# Numerical Analysis: Trapezoidal and Simpson's Rule

Natasha S. Sharma, PhD

# Mathematical question we are interested in numerically answering



- How to we evaluate

$$I = \int_a^b f(x)\ dx?$$

- Calculus tells us that if F(x) is the antiderivative of a function $f(x)$ on the interval $[a, b]$, then

$$I = \int_a^b f(x)\ dx = F(x)|_a^b = F(b) - F(a).$$

- Practically, most integrals cannot be evaluated using this approach. For example,

$$\int_0^1 \frac{dx}{1+x^5}$$

has a complicated antiderivative and it easier to adopt a numerical method to approximate this integral.

# Numerical Integration: A General Framework

If you cannot solve a problem, then replace it with a "near-by" problem that you can solve!

Our problem:

$$\text{Evaluate } I = \int_a^b f(x)\, dx.$$

To do so, many of the numerical schemes are based on replacing $f(x)$ with some approximate function $\tilde{f}(x)$ so that

$$I \approx \int_a^b \tilde{f}(x)\, dx = \tilde{I}.$$

Example: $\tilde{f}(x)$ could be an easy to integrate function approximating $f(x)$.

# Numerical Integration: A General Framework



Then, the approximation error in this case is

$$E = I - \tilde{I} = \int_a^b \left(f(x) - \tilde{f}(x)\right)\, dx$$
$$\leq (b-a) \max_{a \leq x \leq b} |f(x) - \tilde{f}(x)|$$

The inequality above tells us that the approximation error $E$ depends on:

1. the maximum error in the approximating $f(x)$ that is $\max_{a \leq x \leq b} |f(x) - \tilde{f}(x)|$, and
2. $(b-a)$, the width of the interval.

Next Goal: How to choose $\tilde{f}(x)$?

# Polynomial Approximations to $f(x)$



Goal Choose an approximation $\tilde{f}(x)$ to $f(x)$ that is easily integrable and a good approximation to $f(x)$.

Two natural candidates:

1. Taylor polynomials approximating $f(x)$.
   One caveat: We need $f(x)$ to have derivatives at "a" to exist of a higher order to improve the approximation!
2. Interpolating polynomials approximating $f(x)$.

# Example

## Example

Consider evaluating

$$I = \int_0^1 e^{x^2}\, dx$$

Use the Taylor expansion to approximate $f(x) = e^{x^2}$.
That is,

$$f(x) = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!} + \underbrace{\frac{t^{n+1}}{(n+1)!}e^c}_{\text{remainder term } R_n(x)}, \quad t = x^2,$$

where $c$ is an unknown number between 0 and $t = x^2$.

**Solution:**

$$I = \int_0^1 \left(1 + x^2 + \frac{x^4}{2!} + \cdots + \frac{x^{2n}}{n!}\right) dx + \int_0^1 \frac{x^{2(n+1)}}{(n+1)!} e^c \, dx.$$

Taking $n = 3$, we have

$$I = 1 + \frac{1}{3} + \frac{1}{10} + \frac{1}{42} + E = 1.4571 + E,$$

where $E = \int_0^1 \frac{x^{2(n+1)}}{(n+1)!} e^c \, dx$ and we need a bound on this "remainder" term.

$$0 < E \leq \frac{e}{24} \int_0^1 x^8 \, dx = \frac{e}{216} = 0.0126.$$

# Using Interpolating Polynomials

In spite of the simplicity of the above example, it is generally more difficult to do numerical integration by constructing Taylor polynomial approximations than by constructing polynomial interpolates.
Thus, we construct the function $\tilde{f}(x)$ as the polynomial interpolating $f(x)$ such that

$$\int_a^b f(x)\,dx \approx \int_a^b \tilde{f}(x)\,dx.$$

Case I: Use linear interpolating polynomial $p_1(x)$ approximating $f(x)$ at two points. We pick $a$ and $b$.

