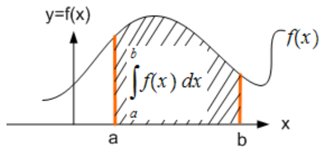
Numerical Integration

In physics and engineering of any fields, definite integral overwhelms the use of indefinite integral and, more often than not, such definite integrals can rarely be carried out by hand or analytically –

fields due to electrical charge configuration or mass distribution, center of gravity of space shuttle, lift and drag experienced by an aircraft and the list continues without end.

Numerically, try an approximate solution to a definite integral up to desired precision



$$\mathcal{I} = \int_a^b f(x)\,dx$$

where $f(x)$ is a (piecewise) continuous, well-behaved (smooth) function over interval $[a, b]$ of our interest. Definite integration is obviously the area under the curve as shown.

A few reasons for doing integration numerically

1. $f(x)$ may be known only at certain points
2. $f(x)$ is known but analytical integration may be too difficult or at times impossible to carry out
3. integration can be carried out analytically but numerically it may be far more easier to a given accuracy.

Numerical integration is often called *quadrature* especially when it is applied to one dimensional integration.

Numerical integration amounts to evaluating the integrand at finite set of points (may or may not be equally spaced) bounded by interval $[a, b]$ and doing a weighted sum of these values,

$$\mathcal{I} = \int_a^b f(x)\,dx \approx \sum_{n=1}^{N} w(x_n)\,f(x_n) \equiv \mathcal{I}_N$$

where $w(x_n)$ is weight function and N is the number of integration points. Choice of N depends on maximum error $|\mathcal{I} - \mathcal{I}_N| < \epsilon$ we desire.

The origin of weight function is Lagrange's interpolating polynomial,

$$p(x) = f(x_0) \cdot \ell_0(x) + f(x_1) \cdot \ell_1(x) + \cdots + f(x_N) \cdot \ell_N(x)$$

where $\ell_n(x)$ are n-degree polynomials. Integrating the polynomial,

$$\begin{aligned}
\mathcal{I}_N &= \int_a^b p(x)\, dx = \int_a^b \sum_{n=0}^{N} f(x_n) \cdot \ell_n(x)\, dx \\
&= \sum_{n=0}^{N} \left(f(x_n) \cdot \int_a^b \ell_n(x)\, dx \right) \equiv \sum_{n=0}^{N} f(x_n) \cdot w(x_n)
\end{aligned}$$

Techniques to explore –

- Midpoint
- Trapezoidal
- Simpson
- Monte Carlo
- Gaussian Quadrature

Strategy : approximate $f(x)$ with polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots$.

Midpoint : only constant a_0
Trapezoidal : up to linear $a_0 + a_1 x$
Simpson : up to quadratic $a_0 + a_1 x + a_2 x^2$

Midpoint method

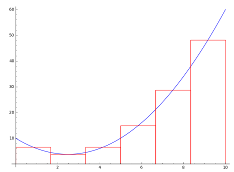
1. Divide the integration range $[a, b]$ in N equal parts of width h

$$h = (b - a)/N$$

2. Determine midpoints of each intervals *i.e.* the integration points

$$x_1 = \frac{(a) + (a + h)}{2}, \; x_2 = \frac{(a + h) + (a + 2h)}{2}, \; x_3 = \frac{(a + 2h) + (a + 3h)}{2}, \ldots$$

   $f(x)$ is evaluated only at x_n, assumption being the area of each rectangle evaluated at each integration point $h \times f(x_n)$ approximates the area under the curve over that sub-interval.



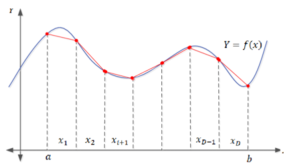
3. The sum of all such rectangular area is

$$\mathcal{M}_N = \sum_{n=1}^{N} h \cdot f(x_n) \;\Rightarrow\; \lim_{N\to\infty} \mathcal{M}_N = \int_a^b f(x)\, dx$$

   where $w(x_n) = w = 1$ or h, constant for all x_n

Trapezoidal method

Estimate definite integrals using straight lines connecting consecutive $f(x_n)$ and generating trapezoids whose areas are expected to better approximate area under the curve over each interval.



