Numerical Integration

Basics and Newton-Cotes

Sachin Shanbhag

Department of Scientific Computing
Florida State University,
Tallahassee, FL 32306.

References

- S. Chapra and R. Canale, Numerical Methods for Engineers
- Carnahan and Wilkes, Applied Numerical Methods, University of Michigan, Class Notes, 1996.
- Pal, Numerical Analysis for Scientists and Engineers, 2007.
- Holistic Numerical Methods Institute webpage
- Samir Al-Amer, Class Notes
- wikipedia.org

Motivation: Why numerical integration?

- Numerical evaluation of an analytical solution may be problematic

$$\int_0^{\infty} \frac{dx}{\sqrt{x}\cosh(x)} = 2\sqrt{\pi}\sum_{k=0}^{\infty} \frac{(-1)^k}{\sqrt{2k+1}}$$

- Many integrals cannot be evaluated analytically

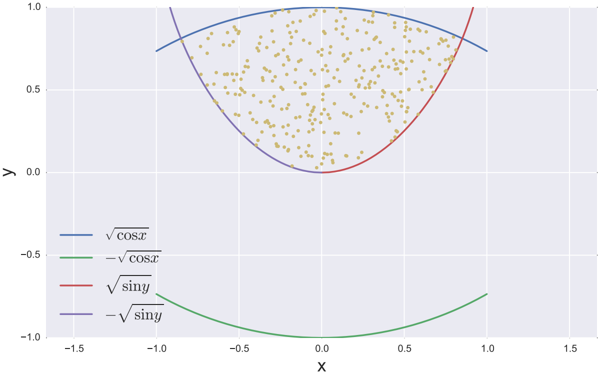
$$I = \int_a^b e^{-x^2} dx$$

This is related to the so-called “error function”

Motivation: Why numerical integration?

- Integral may be multidimensional/have complicated boundaries

$$I = \iint_{x^2<\sin(y),y^2<\cos(x)} e^{\sqrt{x^2+y^2}} dxdy$$

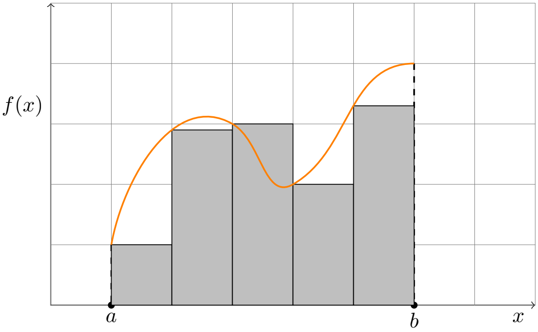


Riemann Sum

In the differential limit, an integral is equivalent to a summation

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=0}^{n} f(x_i)\Delta x \approx \sum_{i=0}^{N-1} f(x_i)\Delta x$$

where N is a large, but finite, number



Orientation

- Newton-Cotes
  - interpolating polynomials
  - trapezoidal, Simpson's rules, Bode's rule
  - Multistep rules
- Romberg Integration
  - extrapolation technique
  - successive refinement
- Gaussian Quadrature
  - optimize grid spacing to obtain higher order accuracy
  - unevenly spaced points
- Singularities
- Monte Carlo Integration

Newton-Cotes

- The function to be integrated is approximated by a polynomial of order n, in the interval $[a, b]$

$$f(x) \approx p_n(x) = \sum_{i=0}^{n} a_i x^i$$

  This can be done by using our knowledge of interpolating polynomials
- Getting the integral of the polynomial is easy

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} a_i \int_a^b x^i dx = \sum_{i=0}^{n} a_i \frac{b^{i+1} - a^{i+1}}{i+1}$$

