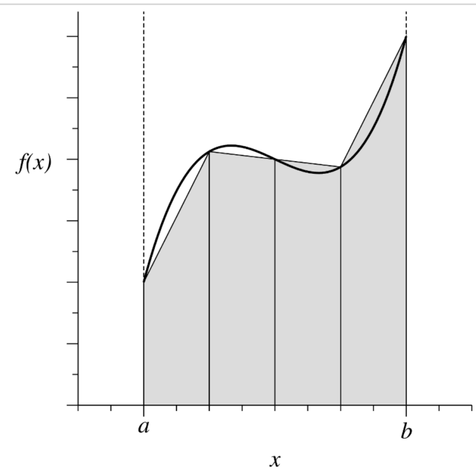
Slide Type Slide

- Think of integrals as areas under curves.
- Approximate these areas in terms of simple shapes (rectangles, trapezoids, rectangles with parabolic tops)
- Simplest case: Riemann sum, $I \approx \sum_k f(x_k)h$, with $h =$ slice width (bottom-left panel below)

Trapezoidal rule

- Break up interval into N slices,
- Approximate function as segments on each slice.

From Newman: fig. 5.1b



- N slices from a to b means that slice width:
$$h = (b-a)/N$$
- area of k^{th} slice's trapezoid: (Rectangle + Triangle)
$$A_k = f(x_k)h + \frac{h[f(x_k+h) - f(x_k)]}{2}$$
$$= \frac{h[f(x_k) + f(x_k+h)]}{2}.$$
- Total area (our approximation for the integral) (and using $x_k = a + kh$):
$$\boxed{I(a,b) \approx h\left[\frac{1}{2}f(a) + \frac{1}{2}f(b) + \sum_{k=1}^{N-1} f(a+kh)\right]}.$$
- Note how it is almost a Riemann sum, except for beginning and end points. Yet, the differences will be significant!

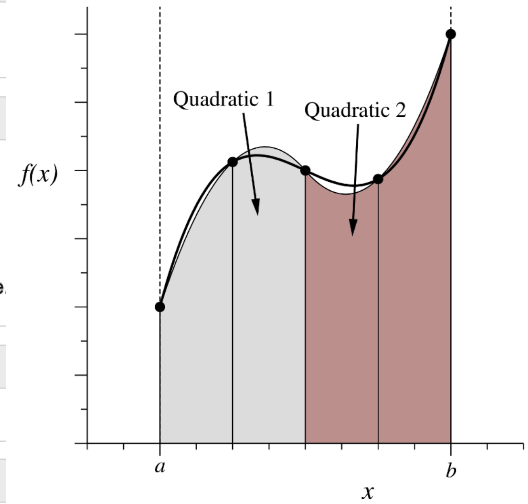
Slide Type Sub-Slide

$$\text{Recall } I(a,b) \approx h\left[\frac{1}{2}f(a) + \frac{1}{2}f(b) + \sum_{k=1}^{N-1} f(a+kh)\right].$$

Simpson's rule

- Break up interval into N slices,
- approximate function as a **quadratic** for every 2 slices
- need 2 slices because you need 3 points to define a quadratic
- more slices $\Rightarrow$ better approximation to function
- Number of slices need to be even! If uneven, either discard one, or use trapezoidal rule on one slice.

From Newman: fig. 5.2



- Area of each 2-slice quadratic (see text for formula):

$$A_k = \frac{h}{3}\{f[a+(2k-2)h] + 4f[a+(2k-1)h] + f(a+2kh)\}\,.$$

- Adding up the slices:

$$\boxed{I(a,b) \approx \frac{h}{3}\left[f(a) + f(b) + 4\sum_{\substack{k\ odd\\ 1\ldots N-1}} f(a+kh) + 2\sum_{\substack{k\ even\\ 2\ldots N-2}} f(a+kh)\right]}.$$

- In Python, you can easily sum over even and odd values: `for k in range(1, N, 2)` for the odd terms, and `for k in range(2, N, 2)` for the even terms.

