Numerical Integration
Trapezoid and Simpson's Rules

P. Sam Johnson

October 21, 2014

Overview

We derive and analyse numerical methods for evaluating definite integrals. The integrals are mainly of the form

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

with $[a, b]$.

Most such integrals cannot be evaluated explicitly. It is often faster to integrate them numerically rather than evaluating them exactly using a complicated antiderivative of $f(x)$.

The approximation of $I(f)$ is usually referred to as **numerical integration** or **quadrature**.

There are many numerical methods for evaluating

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

but most can be made to fit within the following simple framework.

For the integrand $f(x)$, find an approximating family $\{f_n(x) : n \geq 1\}$ and define

$$I_n(f) = \int_a^b f_n(x)\, dx = I(f_n).$$

We usually require the approximations $f_n(x)$ to satisfy

$$\|f - f_n\|_\infty \to 0 \quad \text{as} \quad n \to \infty.$$

For the error,

$$\begin{aligned} E_n(f) &= I(f) - I_n(f) \\ &= \int_a^b [f(x) - f_n(x)]\, dx. \end{aligned}$$

Hence

$$\begin{aligned} |E_n(f)| &\leq \int_a^b |f(x) - f_n(x)|\, dx. \\ &\leq (b-a)\|f - f_n\|_\infty. \end{aligned}$$

Most numerical integration methods can be viewed within this framework.

Most numerical integrals $I_n(f)$ will have the following form when they are evaluated :

$$I_n(f) = \sum_{j=1}^{n} w_{j,n} f(x_{j,n}) \quad n \geq 1.$$

The coefficients $w_{j,n}$ are called the **integration weights** or **quadrature weights**; and the points $x_{j,n}$ are the **integration nodes**, usually chosen in $[a, b]$.

Standard methods have nodes and weights that have simple formulas or else they are tabulated in tables that are readily available.

Thus there is usually no need to explicitly construct the functions $f_n(x)$ of

$$I(f_n) = \int_a^b f_n(x)\, dx$$

although their role in defining $I_n(f)$ may be useful to keep in mind.

Most numerical integration formulas are based on defining $f_n(x)$ in

$$I(f_n) = \int_a^b f_n(x)\, dx$$

by using polynomial or piecewise polynomial interpolation.

Formulas using such interpolation **with evenly spaced node points** are derived and discussed.

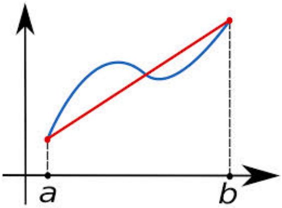
Trapezoidal rule

The simple trapezoidal rule is based on approximating $f(x)$ by the straight line joining $(a, f(a))$ and $(b, f(b))$.

By integrating the formula for this straight line, we obtain the approximation

$$I_1(f) = \left(\frac{b-a}{2}\right)[f(a) + f(b)].$$

This is of course the area of the trapezoid shown below.



Note that the equation of the line joining $(a, f(a))$ and $(b, f(b))$ is

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

We further assume that $f(x)$ is **twice continuously differentiable** on $[a, b]$. The error for the trapezoidal rule is

$$\begin{aligned} E_1(f) &= \int_a^b f(x)\, dx - \left\{ \frac{f(a)+f(b)}{2}(b-a) \right\} \\ &= \int_a^b (x-a)(x-b) f[a,b,x]\, dx. \end{aligned}$$

Here $f[a, b, x]$ is the 3-rd order Newton divided difference of f about the nodes a, b and x.

Using the integral mean value theorem

$$\begin{aligned} E_1(f) &= f[a,b,\xi] \int_a^b (x-a)(x-b)\ dx \quad \text{for some } \xi \in [a,b]. \\ &= \left[\frac{1}{2} f''(\eta)\right] \left[\frac{-1}{6}(b-a)^3\right] \quad \text{for some } \eta \in [a,b]. \end{aligned}$$

Thus the error is

$$E_1(f) = -\frac{(b-a)^3}{12} f''(\eta) \quad \text{for some } \eta \in [a,b].$$

If the value $b - a$ is not sufficiently small, the trapezoidal rule is not of much use.

