4.8

NUMERICAL INTEGRATION

The general problem of numerical integration may be stated as follows. Given a set of data points $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ of a function $y = f(x)$, where $f(x)$ is not known explicitly, it is required to compute the value of the definite integral

$$I = \int_a^b y \, dx$$

The idea is to replace $f(x)$ by an interpolating polynomial $\emptyset(x)$ and obtain, on integration, an approximate value of the definite integral. Thus, different integration formulae can be obtained depending upon the type of the interpolation formula used. We derive in this module a general formula for numerical integration using Newton's forward difference formula.

Let the interval $[a, b]$ be divided into n equal subintervals such that $a = x_0 < x_1 < x_2 < \cdots x_n = b$. Clearly, $x_n = x_0 + nh$. Hence the integral becomes

$$I = \int_{x_0}^{x_n} y \, dx.$$

Approximating y by Newton's forward difference formula, we obtain

$$I = \int_{x_0}^{x_n} \left[y_0 + p\Delta y_0 + \frac{p(p-1)}{2}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{6}\Delta^3 y_0 + \cdots \right] dx.$$

Since $x = x_0 + ph, dx = hdp$ and hence the above integral becomes

$$I = h\int_0^n \left[y_0 + p\Delta y_0 + \frac{p(p-1)}{2}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{6}\Delta^3 y_0 + \cdots\right] dp,$$

which gives on simplification

$$\int_{x_0}^{x_n} y\, dx = nh\left[y_0 + \frac{n}{2}\Delta y_0 + \frac{n(2n-3)}{12}\Delta^2 y_0 + \frac{n(n-2)^2}{24}\Delta^3 y_0 + \cdots\right].$$

From this general formula, we can obtain different integration formulae by putting $n = 1, 2, 3, \ldots,$ etc We derive here a few of these formulae but it should be remarked that the trapezoidal and Simpson's $1/3$ rules are found to give sufficient accuracy for use in practical problems.

Trapezoidal Rule

Setting $n = 1$ in the general formula, all differences higher than the first will become zero and we obtain

$$\int_{x_0}^{x_1} y\, dx = h\left(y_0 + \frac{1}{2}\Delta y_0\right) = h\left[y_0 + \frac{1}{2}(y_1 - y_0)\right] = \frac{h}{2}(y_0 + y_1).$$

For the next interval $[x_1, x_2]$, we deduce similarly

$$\int_{x_1}^{x_2} y\, dx = \frac{h}{2}(y_1 + y_2)$$

and so on. For the last interval $[x_{n-1}, x_n]$, we have

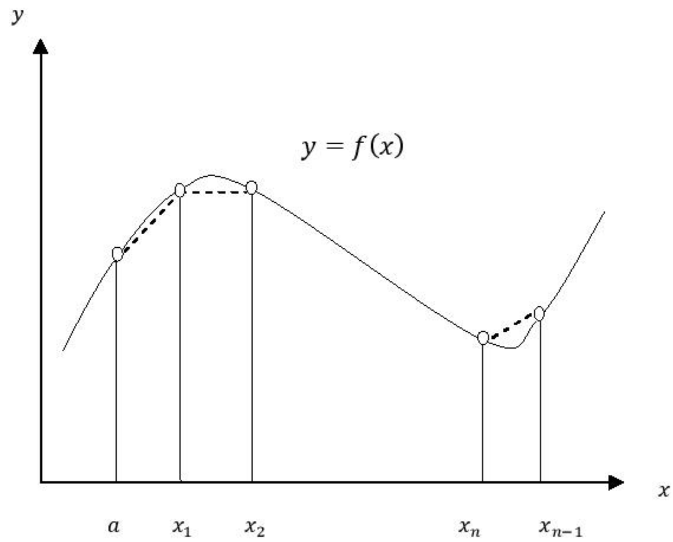
$$\int_{x_{n-1}}^{x_n} y\, dx = \frac{h}{2}(y_{n-1} + y_n).$$

Combining all these expressions, we obtain the rule

$$\int_{x_0}^{x_n} y\, dx = \frac{h}{2}[y_0 + 2(y_1 + y_2 + \cdots + y_{n-1}) + y_n],$$

which is known as the trapezoidal rule.

The geometrical significance of this rule is that the curve $y = f(x)$ is replaced by n straight lines joining the points (x_0, y_0) and (x_1, y_1); (x_1, y_1) and (x_2, y_2); ... ; (x_{n-1}, y_{n-1}) and (x_n, y_n). The area bounded by the curve $y = f(x)$, the ordinates $x = x_0$ and $x = x_n$, and the $x - axis$ is then approximately equivalent to the sum of the areas of the n trapeziums obtained.



