# Lecture 3
# Numerical Integration

There are many circumstances in mathematical modelling when we need to evaluate an integral. We may need to do so because this is the explicit answer to some problem we have directly, or because it is a step on the way to solving a differential equation. It is worth stating that often clever reformulation of differential equations leaves them in integral form, that is, the differential equation is easier to solve analytically, but the integral equation gives a more physical feel to the solution.

## Rectangular formula

The basic idea to evaluate a definite integral is to replace the integral with a finite sum

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i). \tag{1}$$

This approximate equality is called a quadrature formula. Points $x_i$ are called nodes and $A_i$ are coefficients of a quadrature formula.

We shall derive first the simplest quadrature formula which is called the rectangular formula. Let us introduce a uniform grid for the segment $[a, b]$ with step $h$

$$\Omega_h = \{x_i = a + ih, \quad i = 0, 1, ..., n, \quad h = (b - a)/n\}. \tag{2}$$

It is clear that

$$\int_a^b f(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx. \tag{3}$$

Let us replace the integral

$$\int_{x_{i-1}}^{x_i} f(x)dx$$

with $f(x_{i-1/2})h$, where $x_{i-1/2} = x_i - h/2$. Geometrically, this means that we replace the square of the curvilinear trapezium with a square of corresponding rectangular. We obtain

$$\int_{x_{i-1}}^{x_i} f(x)dx \approx f(x_{i-1/2})h,$$

which is known as a rectangular formula for a partial segment $[x_{i-1}, x_i]$. Using Taylor's series we can easily evaluate a corresponding error

$$R_i = \int_{x_{i-1}}^{x_i} (f(x) - f(x_{i-1/2}))dx = \int_{x_{i-1}}^{x_i} \frac{(x - x_{i-1/2})^2}{2} f''(\zeta_i)dx,$$

as

$$f(x) = f(x_{i-1/2}) + f'(x_{i-1/2})(x - x_{i-1/2}) + f''(\zeta_i)\frac{(x - x_{i-1/2})^2}{2},$$

where $\zeta_i \in [x_{i-1}, x_i]$. If $M_{2,i} = \max |f''(x)|$ at $x \in [x_{i-1}, x_i]$, we have estimate for the error

$$|R_i| \leq M_{2,i} \int_{x_{i-1}}^{x_i} \frac{(x - x_{i-1/2})^2}{2} dx = \frac{h^3}{24} M_{2,i}. \tag{4}$$

As a result of summation of contributions for all partial segments we obtain composite rectangular formula which looks like this

$$\int_a^b f(x)dx = h \sum_{i=0}^{n} f(x_{i-1/2}) + R, \tag{5}$$

where the estimate for the error $R$ is given by

$$|R| \leq \frac{h^2(b-a)}{24} M_2, \tag{6}$$

where $M_2 = \max |f''(x)|$ at $x \in [a, b]$.

## Newton-Cotes formulae and derivations

We now consider a polynomial interpolation for a function $f(x)$:

$$p(x) = \sum_{i=0}^{n} f(x_i) l_i(x), \qquad l_i(x) = \prod_{j=0,\ j\neq i}^{n} \frac{x - x_j}{x_i - x_j}, \qquad i, j = 0, 1, 2, ..., n \quad (7)$$

with nodes $x_i$ and corresponding function values $f(x_i)$, and substitute $p(x)$ into the integrand. So, we have

$$\int_a^b f(x)dx \approx \int_a^b p(x)dx = \sum_{i=0}^{n} f(x_i) \int_a^b l_i(x)dx = \sum_{i=0}^{n} A_i f(x_i), \qquad (8)$$

where

$$A_i = \int_a^b l_i(x)dx.$$

A formula of this form is called a **Newton-Cotes formula** if the nodes are equally spaced.