1. We have N intervals each of width h which forms the width of each trapezoid. Endpoints of each interval be at $x_0, x_1, x_2, \dots x_N$ where

$$x_0 = a,\ x_1 = x_0 + h,\ x_2 = x_0 + 2h,\ \dots\ x_n = x_0 + Nh = b$$

2. Evaluate $f(x_n)$ and calculate area of each trapezoid,

$$\mathcal{T}_n = \frac{h}{2}\Big(f(x_{n-1}) + f(x_n)\Big)$$

3. Sum over all the $\mathcal{T}_n$'s to approximate the integral

$$\mathcal{T}_N = \sum_{n=1}^{N} \mathcal{T}_n = \frac{h}{2}\Big([f(x_0) + f(x_1)] + [f(x_1) + f(x_2)] + \cdots + [f(x_{N-1}) + f(x_N)]\Big)$$

Hence the approximation to the integral is,

$$\begin{aligned} \mathcal{T}_N &= \frac{h}{2}\Big(f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{N-1}) + f(x_N)\Big) \\ &= \sum_{n=1}^{N} w(x_n)\cdot f(x_n) \quad \Rightarrow \quad \lim_{N\to\infty} \mathcal{T}_N = \int_a^b f(x)\,dx \end{aligned}$$

where the weight function $w(x_0) = w(x_N) = 1$ or $h/2$ and $w(x_1) = w(x_2) = \cdots = w(x_{N-1}) = 2$ or h.

Contrary to expectation, trapezoidal rule tends to be less accurate than midpoint rule, particularly when the curve is strictly concave or convex over the integration limits.

Exercise : For $N = 4, 8, 12, 20$ using both Midpoint and Trapezoidal to evaluate and compare with the analytical result

$$\int_1^2 \frac{1}{x}\,dx = 0.69314718 \qquad \int_0^{\pi/2} x\,\cos x\,dx = \frac{\pi}{2} - 1$$

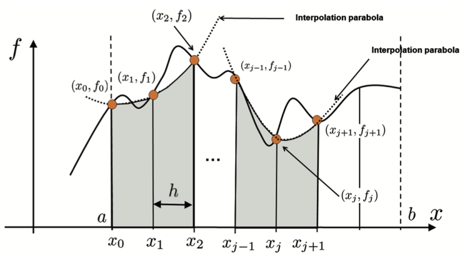
Simpson's method

Midpoint method estimates area under the curve by rectangle *i.e.* piecewise constant function.

Trapezoidal method estimates area under the curve by trapezium *i.e.* piecewise linear function.

Simpson method estimates area under the curve by piecewise quadratic function $\Rightarrow$ need 3-points – two are the (sub) interval boundaries and third is the midpoint of these twos,

$$\int_{x_0}^{x_2} f(x)\,dx \rightarrow (x_0, f(x_0)), (x_2, f(x_2)), (x_1, f(x_1)) \quad \text{where} \quad x_1 = \frac{x_0 + x_2}{2}$$



Upon integrating and using $h = (x_2 - x_0)/2$, we obtain

$$\begin{aligned}
\int_{x_0}^{x_2} f(x)\,dx &\approx \int_{x_0}^{x_2} \Big(a_2x^2 + a_1x + a_0\Big)\,dx = \left(\frac{a_2}{3}x^3 + \frac{a_1}{2}x^2 + a_0x\right)\Big|_{x_0}^{x_2} \\
&= \frac{a_2}{3}\left(x_2^3 - x_0^3\right) + \frac{a_1}{2}\left(x_2^2 - x_0^2\right) + a_0\left(x_2 - x_0\right) \\
&= \frac{x_2 - x_0}{6}\Big[2a_2\left(x_2^2 + x_2x_0 + x_0^2\right) + 3a_1\left(x_2 + x_0\right) + 6a_0\Big] \\
&= \frac{h}{3}\Big[\left(a_2x_2^2 + a_1x_2 + a_0\right) + \left(a_2x_0^2 + a_1x_0 + a_0\right) \\
&\qquad + a_2\left(x_2^2 + 2x_2x_0 + x_0^2\right) + 2a_1\left(x_2 + x_0\right) + 4a_0\Big] \\
\int_{x_0}^{x_2} f(x)\,dx &\approx \frac{h}{3}\Big[f(x_2) + f(x_0) + a_2\left(2x_1\right)^2 + 2a_1\left(2x_1\right) + 4a_0\Big] \\
&= \frac{h}{3}\Big[f(x_2) + f(x_0) + 4(a_2x_1^2 + a_1x_1 + a_0)\Big] \\
&= \frac{h}{3}\Big[f(x_0) + 4f(x_1) + f(x_2)\Big]
\end{aligned}$$