The error of the trapezoidal formula can be obtained in the following way. Let $y = f(x)$ be continuous, well-behaved, and possess continuous derivatives in $[x_0, x_n]$. Expanding y in a Taylor's series around $x = x_0$, we obtain

$$\int_{x_0}^{x_1} y\,dx = \int_{x_0}^{x_1} \left[y_0 + (x - x_0)y_0' + \frac{(x-x_0)^2}{2} y_0'' + \cdots \right] dx$$

$$= hy_0 + \frac{h^2}{2} y_0' + \frac{h^3}{6} y_0'' + \cdots$$

Similarly,

$$\frac{h}{2}(y_0 + y_1) = \frac{h}{2}\left(y_0 + y_0 + hy_0' + \frac{h^2}{2}y_0'' + \frac{h^3}{6}y_0'' + \cdots\right)$$

$$= hy_0 + \frac{h^2}{2}y_0' + \frac{h^3}{4}y_0'' + \cdots,$$

so

$$\int_{x_0}^{x_1} y\,dx - \frac{h}{2}(y_0 + y_1) = -\frac{1}{12}h^3y_0'' + \cdots,$$

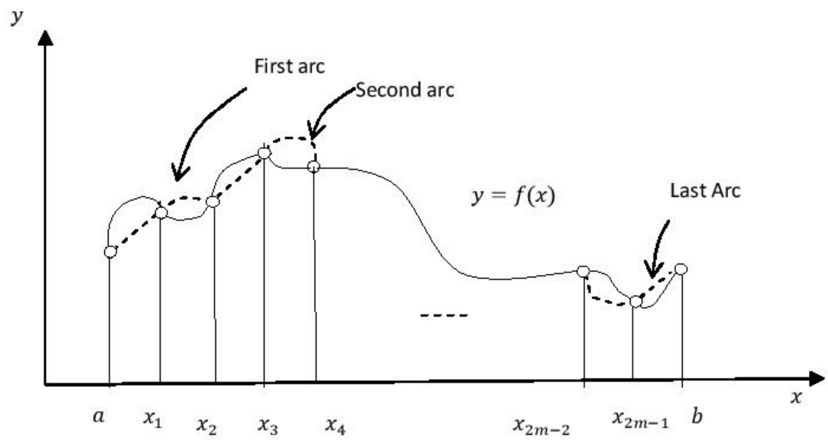
which is the error, in the interval $[x_0, x_1]$. Proceeding in a similar manner we obtain the errors in the remaining subintervals, i.e, $[x_1, x_2], [x_2, x_3], \ldots$ and $[x_{n-1}, x_n]$. We thus have

$$E = -\frac{1}{12}h^3(y_0'' + y_1'' + \cdots + y_{n-1}''),$$

where E is the total error. Therefore by determining the maximum and minimum values of y'' in $[x_0, x_n]$ one can obtain bounds for the error.

Simpson's 1/3-Rule

This rule is obtained by putting $n = 2$ in the above general formula, i.e. by replacing the curve by $n/2$ arcs of second-degree polynomials or parabolas.



We have then

$$\int_{x_0}^{x_2} y\,dx = 2h\left(y_0 + \Delta y_0 + \frac{1}{6}\Delta^2 y_0\right) = \frac{h}{3}(y_0 + 4y_1 + y_2).$$

Similarly,

$$\int_{x_2}^{x_4} y\,dx = \frac{h}{3}(y_2 + 4y_3 + y_4)$$

.

.

.

and finally

$$\int_{x_{n-2}}^{x_n} y\,dx = \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n).$$

Summing up, we obtain

$$\int_{x_0}^{x_n} y\,dx = \frac{h}{3}[y_0 + 4(y_1 + y_3 + y_5 + \cdots + y_{n-1}) + 2(y_2 + y_4 + y_6 + \cdots + y_{n-2}) + y_n],$$

which is known as Simpson's 1/3-rule, or simply Simpson's rule. It should be noted that this rule requires the division of the whole range into an even number of subintervals of width h.

Following the method outlined in the case of trapezoidal rule, it can be shown that the error in Simpson's rule is given by

$$E = \int_a^b y\,dx - \frac{h}{3}[y_0 + 4(y_1 + y_3 + y_5 + \cdots + y_{n-1}) + 2(y_2 + y_4 + y_6 + \cdots + y_{n-2}) + y_n]$$

$$= -\frac{1}{180}h^5\left(y_0^{iv} + y_1^{iv} + \cdots + y_{n-1}^{iv}\right),$$

Therefore by determining the maximum and minimum values of y^{iv} in $[x_0, x_n]$ one can obtain bounds for the error.