## Trapezium rule

The simplest case occurs if we set $n = 1$ with the nodes $x_0 = a$ and $x_1 = b$. Hence

$$l_0(x) = \frac{b - x}{b - a}, \qquad l_1(x) = \frac{x - a}{b - a}.$$

Consequently

$$A_0 = \int_a^b l_0(x)dx = \frac{b - a}{2} = \int_a^b l_1(x)dx = A_1.$$

The corresponding quadrature formula is

$$\int_a^b f(x)dx = \frac{b - a}{2} \cdot [f(a) + f(b)] + R. \qquad (9)$$

This is known as trapezium rule, where the error

$$R = -f''(\zeta)(b - a)^3/12,$$

can be derived by integrating

$$f(x) - p_1(x) = f''(\zeta)(x-a)(x-b)/2, \qquad \zeta \in [a,b].$$

If the interval $[a,b]$ is partitioned as before, we obtain the composite trapezium rule with non-uniformly spaced nodes:

$$\int_a^b f(x)dx \approx \frac{1}{2}\sum_{i=1}^{n}(x_i - x_{i-1}) \cdot [f(x_{i-1}) + f(x_i)]. \tag{10}$$

With uniform spacing $h = (b-a)/n$ we have

$$\int_a^b f(x)dx = \frac{h}{2}\left(f(a) + 2\sum_{i=1}^{n-1} f(a+ih) + f(b)\right) + R, \tag{11}$$

and the error estimate looks like this

$$|R| \le \frac{(b-a)h^2 M_2}{12}, \tag{12}$$

where $M_2 = \max|f''(x)|$ at $x \in [a,b]$.

## Simpson's rule

Let us consider a formula

$$\int_0^1 f(x)dx = A_0 f(0) + A_1 f(\frac{1}{2}) + A_2 f(1) \tag{13}$$

with the undetermined coefficients $A_0,\ A_1,\ A_2$. We determine these coefficients requiring that this formula is exact for the trial functions $f(x) = 1,\ x,\ x^2$. Thus, we obtain

$$1 = \int_0^1 dx = A_0 + A_1 + A_2,$$

$$\frac{1}{2} = \int_0^1 x dx = \frac{1}{2}A_1 + A_2,$$

$$\frac{1}{3} = \int_0^1 x^2 dx = \frac{1}{4}A_1 + A_2.$$

The solution of the system is $A_0 = 1/6,\ A_1 = 2/3,\ A_2 = 1/6$. Similar calculations for an arbitrary interval $[a, b]$ leads to the **Simpson's rule**

$$\int_a^b f(x)dx = \frac{b-a}{6}\left[f(a) + 4f(\frac{a+b}{2}) + f(b)\right] + R, \tag{14}$$

where the error is

$$R = -\frac{f^{(4)}(\zeta)}{90}\left[\frac{b-a}{2}\right]^5, \qquad \zeta \in [a, b].$$

Note that the Simpson's rule is exact formula for any cubic polynomial as $f^{(4)}(x) = 0$. The formula for the error can be derived by analysing the representation

$$\int_a^{a+2h} f(x)dx \approx \frac{h}{3}\left[f(a) + 4f(a+h) + f(a+2h)\right]$$

with $h = (b-a)/2$. Applying the Taylor's series up to the order $O(h^5)$ to the right-hand side of the equality, we obtain

$$2hf(a) + +2h^2 f'(a) + \frac{4}{3}h^3 f''(a) + \frac{2}{3}h^4 f'''(a) + \frac{100}{3 \cdot 5!}h^5 f^{(4)}(\zeta).$$

The same procedure applied to the left-hand side gives

$$2hf(a) + +2h^2 f'(a) + \frac{4}{3}h^3 f''(a) + \frac{2}{3}h^4 f'''(a) + \frac{32}{5!}h^5 f^{(4)}(\zeta).$$

Difference of these results yields the desired formula for error

$$R = -\frac{f^{(4)}(\zeta)}{90}h^5.$$

A composite Simpson's rule using an even number of subintervals with uniform spacing of nodes

$$x_i = a + ih, \qquad h = (b-a)/n, \qquad i = 0, 1, ..., n$$

is often used. This looks like this

$$\int_a^b f(x)dx = \frac{h}{3}\left[f(a) + 2\sum_{i=2}^{n/2} f(x_{2i-2}) + 4\sum_{i=1}^{n/2} f(x_{2i-1}) + f(b)\right] + R, \quad (15)$$

where the estimate for the error is given by

$$|R| \leq \frac{M_4}{180}(b-a)h^4, \quad (16)$$

where $M_4 = \max |f^{(4)}(x)|$ at $x \in [a,b]$.