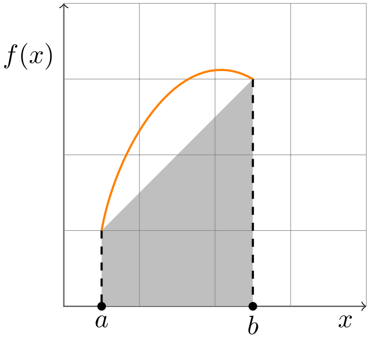
Newton-Cotes: Trapezoid Rule

First ordered polynomial

$$f(x) \approx p_1(x) = a_0 + a_1 x = f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

This implies that the area of the gray shaded region:

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



Trapezoid Rule: Error

From Taylor series expansion:

$$\begin{aligned}\epsilon &= \int_a^b f(x)dx - \int_a^b p_n(x)dx \\ &= \frac{1}{n+1!}\int_a^b f^{(n+1)}(\xi)\left(\prod_{l=0}^{n}(x-x_l)\right)dx.\end{aligned}$$

For trapezoidal rule $n = 1$,

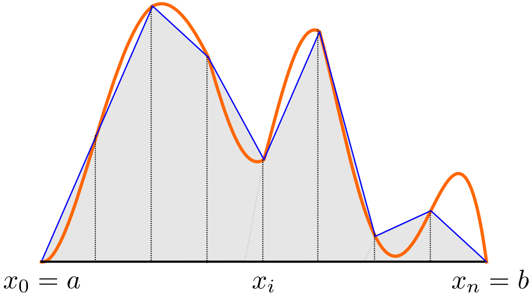
$$\begin{aligned}\epsilon &= \frac{1}{2}\int_a^b f''(\xi)(x-a)(x-b)dx \\ &= \frac{1}{2}f''(\eta)\int_a^b (x-a)(x-b)dx \\ &= -\frac{(b-a)^3}{12}f''(\eta)\end{aligned}$$

where $\eta \in [a, b]$, and the mean-value theorem is applied.

Multistep Trapezoidal

Divide interval $[a, b]$ into n segments, which may not necessarily be equal:

$$a = x_0 < x_1 < x_2 < ... < x_n = b$$



Composite formula

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

Multistep Trapezoidal: Error

Equally spaced points: $h = x_{i+1} - x_i$

$$\int_a^b f(x)dx \approx h\left[\frac{f(x_0)+f(x_n)}{2} + \sum_{i=1}^{n-1} f(x_i)\right]$$

From Taylor series expansion, for equally spaced case, the error can be shown to be:

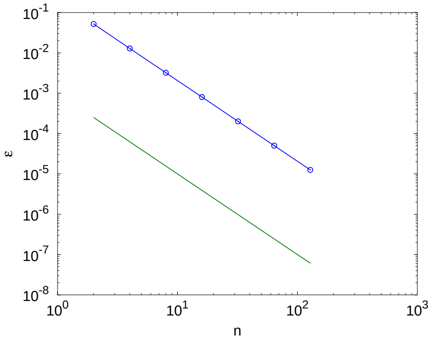
$$\epsilon = \frac{(b-a)^3}{12n^2} \max_{\xi\in[a,b]} f''(\xi)$$

Trapezoidal Rule: Error

Consider the example:

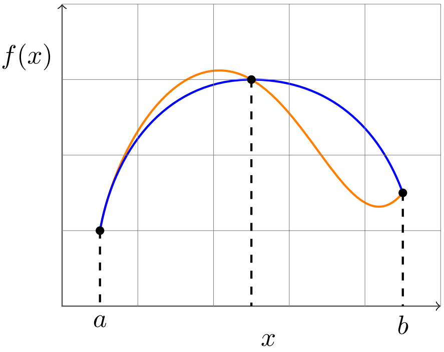
$$I = \int_{0}^{\pi/2} \cos(x)dx = 1$$

with $n = 2,\ 4,\ 8,\ 16,\ 32,\ 64,\ 128.$



Simpson's 1/3 Rule

In the single step version, we use a second order polynomial (quadratic) to pass through three points in the domain.

