Numerical Integration

A. Donev, Spring 2021

Most integrals cannot be computed analytically using standard functions, and require introducing special functions such as the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

How do we compute

$$J = \int_a^b f(x)\, dx$$

numerically – called numerical integration or quadrature

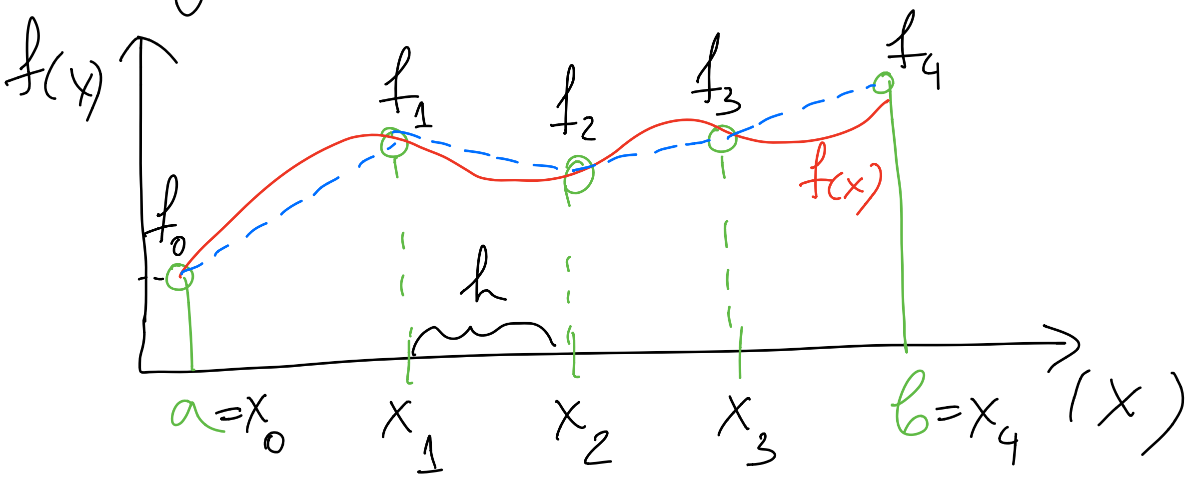
(comes from old days when gridded paper was used to count number of squares under a graph)

Simple idea:

Approximate function by (piecewise or global) polynomial and integrate polynomial

Note: Same works for numerical differentiation, as you saw in Worksheets. But, differentiation is easy so we mostly use numerical differentiation for solving differential equations (not covered in this course)

E.g. Piecewise linear interpolation:



$$x_k = a + k \cdot h$$

$$h = \frac{b-a}{n} = \text{grid spacing}$$

Since we approximate area under the graph of $f(x)$ by a sum of areas of trapezoids we call this the composite trapezoidal rule:

$$I \simeq h \cdot \left(\frac{f_0+f_1}{2} + \frac{f_1+f_2}{2} + \frac{f_2+f_3}{2} + \frac{f_3+f_4}{2} \right)$$

$$I \simeq h\left(\frac{f_0+f_n}{2}+\sum_{j=1}^{n-1} f_j\right)$$

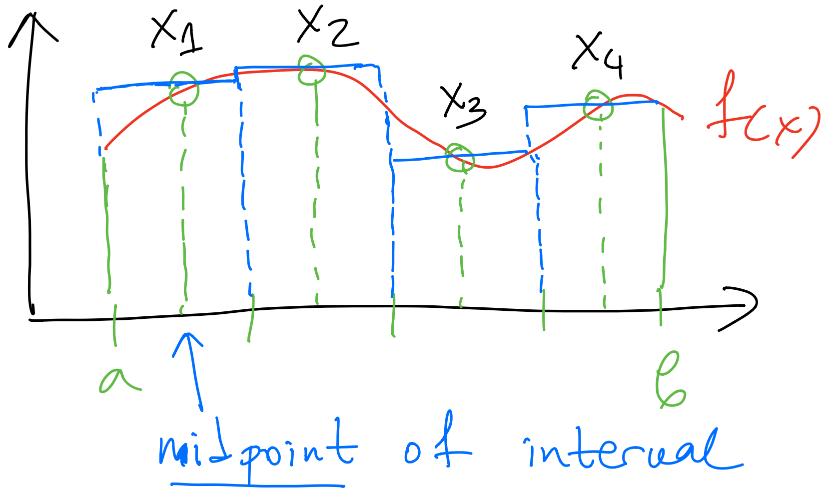
where $f_j = f(x_j)$

Usually written as

$$I = \int_a^b f(x)\,dx \simeq h\sum_{j=0}^{n} f(x_j)$$

$\boxed{\text{Trapezoidal}}$ $$- \frac{h}{2}\left(f(a)+f(b)\right)$$

It is very common to see student codes that miss this endpoint correction. If you don't want a correction at the end points, use instead the midpoint quadrature rule