Simpson's 3/8-Rule

Setting $n = 3$ in the general formula, we observe that all the differences higher than the third will become zero and we obtain

$$\int_{x_0}^{x_3} y\,dx = 3h\left(y_0 + \frac{3}{2}\Delta y_0 + \frac{3}{4}\Delta^2 y_0 + \frac{1}{8}\Delta^3 y_0\right)$$

$$= 3h\left[y_0 + \frac{3}{2}(y_1 - y_0) + \frac{3}{4}(y_2 - 2y_1 + y_0) + \frac{1}{8}(y_3 - 3y_2 + 3y_1 - y_0)\right]$$

$$\frac{3h}{8}(y_0 + 3y_1 + 3y_2 + y_3).$$

Similarly

$$\int_{x_3}^{x_6} y\,dx = \frac{3h}{8}(y_3 + 3y_4 + 3y_5 + y_6)$$

and so on. Summing up all these, we obtain

$$\int_{x_0}^{x_n} y\,dx = \frac{3h}{8}[(y_0 + 3y_1 + 3y_2 + y_3) + (y_3 + 3y_4 + 3y_5 + y_6) + \ldots + (y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n)]$$

$$= \frac{3h}{8}(y_0 + 3y_1 + 3y_2 + 2y_3 + 3y_4 + 3y_5 + 2y_6 + \cdots + 2y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n).$$

This rule, called Simpson's (3/8)-rule, is not so accurate as Simpson's rule. The total error in this case is

$$E = -\frac{3}{80}h^5\left(y_0^{iv} + y_1^{iv} + \cdots + y_{n-1}^{iv}\right),$$

Therefore by determining the maximum and minimum values of y^{iv} in $[x_0, x_n]$ one can obtain bounds for the error.

Solved Problems

Problem 1: Find, from the following table, the area bounded by the curve and the $x-$axis from $x = 7.47$ to $x = 7.52$.

x	$f(x)$
7.47	1.93
7.48	1.95
7.49	1.98
7.50	2.01
7.51	2.03
7.52	2.06

Solution: We know that

$$\text{Area} = \int_{7.47}^{7.52} f(x)dx$$

with $h = 0.01$, the trapezoidal rule gives

Area= $\frac{0.01}{2}[1.93 + 2(1.95 + 1.98 + 2.01 + 2.03) + 2.06] = 0.0996.$

Problem 2: Evaluate

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

correct to three decimal places by using both the trapezoidal and Simpson's rules with $h = 0.5,\ 0.25$ and 0.125 respectively.

Solution:

i. $h = 0.5$: The values of x and y are tabulated below:

x	y
0.0	1.0000
0.5	0.6667
1.0	0.5000

a) Trapezoidal rule gives

$$I = \frac{1}{4}[1.0000 + 2(0.6667) + 0.5] = 0.70835.$$

b) Simpson's rule gives

$$I = \frac{1}{6}[1.0000 + 4(0.6667) + 0.5] = 0.6945.$$

ii. $h = 0.25$: The tabulated values of x and y are given below:

x	y
0.00	1.0000
0.25	0.8000
0.50	0.6667
0.75	0.5714
1.00	0.5000

a) Trapezoidal rule gives

$$I = \frac{1}{8}[1.0 + 2(0.8000 + 0.6667 + 0.5714) + 0.5]$$

$$= 0.6970$$

b) Simpson's rule gives

$$I = \frac{1}{12}[0.1 + 4(0.8000 + 0.5714) + 2(0.6667) + 0.5]$$

$$= 0.6932.$$

iii. Finally, we take $h = 0.125$: the tabulated values of x and y are

x	y

0	1.0
0.125	0.8889
0.250	0.8000
0.375	0.7273
0.5	0.6667
0.625	0.6154
0.750	0.5714
0.875	0.5333
1.0	0.5

a) Trapezoidal rule gives

$$I = \frac{1}{16}[1.0 + 2(0.8889) + 0.8000 + 0.7273 + 0.6667 + 0.6154 + 0.5714 + 0.5333) + 0.5]$$

$$= 0.6941.$$

b) Simpson's rule gives

$$I = \frac{1}{24}[1.0 + 4(0.8889 + 0.7273 + 0.6154 + 0.5333) + 2(0.8000 + 0.6667 + 0.5714) + 0.5]$$

$$= 0.6932.$$

Hence the value of I may be taken to be equal to 0.693, correct to three decimal places. The exact value of I is $\log_e 2$, which is equal to 0.693147 ….. This example demonstrates that, in general, Simpson's rule yields more accurate results than the trapezoidal rule.

EXERCISE

1. Evaluate

   A. $\int_0^{\pi} t \sin t \, dt$

   B. $\int_{-2}^{2} \frac{t}{5+2t} dt$

   using the trapezoidal rule.

2. Estimate the value of the integral $\int_1^3 \frac{1}{x} dx$ by Simpson's rule, with 4 stripes and 8 strips, respectively. Determine the error by direct integration.
3. Compute the value of $I = \int_0^1 \frac{dx}{1+x^2}$ by using the trapezoidal rule with $h = 0.5, 0.25$ and 0.125.
4. Evaluate $\int_0^1 e^{-x^2} dx$ by means of trapezoidal and Simpson's rule with $n = 10$. Also obtain error bounds in each method.

ANSWERS

1. (A). 3.14 (B). -0.747.
2. 1.1000, error$= 0.0014$; 1.0987, error$= 0.0001$.
3. $0.77500, 0.78279, 0.78475$.
4. Trapezoidal rule: 0.746211; $-0.000614 \leq E \leq 0.001667$.
   Simpson's rule: 0.746825; $-0.000007 \leq E \leq 0.000005$.