Slide Type

Newton-Cotes formulas

Slide Type Fragment

Trapezoid and Simpson's Rules are part of a more general set of integration rules:

- Break your interval into small **equal** sub-intervals,
- approximate your function by a polynomial of some degree, e.g.
  - 0 for Riemann Sum (mid-point rule, just summing all elements and multiplying by h)
  - 1 for Trapezoidal
  - 2 for Simpson

on that sub-interval.

- this class of methods leads to Newton-Cotes (N-C) formulas.

Slide Type Sub-Slide

- All Newton-Cotes formulas can be written in the form:

$$\int_a^b f(x)dx \approx \sum_{k=1}^{N+1} w_k f(x_k).$$

- w_k: "weights".
- x_k: "sample points". Notice above we are using $N+1$ points (N slices) to sample.
- N-C formulas of degree N: exact for polynomials of degree N (which require $N+1$ points to determine)
- For N-C formulas, the sample points are **evenly spaced**.

Generalization

Degree	Polynomial	Coefficients
1 (trapezoidal)	Straight line	$\frac{1}{2}, 1, 1, \ldots, 1, \frac{1}{2}$
2 (Simpson)	Parabola	$\frac{1}{3}, \frac{4}{3}, \frac{2}{3}, \frac{4}{3}, \ldots, \frac{2}{3}, \frac{4}{3}, \frac{1}{3}$
3	Cubic	$\frac{3}{8}, \frac{9}{8}, \frac{9}{8}, \frac{3}{4}, \frac{9}{8}, \frac{9}{8}, \frac{3}{4}, \ldots, \frac{9}{8}, \frac{3}{8}$
4	Quartic	$\frac{14}{45}, \frac{64}{45}, \frac{8}{15}, \frac{64}{45}, \frac{28}{45}, \frac{64}{45}, \ldots, \frac{64}{45}, \frac{14}{45}$

Slide Type

Error estimation

Slide Type Sub-Slide

- If you tried the trapezoidal integration routine, you noticed that the error (difference between true value of the integral and computed value) goes down as N increases.
- How fast?

Slide Type Sub-Slide

Euler-MacLaurin formulas for error

Based on Taylor expansions.

Example, for the trapezoidal rule:

$$I(a,b) = \int_a^b f(x)dx \underbrace{=}_{look!} \underbrace{h\left[\frac{1}{2}f(a) + \frac{1}{2}f(b) + \sum_{k=1}^{N-1} f(a+kh)\right]}_{\text{the method}} + \underbrace{\epsilon}_{\text{the error}}$$

Slide Type Sub-Slide

- Trapezoidal rule is a "1st-order" integration rule, i.e. accurate up to and including terms proportional to h. Leading order approximation error is of order h^2:
$$\boxed{\epsilon = \frac{h^2[f'(a) - f'(b)]}{12} + h.o.t.}$$
- Simpson's rule is a "3rd-order" integration rule, i.e., accurate up to and including terms proportional to h^3. Leading order approximation error is of order h^4 (even though we go from segments to quadratics!)
$$\boxed{\epsilon = \frac{h^4[f'''(a) - f'''(b)]}{180} + h.o.t.}$$

  (we won't worry about deriving these)

Adaptive methods

What if you don't know f', f''', etc.? If you know the order of the error, there is another way: compute the integral using N intervals, then double N and compute the integral again. Based on the order, a formula can be derived relating the error ϵ on the latter result to the difference between the results.

- e.g. for trapezoidal rule, we can (but won't) derive the following formula: $\epsilon = (I_{2N} - I_N)/3$
- for Simpson's, $\epsilon = (I_{2N} - I_N)/15$
- when you're re-computing the integral with N doubled: re-use some of your results from the previous computation (since the sample points for the previous estimate are nested inside the points for the new estimate)
- you can add ϵ (as calculated above) to I_{2N}, to obtain a new estimate that is accurate to two more orders!

