Lecture 22

Integration: Newton Cotes

Owen L. Lewis

Department of Mathematics and Statistics
University of New Mexico

Nov. 14, 2024

This lecture

- (Done) We can interpolate $f(x)$
- (Done) We can differentiate $f(x)$
- Next, we look at integrating $f(x)$

Why?

- Often $f(x)$ is only known via samples (known at a certain number of points).

Given a collection of points (knots) $x_0, x_1, \ldots x_n$ and the value of some function $y_0 = f(x_0), y_1 = f(x_1), \ldots y_n = f(x_n)$.

- Often the anti-derivative of $f(x)$ is not known.

We have a formula for $f(x)$, but not $F(x) = \int f(x)\,dx$. Or we can evaluate $f(x)$, but not $F(x)$.

$$\text{Example: } f(x) = e^{-x^2}, \quad \int e^{-x^2}\,dx = \operatorname{erf}(x)?$$

Integrals

We seek a way to calculate to the following quantity

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

The Fundamental Theorem of Calculus states that

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

where F is the antiderivative of f. We don't know F, so we approximate the integral operation.

Integration

What is the (definite) integral $\int_a^b$?

Integration

What is the (definite) integral $\int_a^b$?

- Let P be a partition of $[a, b]$ into $n+1$ distinct and ordered points with $x_0 = a$ and $x_n = b$.
- For interval $[x_i, x_{i+1}]$ let m_i be a lower bound for $f(x)$ on that interval
- For interval $[x_i, x_{i+1}]$ let M_i be an upper bound for $f(x)$ on that interval
- Lower Sum:
$$L(f; P) = \sum_{i=0}^{n-1} m_i(x_{i+1} - x_i)$$
- Upper Sum:
$$U(f; P) = \sum_{i=0}^{n-1} M_i(x_{i+1} - x_i)$$

Integration

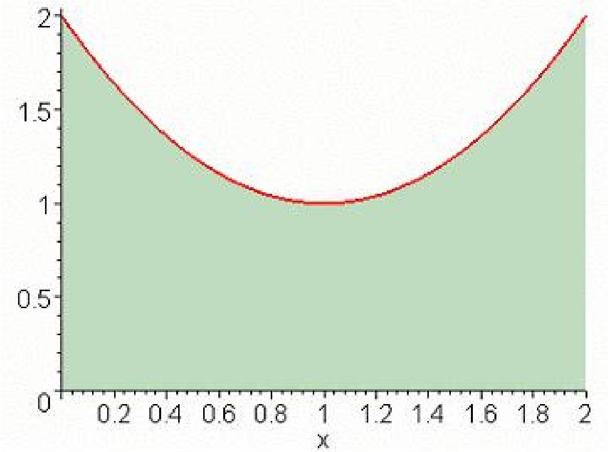
- The low sum always under-approximates the integral
- The upper sum always over-approximates the integral