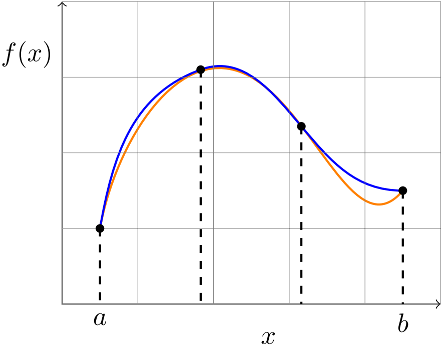


If points are equispaced, then

$$h = \frac{(b-a)}{2}$$

Simpson's 3/8 Rule

In the single step version, we use a third order polynomial (quadratic) to pass through 4 points in the domain.



If points are equispaced, then

$$h = \frac{(b-a)}{3}$$

Error: Single Step

Assuming equispaced points $h = x_{i+1} - x_i$ with $a = x_0$ and $b = x_n$ the endpoints, and $f_i = f(x_i)$, we can summarize the first few Newton-Cotes formulas:

Degree	Common name	Formula	Error
1	Trapezoid rule	$\frac{b-a}{2}(f_0 + f_1)$	$-\frac{(b-a)^3}{12} f^{(2)}(\xi)$
2	Simpson's 1/3 rule	$\frac{b-a}{6}(f_0 + 4f_1 + f_2)$	$-\frac{(b-a)^5}{2880} f^{(4)}(\xi)$
3	Simpson's 3/8 rule	$\frac{b-a}{8}(f_0 + 3f_1 + 3f_2 + f_3)$	$-\frac{(b-a)^5}{6480} f^{(4)}(\xi)$
4	Boole's rule	$\frac{b-a}{90}(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4)$	$-\frac{(b-a)^7}{1935360} f^{(6)}(\xi)$

Note that there is something special about Simpson's $1/3$ rule.

Simpson's 1/3 Multistep

Assuming equally spaced points divide interval $[a, b]$ into even number of strips ($n = 2, 4, 6$, etc.)

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6}\left[f(x_0)+2\sum_{j=1}^{n/2-1} f(x_{2j})+4\sum_{j=1}^{n/2} f(x_{2j-1})+f(x_n)\right]$$

With $n = 6$ for example

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6}\left[f_0 + 2f_1 + 4f_2 + 2f_3 + 4f_4 + 2f_5 + f_6\right]$$

Multistep Error Bounds

- Trapezoidal

$$e = \frac{(b-a)^3}{12n^2} \max_{x \in [a,b]} f''(x)$$

- Simpson's 1/3

$$e = \frac{(b-a)^5}{2880n^4} \max_{x \in [a,b]} f^{(4)}(x)$$

- Simpson's 3/8

$$e = \frac{(b-a)^5}{6480n^4} \max_{x \in [a,b]} f^{(4)}(x)$$

Newton Cotes

- Prefer odd-point (= even degree polynomial) formulas over even point
  Note: Simpson's $1/3$ rule popularity
- Very high order schemes are undesirable large weights with alternating signs
- Essentially the same reason that interpolation fails with high order schemes
- In practice n is almost never greater than 8.
- Summary: In Newton-Cotes you have two choices: the order of the polynomial (first = trapezoidal, second = Simpson's $1/3$, etc.), and the number of strips you want to discretize the domain into (n).

Romberg Integration

- Trapezoid formula with n subintervals gives error $\epsilon \sim 1/n^2$
- Say, I repeat the calculation with $2n$ subintervals
- Can I combine to get a better estimate?