Slide Type

Often we want to calculate the value of an integral to a given accuracy (e.g. 4 decimal places), and don't know beforehand what value of N will be required to achieve this. We could just start with a really huge N, but this is computationally expensive. Instead use an adaptive method:

1. Evaluate integral using a small N
2. Double N, evaluate again, and calculate error using formula above
3. If error doesn't satisfy our accuracy criterion, repeat step 2. Keep repeating until accuracy is achieved.

Slide Type

Romberg Integration

Slide Type

An example of "Richardson extrapolation", a technique in which higher-order estimates of quantities are calculated iteratively or recursively from lower-order ones

If you're unfamiliar with recursion: that's OK, I will try to implement algorithms iteratively rather than recursively in the code in this course, even though recursion is sometimes more computationally efficient.

Slide Type

Define $R_{i,m}$ as the estimate calculated at the ith round of the doubling procedure (in the adaptive method above), with an error of order h^{2m}. Recursion relation:

$$R_{i,m+1} = R_{i,m} + \frac{R_{i,m} - R_{i-1,m}}{4^m - 1}$$

Slide Type

1. Calculate $R_{1,1}$ and $R_{2,1}$ using the trapezoidal rule.
2. Calculate $R_{2,2}$ by recursion relation using the results from step 1
3. Calculate $R_{3,1}$ using trapezoidal rule
4. Calculate $R_{3,2}$ by recursion relation using the results from steps 1 and 3, then calculate $R_{3,3}$ by recursion relation using $R_{3,2}$ and the result from step 2.
5. At each successive step i, compute one more trapezoidal rule estimate $R_{i,1}$. Then compute $R_{i,2}$ through $R_{i,i}$ by recursion relation using results of previous steps. Also compute the error on $R_{i,i}$ and stop when it's accurate enough.

Slide Type

- Note, we are essentially calculating the integral by doing a series expansion in powers of h.
- This works best in cases where the power series converges rapidly
- This doesn't work so well -- so we should use adaptive trapezoidal method instead -- if the integrand is poorly behaved, e.g.
  - has large rapid fluctuations
  - has singularities

Perhaps a picture will help make the process clearer. This diagram shows which values $R_{i,m}$ are needed to calculate further Rs:

$$
\begin{array}{lllllll}
I_1 \equiv R_{1,1} & & & & & & \\
& \searrow & & & & & \\
I_2 \equiv R_{2,1} & \to & R_{2,2} & & & & \\
& \searrow & & \searrow & & & \\
I_3 \equiv R_{3,1} & \to & R_{3,2} & \to & R_{3,3} & & \\
& \searrow & & \searrow & & \searrow & \\
I_4 \equiv R_{4,1} & \to & R_{4,2} & \to & R_{4,3} & \to & R_{4,4} \\
& \searrow & & \searrow & & \searrow & & \searrow
\end{array}
$$

Each row here lists one trapezoidal rule estimate I_i followed by the other higher-order estimates it allows us to make. The arrows show which previous estimates go into the calculation of each new one via Eq. (5.51).

Note how each fundamental trapezoidal rule estimate I_i allows us to go one step further with calculating the $R_{i,m}$. The most accurate estimate we get from the whole process is the very last one: if we do n levels of the process, then the last estimate is $R_{n,n}$ and is accurate to order h_n^{2n}.

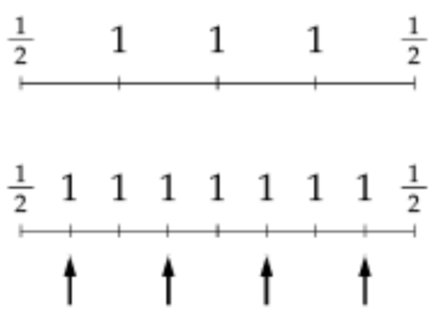


Figure 5.3: Doubling the number of steps in the trapezoidal rule. Top: We evaluate the integrand at evenly spaced points as shown, with the value at each point being multiplied by the appropriate factor. Bottom: when we double the number of steps, we effectively add a new set of points, half way between the previous points, as indicated by the arrows.