# Using Linear Interpolating Polynomails



$\tilde{f}(x) = p_1(x)$ where

$$p_1(x) = \frac{(b-x)f(a) + (x-a)f(b)}{b-a}.$$

Thus,

$$\begin{aligned}
\int_a^b f(x)\ dx \approx \int_a^b p_1(x)\ dx \\
\int_a^b \frac{(b-x)f(a) + (x-a)f(b)}{b-a}\ dx \\
= \frac{b-a}{2}\Big[f(a) + f(b)\Big] \equiv T_1(f)
\end{aligned}$$

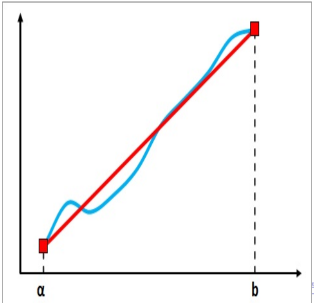
# Trapezoidal Rule

## Definition (Trapezoidal Rule)

The integration rule

$$\int_a^b f(x)\, dx \approx \frac{b-a}{2}\Big[f(a) + f(b)\Big] = T_1(f)$$

is called the trapezoidal rule.



# Example using Trapezoidal Rule

## Example

Evaluate

$$\int_0^{\pi/2} \sin x \; dx$$

using the trapezoidal rule.

$$\int_0^{\pi/2} \sin x \; dx \approx \frac{\pi}{4}\Big[\sin 0 + \sin(\pi/2)\Big] = \pi/4 \approx 0.785.$$

Since we know the true value
$I = \int_0^{\pi/2} \sin x \; dx = -\cos(\pi/2) + \cos(0) = 1,$

$$\int_0^{\pi/2} \sin x \; dx - T_1(f) = 1 - 0.785 = 0.215.$$

# How to improve the accuracy of the integration rule?



An intuitive solution is to improve the accuracy of $f(x)$ by applying the Trapezoidal rule on smaller subintervals of $[a, b]$ instead of applying it to the original interval $[a, b]$ that is apply it to integrals of $f(x)$ on smaller subintervals. For example, let $c = \frac{a+b}{2}$ then,

$$\begin{aligned}
\int_a^b f(x)\ dx &= \int_a^c f(x)\ dx + \int_c^b f(x)\ dx \\
&\approx \frac{c-a}{2}\Big[f(a) + f(c)\Big] + \frac{b-c}{2}\Big[f(c) + f(b)\Big] \\
&= \frac{h}{2}\Big[f(a) + 2f(c) + f(b)\Big] \equiv T_2(f),
\end{aligned}$$

where $h = \frac{b-a}{2}$.

# Testing on the previous example



## Example

Evaluate

$$I = \int_0^{\pi/2} \sin x \; dx$$

using the three point trapezoidal rule.
Compute the approximation error $I - T_2(f)$.
Please use the Fundamental theorem of calculus to directly calculate $I$.

$$\int_0^{\pi/2} \sin x \; dx \approx \frac{\pi}{8}\Big[\sin 0 + 2\sin(\pi/4) + \sin(\pi/2)\Big] \approx 0.9481.$$

$$\int_0^{\pi/2} \sin x \; dx - T_2(f) \approx 1 - 0.9481 = 0.0519.$$

# General Trapezoidal Rule $T_n(f)$

1. We saw the trapezoidal rule $T_1(f)$ for 2 points $a$ and $b$.
2. The rule $T_2(f)$ for 3 points involves three equidistant points: $a$, $\frac{a+b}{2}$ and $b$.
3. We observed the improvement in the accuracy of $T_2(f)$ over $T_1(f)$ so inspired by this,
   we would like to apply this rule to $n+1$ equally spaced points
   $$a = x_0 < x_1 < x_2 \cdots x_n = b$$
   with the space between any two points being denoted by $h$ that is
   $$h = x_{i+1} - x_i, \ \ i = 0, \cdots, n.$$

# General Trapezoidal Rule $T_n(f)$

## Definition

$$I \approx h\left[\frac{1}{2}f(a) + f(x_1) + \cdots f(x_{n-1}) + \frac{f(b)}{2}\right] \equiv T_n(f)$$

1. The subscript "n" refers to the number of subintervals being used;
2. the points $x_0,\ x_1, \cdots x_n$ are called the numerical integration node points.