$$L(f;P) \leqslant \int_a^b f(x)\,dx \leqslant U(f;P)$$

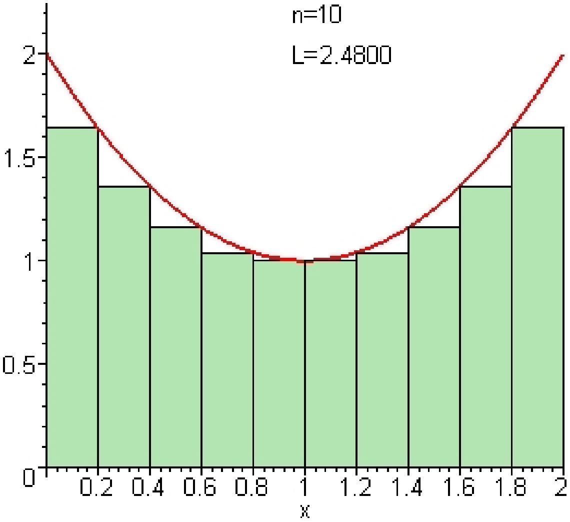
- In the limit, they are equal

$$\lim_{n\to\infty} L(f;P) = \int_a^b f(x)\,dx = \lim_{n\to\infty} U(f;P)$$

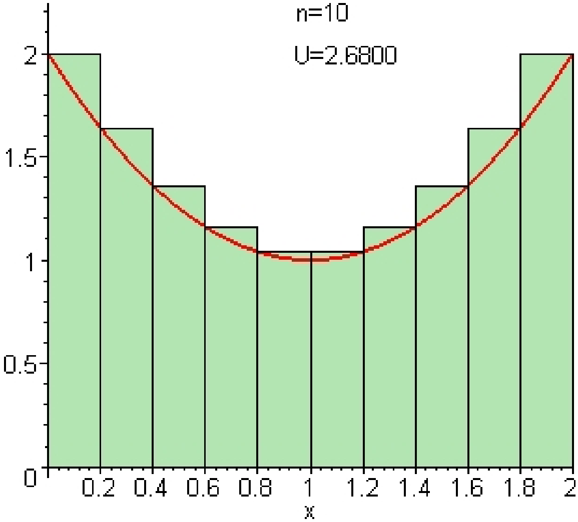
Graphically: Integral



Graphically: Lower sum



Graphically: Upper sum



Left-Riemann, Right-Riemann, Mid-Point

- The upper and lower bounds are often difficult to identify (how do we know m_i or M_i?)
- Use Left-Riemann, Right-Riemann, and Middle Riemann Sums
- Generally the Riemann sum is

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

for $x_i \leqslant z_i \leqslant x_{i+1}$

- $z_i = x_i$ is a Left Riemann Sum
- $z_i = x_{i+1}$ is a Right Riemann Sum
- $z_i = \frac{x_{i+1}+x_i}{2}$ is a Middle Riemann Sum

Goals

We have a way to compute integrals. Why aren't we done?

We don't know if this is a *good* or *fast* way to compute integrals. We also don't know what the error is.

Additionally, we may need a very efficient algorithm to meet a real-time requirement.

Some options that we'll look at are

- Trapezoid Rule
- Composite Trapezoid Rule
- Simpson's Rule
- Composite Simpson's Rule
- General Newton-Cotes Rules

Trapezoid

Goal: Approximate

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

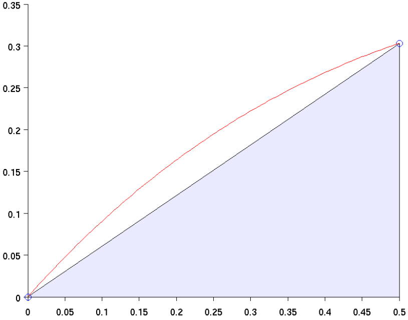
Goal: Approximate area under $f(x)$.

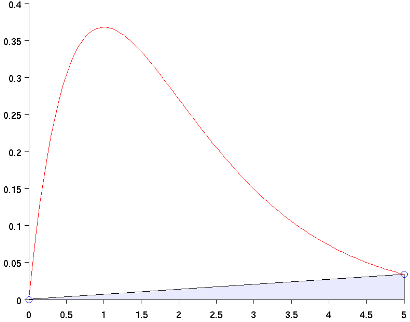
- Old Idea: Left,Right,Midpoint Riemann integration says: Approximate $f(x)$ by a constant function and obtain the area under the constant function.
- New Idea: Trapezoid approximates $f(x)$ by a linear function (degree one polynomial) and obtains the area under the linear function.

Basic Trapezoid

Use endpoints $[a, b]$ to obtain a linear approximation to $f(x)$. The area under this function is the area of a trapezoid:

$$\int_a^b f(x)\,dx \approx \frac{1}{2}(b-a)(f(a)+f(b))$$





Basic Trapezoid

- Trapezoid Rule:

$$\int_{x_1}^{x_2} f(x)\, dx \approx \int_{x_1}^{x_2} p_1(x)\, dx = \frac{1}{2}(y_1 + y_2)h$$

- Where h is the spacing $x_2 - x_1$

Example

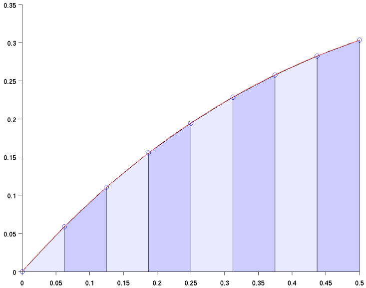
$$\int_1^2 15\, x^2 \approx \frac{1}{2}(15 * 1^2 + 15 * 2^2) * 1$$
$$= \frac{1}{2}(15 + 60) = 37.5$$