## Adaptive Quadrature

Adaptive quadrature methods are intended to compute definite integrals by automatically taking into account the behaviour of the integrand. Ideally, the user supplies only the integrand $f$, the interval $[a,b]$, and the accuracy $\epsilon$ desired for computing the integral

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

The program then divides the interval into subintervals of varying length so that numerical integration on these subintervals will produce results of acceptable precision. The main idea is that if Simpson's rule on a given subinterval is not sufficiently accurate, that interval will be divided into two equal parts, and Simpson's rule will be used on each half. This procedure will be repeated in an effort to obtain an approximation to the integral with the same accuracy over all subintervals involved. At the end, we shall have computed the integral with $n$ applications of Simpson's rule

$$\int_a^b f(x)dx = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)dx = \sum_{i=1}^{n}(S_i + e_i) = \sum_{i=1}^{n} S_i + \sum_{i=1}^{n} e_i, \quad (17)$$

where $S_i$ is the Simpson's formula for the segment $[x_{i-1}, x_i]$ and $e_i$ is the local error. If

$$|e_i| \le \epsilon(x_i - x_{i-1})/(b-a),$$

then the total error will be bounded by

$$\Big|\sum_{i=1}^{n} e_i\Big| \le \sum_{i=1}^{n} |e_i| \le \frac{\epsilon}{b-a}\sum_{i=1}^{n}(x_i - x_{i-1}) = \epsilon. \tag{18}$$

Here is the short description of the Runge's method on how to evaluate local error. Let us assume that using Simpson's rule we have two calculation results with $n$ and $2n$ subintervals $(h = (x_i - x_{i-1})/n)$

$$\int_{x_{i-1}}^{x_i} f(x)dx = S_h(x_{i-1}, x_i) + c_i h^4$$

and

$$\int_{x_{i-1}}^{x_i} f(x)dx \approx S_{h/2}(x_{i-1}, x_i) + c_i \frac{h^4}{2^4}.$$

Eliminating the unknown coefficient $c_i$, we have

$$\int_{x_{i-1}}^{x_i} f(x)dx - S_{h/2}(x_{i-1}, x_i) \approx \frac{1}{2^4}\left(\int_{x_{i-1}}^{x_i} f(x)dx - S_h(x_{i-1}, x_i)\right),$$

and then,

$$\int_{x_{i-1}}^{x_i} f(x)dx - S_{h/2}(x_{i-1}, x_i) \approx \frac{S_{h/2}(x_{i-1}, x_i) - S_h(x_{i-1}, x_i)}{2^4 - 1}.$$

This approximate equality allows us to obtain estimate for the local error.

## Gaussian Quadrature

Let us require that a quadrature formula with nonuniform nodes spacing is exact for arbitrary polynomial of degree $2n-1$, that is

$$\int_a^b x^s dx = \sum_{i=1}^{n} A_i x_i^s, \qquad s = 0, 1, ..., 2n-1. \tag{19}$$

Thus we have $2n-1$ nonlinear equations for $2n-1$ unknowns $x_1, ..., x_n$ and $c_1, ..., c_n$. This problem has a unique solution and the results is called as **Gaussian quadrature**. For example, calculating

$$\int_{-1}^{1} f(x)dx,$$

if $n=2$, then $s=0,1,2$, and the corresponding system is like that

$$c_1+c_2=2, \qquad c_1x_1+c_2x_2=0, \qquad c_1x_1^2+c_2x_2^2=2/3, \qquad c_1x_1^3+c_2x_2^3=0.$$

Their solution is $c_1=c_2=1$, $x_1=-\frac{1}{\sqrt{3}}$ and $x_2=\frac{1}{\sqrt{3}}$. Finally, we obtain Gaussian quadrature formula for this case

$$\int_{-1}^{1} f(x)dx \approx f(-\frac{1}{\sqrt{3}})+f(\frac{1}{\sqrt{3}}),$$

which is exact result for arbitrary polynomial with degree less than 3.