Here
$$x_j = a + (j - 1/2)\,h \,, \quad j = 1, .., n$$
$$h = \frac{b-a}{n}$$

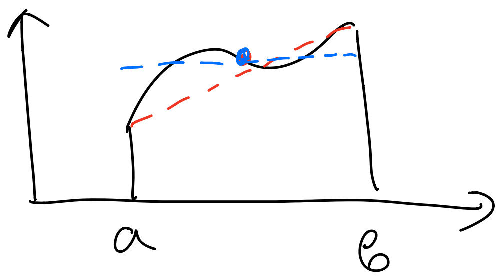
and now we approximate the area by a sum of areas of rectangles

$$I \simeq h \sum_{j=1}^{n} f(x_j) \qquad \boxed{\text{midpoint}}$$

The trapezoidal & midpoint rules are equally accurate & are the most basic formulas

How accurate are these quadrature rules?

Let's consider a <u>single</u> interval / trapezoid / rectangle first:



For midpoint rule: Expand $f(x)$ in a Taylor series around the midpoint, with remainder term, to show that

$$\mathcal{E} = \int_{x-h/2}^{x+h/2} f(t)\,dt - h f(x) = \frac{h^3}{24} f''(\xi)$$

where ξ is in inside interval.

$\boxed{\text{Do on your own!}}$

For trapezoidal rule, use formula for error of linear interpolant $p_1(x)$:

$$f(x) - p_1(x) = \frac{1}{2} f''(\xi(x))(x-a)(x-b)$$

$$\xi(x) \in [a, b]$$

$$\Rightarrow \mathcal{E} = \left(\int_a^b f(x)\,dx\right) - \left(\int_a^b p_1(x)\,dx\right)$$

$$= \frac{1}{2} \int f''(\xi(x))(x-a)(x-b)\,dx$$

$$= \frac{f''(\eta)}{2} \int_a^b (x-a)(x-b)\,dx$$

where $\eta \in [a,b]$ according to the mean-value theorem for integrals. Do intergral to get:

$$\begin{cases} \mathcal{E}_{trap} = -\dfrac{h^3}{12} f''(\xi) \\ \mathcal{E}_{mid} = \dfrac{h^3}{24} f''(\xi) \\ \xi \in [a,b], \quad b-a=h \end{cases}$$

Note these are similar to each other.

This was only <u>one</u> rectangle.

For many rectangles, i.e., for composite trapezoidal rule, in the <u>worst case scenario</u> all errors will have the <u>same</u> sign and will add up,

$$\left| I - I_{trap.} \right| = \left| -\frac{h^3}{12} \sum_{k=1}^{n} f''(\xi_k) \right|$$

where $\xi_k \in [x_{k-1}, x_k]$

Assume $|f''(x)| < M$ for $x \in [a,b]$

$$\Rightarrow e_{trap} = |I - I_{trap}| \leq \frac{h^3}{12} \cdot n \cdot M$$

$$= \frac{h^3}{12} \cdot \frac{(b-a)}{h} \cdot M$$

$$e_{trap} \leq \frac{(b-a)}{12} \cdot M \cdot h^2$$

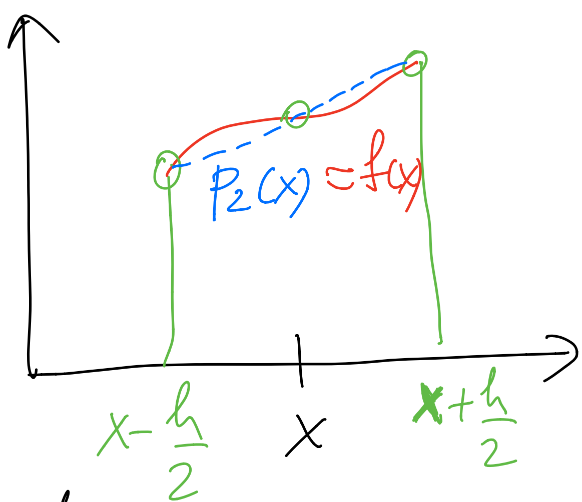
error = $O(h^2)$ for both composite trapezoidal & composite midpoint rules

Composite midpoint / trapezoidal rule is second-order accurate

This means that if you double the number of points/intervals, the error decreases by a factor of 4!