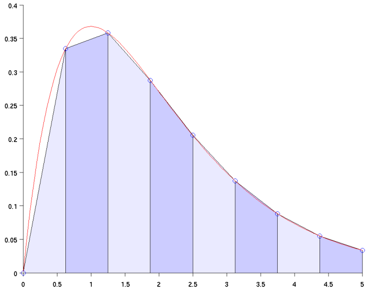
- Analytical answer is $\int_1^2 15\, x^2 = 5\, x^3\big|_1^2 = 40 - 5 = 35.$

Composite Trapezoid

1. Obviously a linear approximation won't cut it, especially for long intervals
2. Use a linear spline and integrate that
3. Consider a partition $P = \{x_0 = a, x_1, x_2, \ldots, x_n = b\}$ of $[a, b]$.
4. In each interval $[x_i, x_{i+1}]$, use the basic Trapezoid:

$$\int_a^b f(x)\,dx \approx \sum_{i=0}^{n-1} \frac{1}{2}(x_{i+1} - x_i)(f(x_i) + f(x_{i+1}))$$





Notes

Composite Trapezoid

- With uniform spacing of P, $h_i = x_{i+1} - x_i = h$ is constant

$$T(f;P) = \int_a^b f(x)\,dx \approx \frac{h}{2}\sum_{i=0}^{n-1}\left(f(x_i) + f(x_{i+1})\right)$$

- This becomes

$$\begin{aligned} T(f;P) &= \int_a^b f(x)\,dx \approx \frac{h}{2}\left(f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)\right) \\ &= \int_a^b f(x)\,dx \approx h\left(\frac{1}{2}f(x_0) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{1}{2}f(x_n)\right) \end{aligned}$$

```
h = (b - a)/n
sum = (f(a) + f(b))/2
for i = 1 to n - 1
    sum = sum + f(x_i)
end
sum = sum · h
```

Example: trap_int_test.m

Test composite trapezoid for

$$\int_0^5 xe^{-x}\,dx$$

Question: What is the order of accuracy? What do we expect for a linear approximation to $f(x)$?

Find h^p numerically with

$$p \approx \frac{log(err^{(k)}/err^{(k-1)})}{log(h^{(k)}/h^{(k-1)})}$$

Run `trap_int_test.m`

Notes

Accuracy

So composite Trapezoid appears to be order 2. Why? Look first at basic Trapezoid:

$$\int_a^b f(x)\,dx \approx \frac{1}{2}(b-a)(f(a)+f(b))$$

Looking at the error

$$\begin{aligned}
E &= \int_a^b f(x) - \frac{(b-x)f(a)}{b-a} - \frac{(x-a)f(b)}{b-a}\,dx \quad \text{(Use Poly. Int. Err. formula)}\\
&= \frac{1}{2}\int_a^b f''(\xi(x))(x-a)(x-b)\,dx \quad \text{note:}(x-a)(x-b) \text{ single signed in } [a,b]\\
&= \frac{f''(\eta)}{2}\int_a^b (x-a)(x-b)\,dx \quad \text{(MVT for Integrals)}\\
&= \frac{f''(\eta)}{2}\left(-\frac{1}{6}(b-a)^3\right)\\
&= -\frac{(b-a)^3 f''(\eta)}{12}
\end{aligned}$$

Accuracy

What about Composite Trapezoid?

$$T(f;P) = \int_a^b f(x)\,dx \approx \frac{h}{2}\sum_{i=0}^{n-1}\left(f(x_i) + f(x_{i+1})\right)$$

The error in each interval $[x_i, x_{i+1}]$ is

$$E_i = -\frac{h^3 f''(\eta_i)}{12}$$

So the total error is

$$\begin{aligned}
\sum_{i=0}^{n-1} E_i = \sum_{i=0}^{n-1} -\frac{h^3 f''(\eta_i)}{12} &= -n\frac{h^3}{12}\sum_{i=0}^{n-1}\frac{1}{n}f''(\eta_i) \\
&= -n\frac{h^3 f''(\eta)}{12}\ (IVT), \quad \text{but } nh = b-a, \\
&= -\frac{(b-a)h^2 f''(\eta)}{12} = O(h^2)
\end{aligned}$$