Similar expressions are obtained for other intervals. For example,

$$\int_{x_2}^{x_4} f(x)dx = \frac{h}{3}\Big[f(x_2) + 4f(x_3) + f(x_4)\Big]$$

Therefore, summing of all such terms in the intervals $[x_0, x_2]$, $[x_2, x_4]$, $\ldots$, $[x_{N-2}, x_N]$, the Simpson's $1/3$ method gives

$$\begin{aligned}
\mathcal{S}_N &= \frac{h}{3}\Big[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 2f(x_{N-2}) + 4f(x_{N-1}) + f(x_N)\Big] \\
&= \frac{h}{3}\left[f(x_0) + 4\sum_{n \text{ odd}}^{N-1} f(x_n) + 2\sum_{n \text{ even}}^{N-2} f(x_n) + f(x_N)\right] \\
&= \sum_{n=1}^{N} w(x_n)\cdot f(x_n) \quad \Rightarrow \quad \lim_{N\to\infty} \mathcal{S}_N = \int_a^b f(x)\, dx
\end{aligned}$$

where, the weight functions are

$$w(x_0) = w(x_N) = 1, \quad w(x_n) = 4 \ \text{ for odd } n \quad \text{and} \quad w(x_n) = 2 \ \text{ for even } n$$

N.B. To use Simpson's 1/3-method there must be

even numbers of intervals and odd numbers of grid points.

There exists Simpson's 3/8-method that makes use of cubic polynomial i.e. retaining terms up to $a_3 x^3$.
An interesting relation

$$\mathcal{S}_{2N} = \frac{2}{3}\mathcal{M}_N + \frac{1}{3}\mathcal{T}_N$$

Error in numerical integration

For numerical integration, we cannot define absolute or relative errors

$$\text{Absolute error: } |\mathcal{I}_N - \mathcal{I}| \quad \text{and Relative error: } \left|\frac{\mathcal{I}_N - \mathcal{I}}{\mathcal{I}}\right| \times 100\%$$

since we do not apriori know $\mathcal{I}$, the reason for doing it numerically.

Instead, we define **upper bound** of the error that each method will yield.
For that, $f(x)$ is continuous over $[a, b]$

- having finite $f''(x)$ for Midpoint and Trapezoidal
- having finite $f''''(x)$ for Simpson

Estimated upper bounds for error in numerical integration schemes are

$$\begin{aligned}
\text{Midpoint : } \text{Error in } \mathcal{M}_N &\leq \frac{(b-a)^3}{24N^2} \left|f''(x)\right|_{\max} \\
\text{Trapezoidal : } \text{Error in } \mathcal{T}_N &\leq \frac{(b-a)^3}{12N^2} \left|f''(x)\right|_{\max} \\
\text{Simpson : } \text{Error in } \mathcal{S}_N &\leq \frac{(b-a)^5}{180N^4} \left|f''''(x)\right|_{\max}
\end{aligned}$$

What if $f'^{\cdots'}(x) = 0$? It does not mean the error is zero but simply indicates that estimate of error upper bound can not be made.

Consider an example – estimate numerically the following integral whose exact analytical answer is available,

$$\int_0^1 x^2\, dx = 0.333$$

Don't worry about error upper bound! Divide integration limit $[0, 1]$ in 4-intervals : $[0, 1/4]$, $[1/4, 2/4]$, $[2/4, 3/4]$, $[3/4, 1]$i, interval length $h = 4$.

Midpoint needs midpoints of these subintervals – $1/8$, $3/8$, $5/8$, $7/8$

$$\begin{aligned} M_4 &= \frac{1}{4}\left[f\left(\frac{1}{8}\right) + f\left(\frac{3}{8}\right) + f\left(\frac{5}{8}\right) + f\left(\frac{1}{8}\right)\right] \\ &= \frac{1}{4}\left[\left(\frac{1}{8}\right)^2 + \left(\frac{3}{8}\right)^2 + \left(\frac{5}{8}\right)^2 + \left(\frac{7}{8}\right)^2\right] = \frac{21}{64} = 0.328 \end{aligned}$$

Absolute error is $|0.333 - 0.328| \approx 0.0052$ and relative error is 1.56%.

