L29: Numerical Integration

Lucas A. J. Bastien

E7 Spring 2017, University of California at Berkeley

April 5, 2017

Version: release

Announcements

Lab 10 is due on April 7 at 12 pm (noon)

Question 2, test case 1: it is okay if `coefficients(4,5)` is `0`

`coefficients(4,5)` is correct if `abs(coefficients(4,5)) < 1e-15`

Today:

- Numerical Integration (Chapter 18)

Friday (April 7):

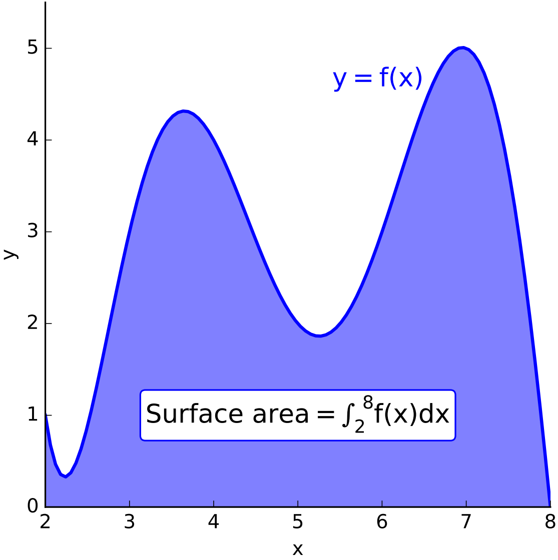
- Project discussion, tips, and recommendations
- Other discussion

Next week:

- Ordinary differential equations (Chapter 19)

Integral: area under the curve

Geometrically, the integral is equal to the "area under the curve"



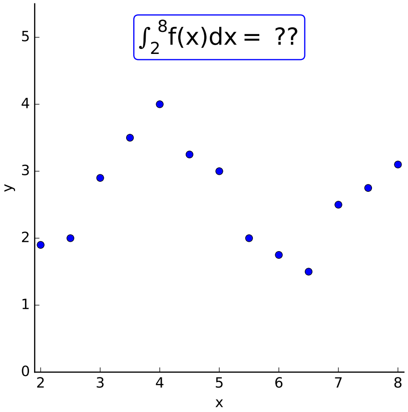
Numerical integration: motivation

We know a number of analytical methods to calculate integrals
For example:

$$\int_0^4 (x^3 - x^2)dx = \left[\frac{x^4}{4} - \frac{x^3}{3}\right]_0^4 = \frac{4^4}{4} - \frac{4^3}{3}$$

But sometimes:

- **We do not know how to calculate the integral analytically**; or
- **We know the value of the function only at certain locations**



General approach to numerical integration

Objective: find an approximation of the following quantity:

$$\int_a^b f(x)dx$$

where f is a real-valued function

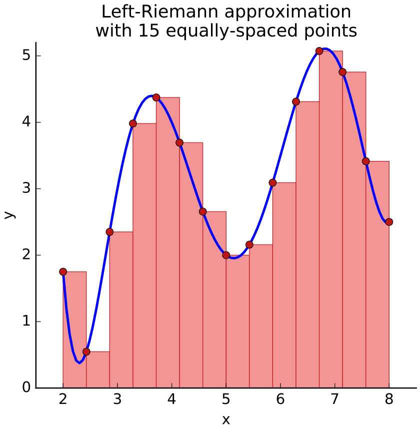
General approach:

1. Divide the interval $[a, b]$ into n intervals: $[x_i, x_{i+1}]$,
   $i \in \{0, 1, 2, \ldots, n-1\}$
   (Note: $x_0 = a$ and $x_n = b$)
2. On each of these "sub-intervals", approximate the integral using approximations of the function (*e.g.*, straight line)
3. Sum the approximated values made on each sub-interval

Method 1: Riemann integral

On each interval $[x_i, x_{i+1}]$, approximate the function as a constant:

- Either: $f(x) \approx f(x_i)$ on this interval ("left-Riemann integral")
- Either: $f(x) \approx f(x_{i+1})$ on this interval ("right-Riemann integral")