Example

How many points should be used to ensure the composite Trapezoid rule is accurate to 10^{-6} for $\int_0^1 e^{-x^2}\,dx$? Need

$$\frac{(b-a)h^2 f''(\eta)}{12} \leqslant 10^{-6}$$

How big is $f''(x)$?

$$\begin{aligned} f(x) &= e^{-x^2} \\ f'(x) &= -2xe^{-x^2} \\ f''(x) &= -2e^{-x^2} + 4x^2e^{-x^2} \\ f'''(x) &= 12xe^{-x^2} - 8x^3e^{-x^2} \end{aligned}$$

So f''' is always positive. So f'' is monotone increasing and thus f'' takes on a maximum at an endpoint: $f''(0) = -2 \quad \Rightarrow \quad |f''(x)| \leqslant 2$. Then, bound becomes

$$\frac{(b-a)2h^2}{12} \leqslant 10^{-6}$$

Or, using the relation $(b-a)/n = h$,

$$h^2 \leqslant 6 \times 10^{-6} \quad \Rightarrow \quad \sqrt{(1/6)}10^3 \leqslant n$$

or $n > 410$.

How do we improve Trapezoid?

- Instead of a linear approximation, use a quadratic approximation
- $\Rightarrow$ Simpson's Rule

Simpson

- Consider $\int_a^b f(x)\,dx$
- Partition $P = \{a, a+h, b = a+2h\}$
- Or, to simplify things, consider $\int_0^{2h} f(x)\,dx$
- With the partition $P = \{0, h, 2h\}$
- Replace $f(x)$ by a quadratic $p(x)$:

$$\begin{aligned} f(x) &\approx p(x) \\ &= f(0) + \frac{f(h) - f(0)}{h}x + \frac{f(2h) - 2f(h) + f(0)}{2h^2}x(x-h) \\ &= \text{Newton form} \end{aligned}$$

- Integrate $\int_0^{2h} p(x)\,dx$:

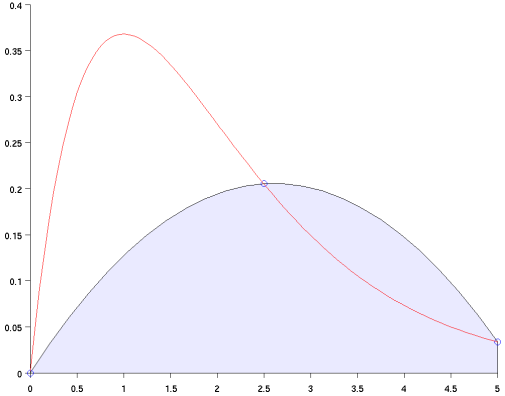
$$\begin{aligned} \int_0^{2h} f(x)\,dx &\approx \int_0^{2h} p(x)\,dx \\ &= \frac{h}{3}\left[f(0) + 4f(h) + f(2h)\right] \end{aligned}$$

Simpson

Since $b - a = 2h$ we have

Basic Simpson's Rule

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right]$$



Composite Simpson

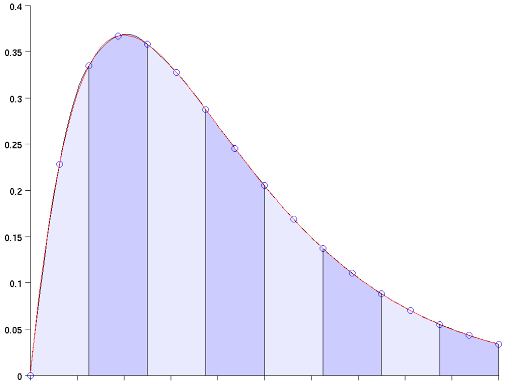
Over a uniform partition $P = x_0, x_1, \ldots, x_n$, use basic Simpson's Rule over each subinterval $[x_{2i}, x_{2i+2}]$

$$\begin{aligned}
\int_a^b f(x)\,dx &= \sum_{i=0}^{n/2-1} \int_{x_{2i}}^{x_{2i+2}} f(x)\,dx \\
&= \sum_{i=0}^{n/2-1} \frac{2h}{6}\left[f(x_{2i}) + 4f(x_{2i+1}) + f(x_{2i+2})\right] \\
&= \frac{h}{3}\left[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \cdots + 4f(x_{n-1}) + f(x_n)\right]
\end{aligned}$$

Simpson

Composite Simpson's Rule

$$\int_a^b f(x)\,dx \approx \frac{h}{3}\left[f(a) + f(b) + 4\sum_{i=1}^{n/2} f(a+(2i-1)h) + 2\sum_{i=1}^{n/2-1} f(a+2ih) \right]$$



Simpson

How accurate is Simpson?