For upper error bound of 0.001

$$f(x) = x^2 \;\Rightarrow\; f^{'}(x) = 2x \;\Rightarrow\; f^{''}(x) = 2 \;\Rightarrow\; |f^{''}(x)|_{\max} = 2$$

$$0.001 = \frac{(1-0)^3}{24N^2}\, 2 \;\Rightarrow\; N = 9$$

The requirement of **upper bound** for error determines N.

Two things to keep in mind –

1. N chosen should be the smallest integer value greater than or equal to N (basically `ceil(N)`)
2. the actual estimate may, in fact, be much better approximation than is indicated by the error upper bound.

It may sound strange, but it is often true, that taking N larger than what obtained from the error bounds can actually deteriorate the accuracy, contrary to what the limit $N \to \infty$ suggests.

Consider the function $f(x) = 1/x$ in the range $[1, 2]$

$$f''(x) = \frac{2}{x^3}, \; f''''(x) = \frac{24}{x^5} \quad \Rightarrow \quad \left|f''(x)\right|_{\text{max}} = 2 \text{ and } \left|f''''(x)\right|_{\text{max}} = 24$$

Hence, for maximum error bound of 0.0001, N is equal to

$$\begin{aligned}
\text{Midpoint :} \qquad & 0.0001 = \frac{(2-1)^3}{24N^2} \times 2 \quad \to \quad N = 29 \\
\text{Trapezoidal :} \qquad & 0.0001 = \frac{(2-1)^3}{12N^2} \times 2 \quad \to \quad N = 41 \\
\text{Simpson :} \qquad & 0.0001 = \frac{(2-1)^5}{180N^4} \times 24 \quad \to \quad N = 7
\end{aligned}$$

Monte Carlo integration

Monte Carlo method is a class of algorithm that involves repeated random sampling to estimate an integration numerically. It is a widely used (perhaps the only) method to estimate higher dimensional *i.e.* multi-variable integrals.

To estimate $\int_a^b f(x)dx$, we define an *estimator*. Given a random variable X drawn from a PDF $p(X)$, the estimator $\mathcal{F}$ is defined as

$$\mathcal{F}_N \equiv \frac{1}{N}\sum_{i=1}^{N}\frac{f(X_i)}{p(X_i)} \quad \Rightarrow \quad \langle \mathcal{F}_N \rangle = \int_a^b \mathcal{F}_N \cdot p(X)\, dX \equiv \int_a^b f(x)\, dx$$

For the particular case of uniform PDF in $[a, b]$

$$\int_a^b p(X)\, dX = \int_a^b \mathcal{C}\, dX = 1 \quad \Rightarrow \quad p(X) = \mathcal{C} = \frac{1}{b-a}$$

$$\mathcal{F}_N = \frac{b-a}{N}\sum_{i=1}^{N} f(X_i)$$

The above expression looks very similar to Midpoint expression.

The steps involve in Monte Carlo integration are

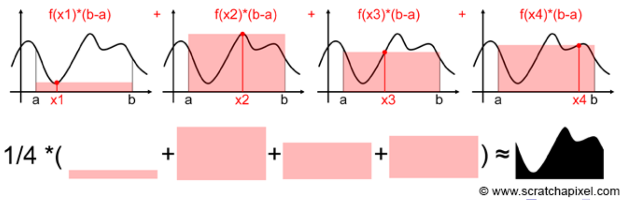
1. Choose a N, say 10 or 20 or 50 or whatever.
2. Draw N number of random variables X_i from its domain $[a, b]$. If your pRNG returns uniform random numbers in the range $[0, 1]$, convert it to $[a, b]$ using

$$X = a + (b - a)\xi \quad \text{where,} \quad \xi \in [0, 1]$$

3. For each X_i calculate $f(X_i)$ and determine $\mathcal{F}_N$ and σ_f using

$$\sigma_f^2 = \frac{1}{N}\sum_{i=1}^{N} f(X_i)^2 - \left(\frac{1}{N}\sum_{i=1}^{N} f(X_i)\right)^2$$

4. Tabulate or plot $\mathcal{F}_N$ versus N. Also keep track of σ_f for each N. As N increases, we will find $\mathcal{F}_N$ converging to a value but decrease in σ_f will be rather slow $\sim 1/\sqrt{N}$.
5. Go to step (1), increase N by 10 or whatever times and repeat.