# Performance of $T_n(f)$



$f(x) = \sin x$ we want to approximate $I = \int_0^{\pi/2} f(x)\, dx$ using the trapezoidal rule $T_n(f)$

| n | $T_n(f)$ | $I - T_n(f)$ | Ratio |
|---|---|---|---|
| 1 | 0.785 | 2.15e-1 | – |
| 2 | 0.948 | 5.18e-2 | 4.13 |
| 4 | 0.987 | 1.29e-2 | 4.03 |
| 8 | 0.997 | 3.21e-3 | 4.01 |
| 16 | .999 | 8.03e-04 | 4.00 |

Note that the errors are decreasing by a constant factor of 4. Why do we always double $n$?

# How to improve the accuracy of the integration rule?



An intuitive solution is to improve the accuracy of $f(x)$ by using a better interpolating polynomial say a quadratic polynomial $p_2(x)$ instead. Let $c = \frac{a+b}{2}$ and $h = \frac{b-a}{2}$ then, the quadratic polynomial is

$$p_2(x) = \frac{(x-c)(x-b)}{(a-c)(a-b)} f(a) + \frac{(x-a)(x-b)}{(c-a)(c-b)} f(c) + \frac{(x-a)(x-c)}{(b-a)(b-c)} f(b).$$

$$\int_a^b f(x)\, dx \approx \int_a^b p_2(x)\, dx = \frac{h}{3}\Big[f(a) + 4f(c) + f(b)\Big] \equiv S_2(f).$$

This is called Simpson's rule.

# Simpson's rule applied to the previous example

## Example

Evaluate

$$\int_0^{\pi/2} \sin x \; dx$$

using the Simpson's rule.

$$\int_0^{\pi/2} \sin x \; dx \approx \frac{\pi/2}{3}\Big[\sin 0 + 4\sin(\pi/4) + \sin(\pi/2)\Big]$$
$$\approx 1.00227$$

$$\int_0^{\pi/2} \sin x \; dx - S_2(f) = -0.00228.$$

# General Simpson's Rule $S_n(f)$



## Definition

$$I \approx \frac{h}{3}\Big[\big(f(a) + 4f(x_1) + 2f(x_2)\big) + 4f(x_3) + 2f(x_4) + 4f(x_5)\cdots$$

$$\cdots 4f(x_{n-1}) + f(b)\Big] \equiv S_n(f)$$

# Performance of $S_n(f)$



For $f(x) = \sin x$ we want to approximate $I = \int_0^{\pi/2} f(x)\, dx$ using the Simpson's rule $S_n(f)$

| n | $S_n(f)$ | $I - S_n(f)$ | Ratio |
|---|---|---|---|
| 2 | 1.00227 | -2.28e-3 | – |
| 4 | 1.00013 | -1.35e-4 | 16.94 |
| 8 | 1.00000 | -8.3e-6 | 16.22 |
| 16 | 1.00000 | -5.17e-7 | 16.06 |

# Error Formulas: Trapezoidal Rule



**Theorem**

*Let $f(x)$ have two continuous derivatives on $[a, b]$. Then,*

$$E_n^T(f) = \int_a^b f(x)\, dx - T_n(f) = -\frac{h^2(b-a)}{12} f''(c_n),$$

*where $c_n$ lies in $[a, b]$.*
*The error decays in a manner proportional to $h^2$.*

Thus doubling n (and halving $h$) should cause the error to decrease by a factor of approximately 4.
This is what we observed with a past example.