Method 1: Riemann integral – the math

$$\int_a^b f(x)dx = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx$$
$$\approx \sum_{i=0}^{n-1} \text{Area of the corresponding rectangle}$$

Left-Riemann integral:

$$= \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i)$$

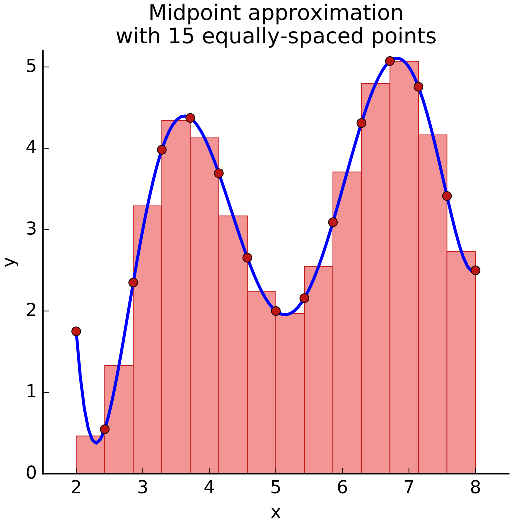
Right-Riemann integral:

$$= \sum_{i=0}^{n-1} f(x_{i+1})(x_{i+1} - x_i)$$

Method 2: Midpoint rule

On each interval $[x_i, x_{i+1}]$, approximate the function as a constant:

- $f(x) \approx f(\mathrm{midpoint})$ on this interval



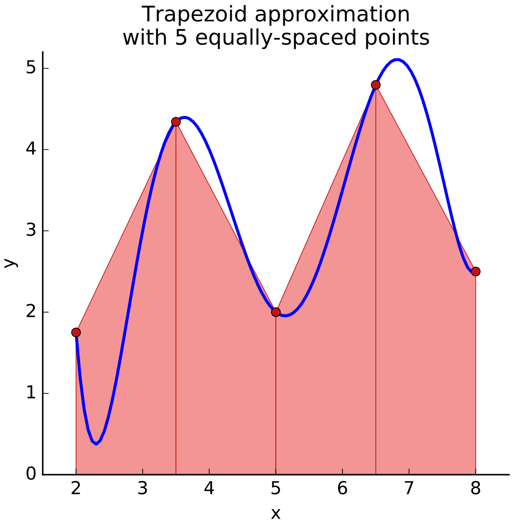
Method 2: Midpoint rule – the math

$$\begin{aligned}
\int_a^b f(x)dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx \\
&\approx \sum_{i=0}^{n-1} \text{Area of the corresponding rectangle} \\
&= \sum_{i=0}^{n-1} f\left(\frac{x_i + x_{i+1}}{2}\right)(x_{i+1} - x_i)
\end{aligned}$$

Limitation: this method requires us to calculate the value of the function at the midpoint of each interval. **We cannot use this method if we only know the value of the function at $x = x_i$, $i = \{0, 1, 2, \ldots, n\}$**

Method 3: Trapezoid rule

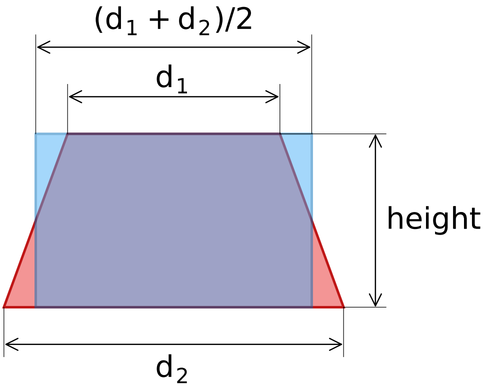
On each interval $[x_i, x_{i+1}]$, approximate the function as a straight line that goes through both (x_i, y_i) and (x_{i+1}, y_{i+1}):