Note that this is just an upper bound on error – a much more precise estimate is provided by the Euler-MacLaurin theorem (see Wikipedia if interested)

To get more accuracy, we can use <u>piecewise quadratic interpolation</u>: approximate function by a parabola over each interval. Let's take one interval only.



$$\int_{x-h/2}^{x+h/2} f(x)\,dx \approx \int_{x-h/2}^{x+h/2} P_2(x)\,dx$$

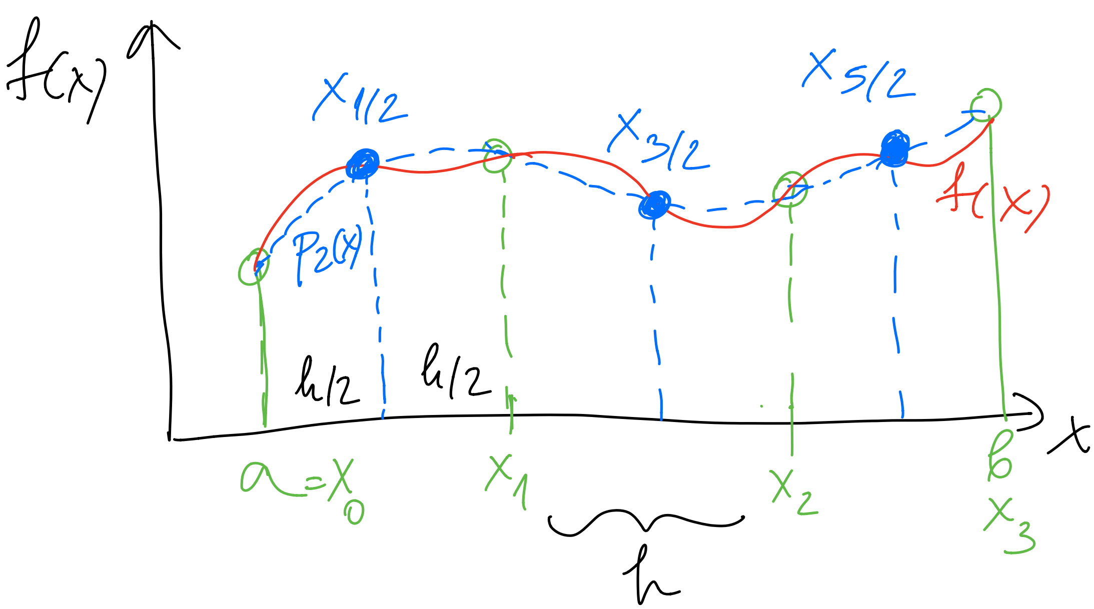
where P_2 is the quadratic interpolant.

You did this calculation in a worksheet, to get

$$\int_{x-h/2}^{x+h/2} f(x)\,dx = \frac{h}{6}\left[f\left(x-\frac{h}{2}\right) + 4f(x) + f(x+h/2)\right]$$

This is called Simpson's rule.

To make this a composite rule, split $[a, b]$ into pieces (piecewise quadratic interpolation)