Error in trapezoidal:

$$E = \frac{(b-a)^3}{12n^2} f''(\xi), \text{ for some } \xi \in [a, b]$$

If we assume $f''(\xi)$ is independent of n

$$E = \frac{C}{n^2}$$

Romberg Integration

- If we assume that C is roughly constant I_{true} = true value of the integral I_n = numerical approximation using n steps,
- We have

$$I_{true} = I_n + E(n)$$

- Therefore,

$$\begin{aligned} E(n) &\approx \tfrac{C}{n^2} &= I_{true} - I_n \\ E(2n) &\approx \tfrac{C}{4n^2} &= I_{true} - I_{2n} \end{aligned}$$

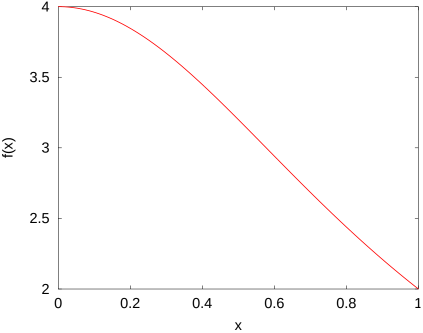
Eliminate C/n^2,

$$\boxed{I_{true,est} = I_{2n} + \frac{I_{2n} - I_n}{3}}$$

Romberg Integration: Example

- Consider the integral

$$I = \int_0^1 \frac{4}{1+x^2}\, dx = \pi.$$



- Let us numerically compute the solution using trapezoidal rule with different number of steps, n

Example

Using multistep trapezoidal rule:

n	I_n	ϵ
1	3.0000	1.4159e-01
2	3.1000	4.1593e-02
4	3.1312	1.0416e-02
8	3.1390	2.6042e-03
16	3.1409	6.5104e-04
32	3.1414	1.6276e-04
64	3.1416	4.0690e-05
128	3.1416	1.0173e-05
256	3.1416	2.5431e-06
512	3.1416	6.3578e-07

You can see that the error goes down roughly by a factor of 4 with every doubling of n

Example

- We can refine an estimate by using say, $I_2 = 3.1000$ and $I_4 = 3.1312$ as

$$I_{true,est} = I_4 + \frac{I_4 - I_2}{3} = 3.1313$$

  which pulls the estimate in the direction of the true solution for essentially no extra work.
- Note: I_2 uses a subset of the same function evaluations as I_4. That is, the grid points x_i that are used in the calculation of I_4 include all of the grid points that are used in the evaluation of I_2.
- This is where a majority of the cost-savings comes from.

Romberg

Trapezoidal			Romberg	
n	$I_n^{(1)}$	ϵ	$I_n^{(2)}$	ϵ
1	3.0000	1.4159e-01		
2	3.1000	4.1593e-02	3.1004	4.1200e-02
4	3.1312	1.0416e-02	3.1313	1.0294e-02
8	3.1390	2.6042e-03	3.1390	2.5735e-03
16	3.1409	6.5104e-04	3.1409	6.4338e-04

Typically, this itself leads to better estimates. But that is not all. We can squeeze out some more performance relatively cheaply.

Error Estimate

- For trapezoidal rule, it can be shown that

$$E = \frac{C_1}{n^2} + \frac{C_2}{n^4} + \frac{C_3}{n^6} + \cdots$$

- For the scheme

$$I_{true,est} = I_{2n} + \frac{I_{2n} - I_n}{3}$$

  the leading error is $\mathcal{O}(1/n^4)$ (like Simpson's rule).
- In fact, in some places, this is how Simpson's is implemented (Numerical Recipes, Press et al.)

Successive Refinement

- General Expression for Romberg Integration

$$I_{2n}^{(k)} = \frac{4^{k-1} I_{2n}^{(k-1)} - I_n^{(k-1)}}{4^{k-1} - 1}$$

- index k represents order of extrapolation
- $I_n^{(1)}$ represents regular trapezoidal with n segments
- $k = 2$ represents values reported above using one application of Romberg integration
- $I_n^{(k)}$ has an error of order $1/n^{2k}$

Example: Successive Refinement

n	$I_n^{(1)}$	$I_n^{(2)}$	$I_n^{(3)}$	$I_n^{(4)}$	$I_n^{(5)}$
1	3.0000				
2	3.1000	3.1333			
4	3.1312	3.1416	3.1421		
8	3.1390	3.1416	3.1416	3.1416	
16	3.1409	3.1416	3.1416	3.1416	3.1416