Area of a Trapezoid

Trapezoid: a quadrilateral with two sides parallel to each other

Area of a trapezoid: height $\times$ average of the lengths of the bases



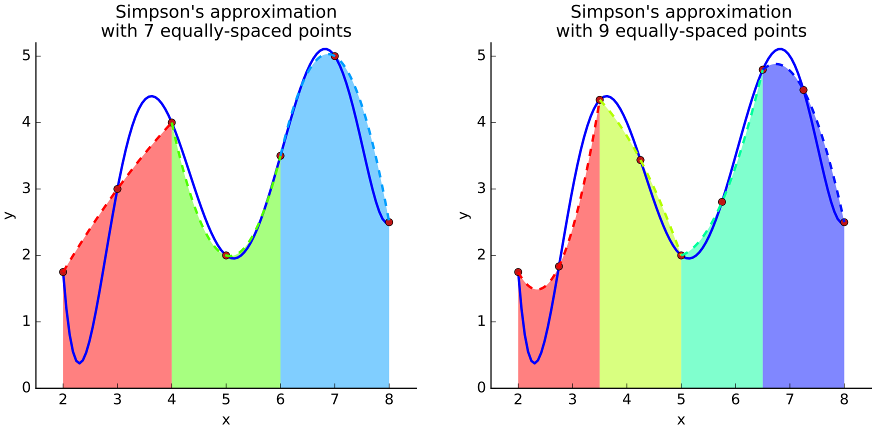
Method 3: Trapezoid rule – the math

$$
\begin{aligned}
\int_a^b f(x)dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx \\
&\approx \sum_{i=0}^{n-1} \text{Area of the corresponding trapezoid} \\
&= \sum_{i=0}^{n-1} \frac{f(x_i) + f(x_{i+1})}{2}(x_{i+1} - x_i) \\
&\text{If the points are equally spaced, with spacing } \Delta x\text{:} \\
&= \frac{\Delta x}{2}\left(f(x_0) + 2\sum_{i=1}^{n-1} f(x_i) + f(x_n) \right)
\end{aligned}
$$

See the textbook for the derivation

Method 4: Simpson's rule

- Group the sub-intervals $[x_i, x_{i+1}]$ two by two
- On each of these pairs of intervals:
  - Fit a parabola
  - Calculate (analytically) the corresponding area under the parabola



Method 4: Simpson's rule – the math

$$\int_a^b f(x)dx = \sum_{i=0}^{n/2-1} \int_{x_{2i}}^{x_{2i+2}} f(x)dx$$

$$\approx \sum_{i=0}^{n/2-1} \text{Area under the corresponding piece of parabola}$$

If the points are equally spaced, with spacing Δx:

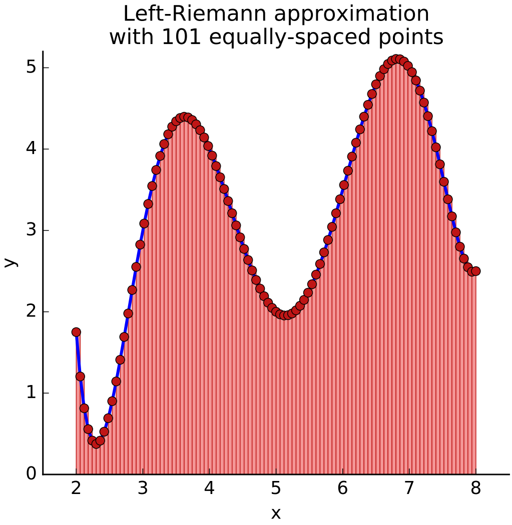
$$= \frac{\Delta x}{3}\left(f(x_0) + 4 \sum_{\substack{i=1 \\ i \text{ is odd}}}^{n-1} f(x_i) + 2 \sum_{\substack{i=2 \\ i \text{ is even}}}^{n-2} f(x_i) + f(x_n) \right)$$

Limitation: the number of sub-intervals $[x_i, x_{i+1}]$ must be even (*i.e.* the number of data points must be odd)

See the textbook for the derivation

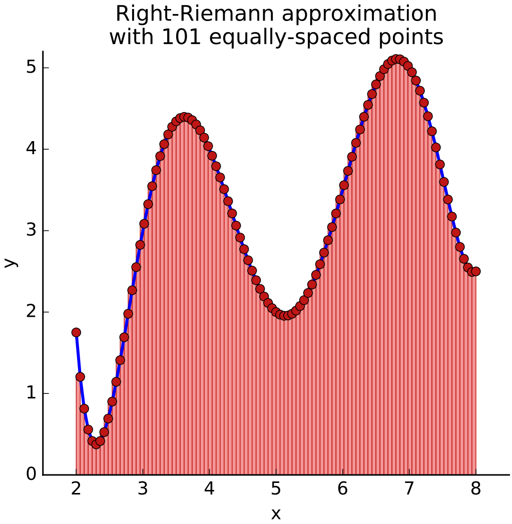
Increase the number of sub-intervals

In general, integral approximations become more accurate as the division of the interval into sub-intervals is refined