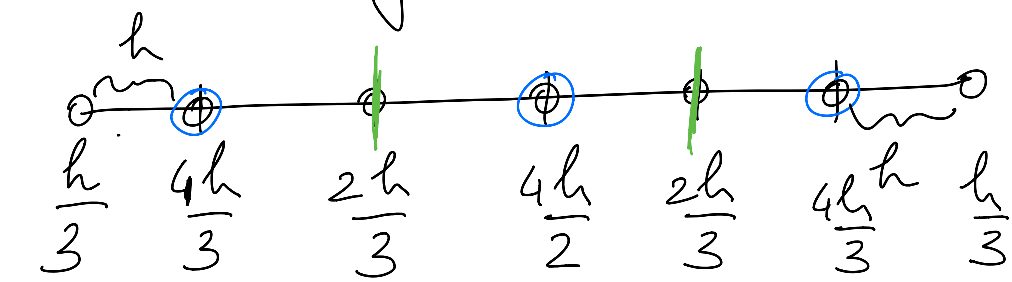


$$
\begin{aligned}
J \approx\ & \frac{h}{6}\left(f_0 + 4 f_{1/2} + f_1\right) \\
& + \frac{h}{6}\left(f_1 + 4 f_{3/2} + f_2\right) \\
& + \frac{h}{6}\left(f_2 + 4 f_{5/2}\ f_3\right) \\
=\ & \frac{h}{6}\Big[f_0 + 4 f_{1/2} + 2 f_1 + 4 f_{3/2} \\
& \quad + 2 f_2 + 4 f_{5/2} + f_3\Big]
\end{aligned}
$$

$$
\begin{aligned}
\int_a^b f(x)\,dx \approx\ & \frac{h}{6}\left(f(a) + f(b)\right) \\
& + \frac{h}{3} \sum_{k=1}^{n-1} f(x_k) \\
& + \frac{2h}{3} \sum_{k=0}^{n-1} f(x_{k+1/2})
\end{aligned}
$$

$\boxed{\text{Simpson quadrature}}$

Another way to write it



Give first and last point weight $h/3$, even points weight $4h/3$, and odd ones $2h/3$ (compare to midpoint which gave all points weight h, and trapezoidal which gave first and last point weight $h/2$ and all others weight h).

Simpson's rule is only slightly more complicated but 4th order accurate so use it!

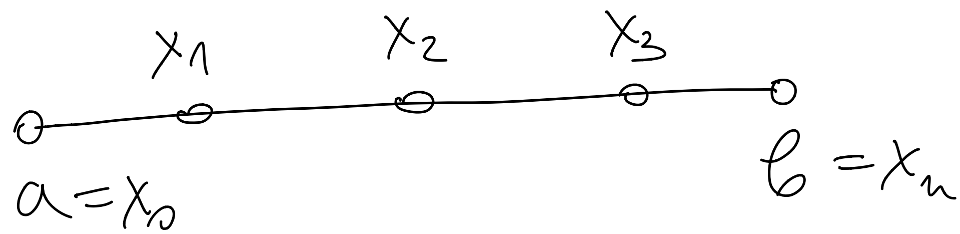
$$\mathcal{E}_{simp} = \left| \mathcal{I} - \mathcal{I}_{simp} \right| = \frac{b-a}{2880} h^4 M$$

$$M = \max_{a \le x \le b} \left| f^{(4)}(x) \right|$$

Double number of points, 16 times smaller error!

We can get even better accuracy by using higher-degree polynomial approximants.

Let's focus on only one interval, but of course one can always split $[a, b]$ into pieces for composite quadrature



Choose interpolation nodes in $[a, b]$ (not necessarily equi-spaced!) and use Lagrange interpolation

$$f(x) \approx p_n(x) = \sum_{j=0}^{n} f(x_j) L_j(x)$$

Lagrange polynomial

$$\Rightarrow \int_a^b f(x)\,dx \approx \int_a^b p_n(x)\,dx$$

$$= \int_a^b f_j L_j(x)\,dx$$

$$\int_a^b f(x)\,dx \simeq \sum_{j=0}^{n} \underbrace{\left(\int_a^b L_j(x)\,dx\right)}_{w_j} f_j$$

$$I = \int_a^b f(x)\,dx \simeq \sum_{j=0}^{n} w_j\, f(x_j)$$

w_j: quadrature weights

x_j: quadrature nodes

$$w_j = \int_a^b L_j(x)\,dx$$

All quadrature rules have this form, but different weights & nodes. Note that once we choose the nodes we can <u>pre-compute the weights $\vec{w}$</u>

(see Newton-Cotes quadrature on Wiki for tables)

The midpoint, trapezoidal & Simpson's rule were just special cases for $n=0$, $n=1$ and $n=2$. For these small n it is obvious where to place the nodes by symmetry, if we want to include the endpoints.

But what is the best choice of nodes & weights?