Gaussian Quadrature

Midpoint, Trapezoidal, Simpson, collectively called closed Newton-Cotes method, whose general form is integrating Lagrange polynomial that interpolates $f(x)$ at N equally spaced points,

$$\mathcal{I} \;=\; \int_a^b f(x)\,dx \approx \sum_{n=1}^{N} w(x_n)\cdot f(x_n) \quad \text{where} \quad w(x_n) = \int_a^b L_n(x)\,dx$$

$$\text{with } k\text{-point} \quad L_n(x) = \prod_{k=1,k\neq n}^{N} \frac{x - x_n}{x_k - x_n}$$

where $x_n =$ nodes and $w(x_n) =$ quadrature weights.

If integration points are not equally spaced, other than they must lie in the interval $[a, b]$, then we have $2n$ parameters : x_n and $w(x_n)$.

Gaussian quadrature rule : express integrand in terms of orthogonal basis polynomial over the given interval. Say, for instance

$$\int_{-1}^{1} f(x)\,dx \;\approx\; \sum_{n=1}^{N} w(x_n)\cdot f(x_n)$$

x_n, at which $f(x_n)$ are evaluated, are roots of an orthogonal polynomial for same interval. It is optimal since it fits all polynomials up to degree $2n-1$ exactly defined by the $2n$ parameters.

In interval $[-1, 1]$, the $f(x)$ is well-approximated by Legendre polynomials $P_n(x)$, where x_n is n-th root of $P_n(x)$ and the weights $w(x_n)$ are

n of $P_n(x)$	x_n roots of $P_n(x)$	weights $w(x_n)$
1	0	2
2	$\pm\sqrt{\frac{1}{3}}$	1
3	$0, \pm\sqrt{\frac{3}{5}}$	$\frac{8}{9}, \frac{5}{9}$
4	$\pm\sqrt{\frac{3}{7} - \frac{2}{7}\sqrt{\frac{6}{5}}}, \pm\sqrt{\frac{3}{7} + \frac{2}{7}\sqrt{\frac{6}{5}}}$	$\frac{18+\sqrt{30}}{36}, \frac{18-\sqrt{30}}{36}$
5	$\dots$	$\dots$

However, if the integral is of type $\int_a^b$, then

$$\int_a^b f(t)\,dt = \frac{b-a}{2}\int_{-1}^{1} f\left(\frac{b-a}{2}x + \frac{b+a}{2}\right) dx$$

$$\int_a^b f(t)\,dt \approx \frac{b-a}{2}\sum_{n=1}^{N} w(x_n)\cdot f\left(\frac{b-a}{2}x_n + \frac{b+a}{2}\right)$$

If, the integrals are of type say $\int_0^\infty$ or $\int_{-\infty}^{\infty}$ instead of $\int_{-1}^{1}$, then the orthogonal polynomials used are Laguerre $L_n(x)$ and Hermite $H_n(x)$.

Steps involved are straight forward : instead of mid points and / or sub-interval boundaries,

1. evaluate $f(x)$ at roots $x_0, x_1, x_2, \ldots, x_N$ of an orthogonal polynomial.
2. use the corresponding weights $w(x_n)$ and perform the sum

$$\sum_{n=1}^{N} w(x_n) \cdot f(x_n) \;\rightarrow\; \int_a^b f(x)\, dx$$

To demonstrate the efficiency of Gaussian quadrature, we evaluate an integral whose analytical result is available for comparison against Simpson to achieve the same level of accuracy

$$\int_{-1}^{1} x\, e^x\, dx = \frac{2}{e} = 0.735\ 758\ 88$$

Simpson		Gaussian Q	
N	$\mathcal{I}_S$	N	$\mathcal{I}_G$
20	0.735 764 50	4	0.735 756 50
40	0.735 759 23	5	0.735 758 87
60	0.735 758 95	6	0.735 758 88
80	0.735 758 90		
100	0.735 758 88		
120	0.735 758 88		