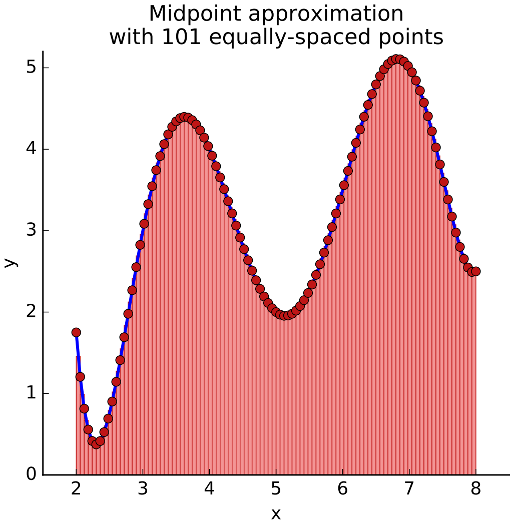
Increase the number of sub-intervals

In general, integral approximations become more accurate as the division of the interval into sub-intervals is refined



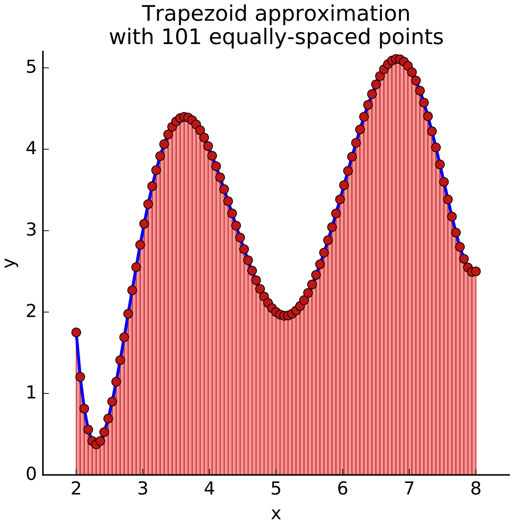
Increase the number of sub-intervals

In general, integral approximations become more accurate as the division of the interval into sub-intervals is refined



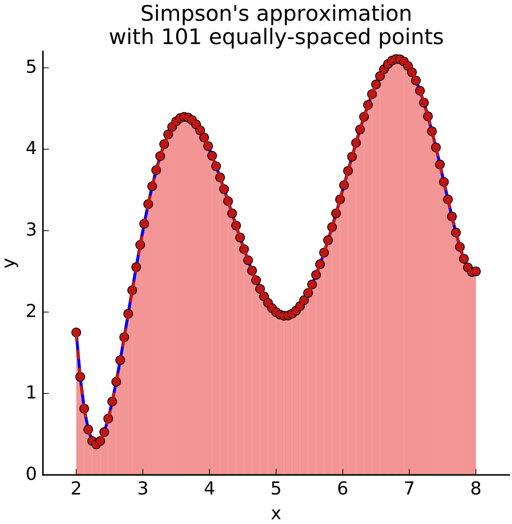
Increase the number of sub-intervals

In general, integral approximations become more accurate as the division of the interval into sub-intervals is refined