# Example

## Example

Consider the task of evaluating

$$I = \int_0^2 \frac{dx}{1+x^2}$$

using the trapezoidal rule $T_n(f)$.
How large should $n$ be chosen in order to ensure that

$$|E_n^T(f)| \leq 5 \times 10^{-6}?$$

# Example



## Proof.

We begin by calculating the derivatives involved:

$$f'(x) = \frac{-2x}{(1+x^2)^2}, \quad f''(x) = \frac{-2+6x^2}{(1+x^2)^3},$$

it is easy to check that

$$\max_{0\leq x\leq 2} |f''(x)| = 2$$

$$\text{thus,} \quad |E_n^T(f)| = |\frac{-h^2(b-a)}{12} f''(c_n)|$$
$$\leq \frac{2h^2}{12}\cdot 2 = \frac{h^2}{3}.$$

We bound $|f''(c_n)|$ since we do not know the exact value of $c_n$ and hence, we must assume the worst possible value of $c_n$ that makes the error formula the largest. □

# Proof



## Proof.

When do we have

$$|E_n^T(f)| \leq 5 \times 10^{-6}?$$

We need to choose $h$ so small that

$$\frac{h^2}{3} \leq 5 \times 10^{-6}$$

which is possible if $h \leq 0.003873$ (verify!). This is equivalent to choosing

$$n = \frac{b-a}{h} = \frac{2-0}{h} \geq 516.4.$$

Thus, $n \geq 517$ will make the error smaller than $5 \times 10^{-6}$. $\square$

# Error Formulas: Simpson's Rule

**Theorem**

*Let $f(x)$ have four continuous derivatives on $[a, b]$. Then,*

$$E_n^S(f) = \int_a^b f(x)\, dx - S_n(f) = -\frac{h^4(b-a)}{180} f^{(4)}(c_n),$$

*where $c_n$ lies in $[a, b]$.*
*The error decays in a manner proportional to $h^4$.*

Thus doubling n should cause the error to decrease by a factor of approximately 16.
This is what we observed with a past example.

# Example



## Example

Consider the task of evaluating

$$I = \int_0^2 \frac{dx}{1+x^2}$$

using the Simpson's rule $T_n(f)$.
How large should $n$ be chosen in order to ensure that

$$|E_n^S(f)| \le 5 \times 10^{-6}?$$

# Proof

## Proof.

We compute the fourth derivative

$$f^{(4)}(x) = 24\frac{5x^4 - 10x^2 + 1}{(1+x^2)^5}$$

$$\max_{0\le x\le 1} |f^{(4)}(x)| = f^{(4)}(0) = 24.$$

$$\text{Thus,} \quad E_n^S(f) = -\frac{h^4(b-a)}{180} f^{(4)}(c_n)$$

$$\le \frac{h^4 \cdot 2}{180} \cdot 24 = \frac{4h^4}{15} \le 5 \times 10^{-6}$$

$$\text{provided} \quad h \le 0.0658 \quad \text{or} \quad n \ge 30.39,$$

thus choosing $n \ge 32$ will give the desired error bound. $\square$

Compare with the trapezoidal rule: $n \ge 517$!

# One more example

Consider the application of trapezoidal and Simpson's rule to approximate

$$\int_0^1 \sqrt{x}\ dx$$

| n | $E_n^T(f)$ | *Ratio* | $E_n^S(f)$ | Ratio |
|---|---|---|---|---|
| 2 | 6.311e-2 | – | 2.86e-2 | – |
| 4 | 2.338e-2 | 2.7 | 1.012e-2 | 2.82 |
| 8 | 8.536e-3 | 2.77 | 3.587e-3 | 2.83 |
| 16 | 3.085e-3 | 2.78 | 1.268e-4 | 2.83 |
| 32 | 1.108e-3 | 2.8 | 4.485e-4 | 2.83 |

Observe that the rate of convergence is slower since $f(x) = \sqrt{x}$ is not sufficiently differentiable on $[0, 1]$. Both converge at a rate proportional to $h^{1.5}$.