Recall composite Trapezoid error, for exact integral I:

$$error = I - T(f, P) = -\frac{1}{12}(b-a)h^2 f''(\xi) = \mathcal{O}(h^2)$$

Prediction? $\mathcal{O}(h^2)$? $\mathcal{O}(h^3)$? $\mathcal{O}(h^4)$? Remember, we are using a quadratic approximation, whereas Trapezoid used linear.

Run `simpson_int_test.m`, and numerically observe the p value.

Note, even using just $N = 500$ points for the first approximation, makes us run into machine precision. It's an accurate method!

Why is composite Simpson $\mathcal{O}(h^4)$?

First analyze basic Simpson's rule by Taylor Series: (around a)

$$f(a+h) = f + hf' + \frac{1}{2!}h^2f'' + \frac{1}{3!}h^3f''' + \frac{1}{4!}h^4f^{(4)} + \frac{1}{5!}h^5f^{(5)} + \dots$$

$$f(a+2h) = f + 2hf' + 2h^2f'' + \frac{4}{3}h^3f''' + \frac{2}{3}h^4f^{(4)} + \frac{4}{15}h^5f^{(5)} + \dots$$

This gives

$$\frac{h}{3}\left[f(a) + 4f(a+h) + f(b)\right] = 2hf + 2h^2f' + \frac{4}{3}h^3f'' + \frac{2}{3}h^4f''' + \frac{5}{18}h^5f^{(4)}$$

Why is composite Simpson $\mathcal{O}(h^4)$?

Consider the antiderivative of $f(x)$:

$$F(x) = \int_a^x f(t)\,dt$$

Taylor series of F:

$$F(a+2h) = F(a) + 2hF'(a) + 2h^2F''(a) + \frac{4}{3}h^3F'''(a) + \frac{2}{3}h^4F^{(4)} + \frac{4}{15}h^5F^{(5)} + \cdots$$

Noting that $F(a+2h) = \int_a^{b=a+2h} f(x)\,dx$, $F(a) = 0$, $F' = f$, $F'' = f'$ and so on,

$$\int_a^{b=a+2h} f(x)\,dx = 2hf + 2h^2f' + \frac{4}{3}h^3f'' + \frac{2}{3}h^4f''' + \frac{4}{15}h^5f^{(4)} + \cdots$$

Comparing this equation with the one on previous slide, basic Simpson's Rule gives an error of

$$-\frac{1}{90}\left(\frac{b-a}{2}\right)^5 f^{(4)}(\xi)$$

Why is composite Simpson $\mathcal{O}(h^4)$?

Error for one interval using basic Simpson's Rule:

$$-\frac{1}{90}\left(\frac{b-a}{2}\right)^5 f^{(4)}(\xi)$$

Over $n/2$ subintervals $[x_{2i}, x_{2i+2}]$ (assuming equispaced points) becomes:

$$\begin{aligned}
err &= \sum_{i=1}^{n/2} -\frac{1}{90}\left(\frac{x_{2i+2}-x_{2i}}{2}\right)^5 f^{(4)}(\xi_i) = -\frac{1}{90}\sum_{i=1}^{n/2}\left(\frac{2h}{2}\right)^5 f^{(4)}(\xi_i) \\
&= -\frac{1}{90}\frac{n}{2}h^5 f^{(4)}(\xi) = -\frac{1}{180}\frac{(b-a)}{h}h^5 f^{(4)}(\xi) \\
&= -\frac{b-a}{180}h^4 f^{(4)}(\xi)
\end{aligned}$$

Composite Simpson's Rule

$$-\frac{b-a}{180}h^4 f^{(4)}(\xi)$$

We "gain" two orders over Trapezoid – got lucky!

Summary

Summary:

- Left/Right Riemann: approximate $f(x)$ by 0-degree $p(x)$ and integrate
- Trapezoid: approximate $f(x)$ by 1-degree $p(x)$ and integrate
- Simpson: approximate $f(x)$ by 2-degree $p(x)$ and integrate