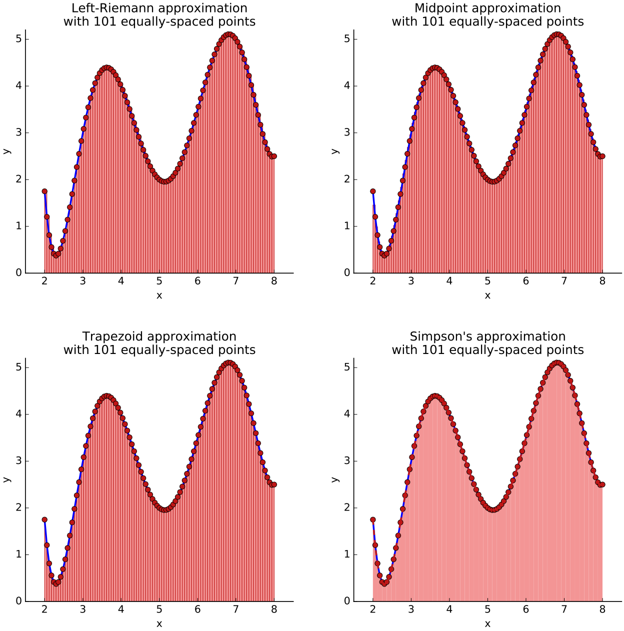


Increase the number of sub-intervals

In general, integral approximations become more accurate as the division of the interval into sub-intervals is refined



Comparison of different integral approximations



Order of the error of the integral approximations

If the error made on the approximation of the integral on one of the sub-intervals $[x_i, x_{i+1}]$ is of order $\mathcal{O}((\Delta x)^m)$, then what is the order of the error made on the integral over the entire interval $[a, b]$?

$$
\begin{aligned}
\text{Overall order} &= \sum_{i=0}^{n-1} \text{individual orders} \\
&= \sum_{i=0}^{n-1} \mathcal{O}((\Delta x)^m) \\
&\approx n \times \mathcal{O}((\Delta x)^m) \\
&= \frac{b-a}{\Delta x} \times \mathcal{O}((\Delta x)^m) \\
&= \mathcal{O}((\Delta x)^{m-1})
\end{aligned}
$$

Note: the explanation presented here is not a rigorous mathematical proof

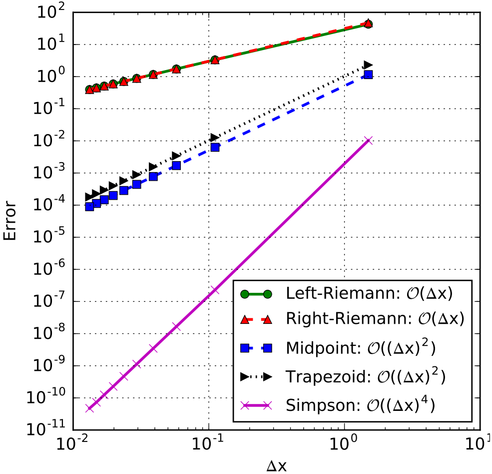
Order of the error of the integral approximations

Method	Order of the error on each sub-interval	Order of the error on the overall interval
Left-Riemann integral	$\mathcal{O}((\Delta x)^2)$	$\mathcal{O}(\Delta x)$
Right-Riemann integral	$\mathcal{O}((\Delta x)^2)$	$\mathcal{O}(\Delta x)$
Midpoint rule	$\mathcal{O}((\Delta x)^3)$	$\mathcal{O}((\Delta x)^2)$
Trapezoid rule	$\mathcal{O}((\Delta x)^3)$	$\mathcal{O}((\Delta x)^2)$
Simpson's rule	$\mathcal{O}((\Delta x)^5)$	$\mathcal{O}((\Delta x)^4)$

See the textbook for the derivations

Order of the error of the integral approximations

Error versus Δx for different integration approximations, when calculating the integral of $x \mapsto \sin(x) + x^2$ from 2 to 8



The slope of the line in a log-log plot indicates the order of the method

IMPORTANT practice question

We use a 2^{nd}- and a 4^{th}-order integration approximation to estimate the integral of a function, using equally-spaced points (spacing is Δx).

Which of the following statements are true?

(A) The error made when using the 4^{th}-order method is always smaller than the error made when using the 2^{nd}-order method

(B) On average, if we reduce Δx by a factor of 2, the error made when using the 2^{nd}-order method is divided by 4

(C) On average, if we reduce Δx by a factor of 2, the error made when using the 4^{th}-order method is divided by 4

(D) On average, if we reduce Δx by a factor of 2, the error made when using the 4^{th}-order method is divided by 16

(E) The error made when using the 4^{th}-order method is always twice as small as the error made when using the 2^{nd}-order method