For such case, we break the integral into a sum of integrals over smal subintervals, and then apply the trapezoidal rule to each of these integrals.

Composite Trapezoidal Rule

Let $n \geq 1$, $h = \frac{b-a}{n}$, and $x_j = a + jh$ for $j = 0, 1, 2, \ldots, n$. Then

$$\begin{aligned} I(f) &= \int_a^b f(x)\, dx = \sum_{j=1}^{n} \int_{x_{j-1}}^{x_j} f(x)\, dx \\ &= \sum_{j=1}^{n} \left\{ \left(\frac{h}{2}\right) [f(x_{j-1}) + f(x_j)] - \frac{h^3}{12} f''(\eta_j) \right\} \end{aligned}$$

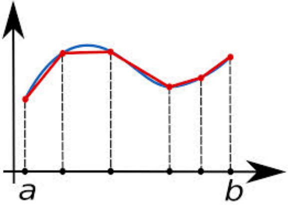
with $x_{j-1} \leq \eta_j \leq x_j$.

The first terms in the sum can be combined to give the **composite trapezoidal rule**,

$$I_n(f) = \frac{h}{2}[f_0 + 2(f_1 + f_2 + \cdots + f_{n-1}) + f_n] \quad n \geq 1$$

with $f(x_j) \equiv f_j$.

Error in Composite Trapezoidal Rule



The error in $I_n(f)$ is given by

$$E_n(f) = I(f) - I_n(f) = \sum_{j=1}^{n} -\frac{h^3}{12} f''(\eta_j) = -\frac{h^3 n}{12} \left[\frac{1}{n} \sum_{j=1}^{n} f''(\eta_j) \right].$$

For the term in brackets,

$$\min_{a \le x \le b} f''(x) \le M \equiv \frac{1}{n} \sum_{j=1}^{n} f''(\eta_j) \le \max_{a \le x \le b} f''(x).$$

Since $f''(x)$ is continuous for $a \leq x \leq b$, it must attain all values between its minimum and maximum at some point of $[a, b]$.

Hence

$$f''(\eta) = M \quad \text{for some } \eta \in [a, b].$$

Thus we can write

$$E_n(f) = \frac{-(hn)h^2}{12} f''(\eta) = \frac{-(b-a)h^2}{12} f''(\eta) \quad \text{for some } \eta \in [a, b].$$

There is no reason why the subintervals $[a_{j-1}, x_j]$ must all have equal length, but it is customary to first introduce the general principles involved in this way.

Although this is also the customary way in which the method is applied, there are situations in which it is desirable to vary the spacing of the nodes.

Simpson's Rule

To improve upon the simple trapezoidal rule, we use a quadratic interpolating polynomial $p_2(f)$ to approximate $f(x)$ on $[a, b]$.

Let $c = \frac{(a+b)}{2}$, and define

$$I_2(f) = \int_a^b \left[\frac{(x-c)(x-b)}{(a-c)(a-b)} f(a) + \frac{(x-a)(x-b)}{(c-a)(c-b)} f(c) + \frac{(x-a)(x-c)}{(b-a)(b-c)} f(b) \right] dx.$$

Carrying out the integration, we obtain

$$I_2(f) = \frac{h}{3} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \quad \text{where } h = \frac{b-a}{2}.$$

This is called **Simpson's rule** or **Simpson's $\frac{1}{3}$-rule**.

Error

For the error,

$$E_2(f) = I(f) - I_2(f) = \int_a^b (x-a)(x-c)(x-b)\ f[a,b,c,x]\ dx.$$

If we further assume that f is four times continuously differentiable on $[a, b]$, we can apply the integral mean value theorem.

Thus

$$E_2(f) = -\frac{h^5}{90} f^{(4)}(\eta) \quad \text{where } \eta \in [a,b].$$

From this we see that $E_2(f) = 0$ if $f(x)$ is a polynomial of degree at most three, even though quadratic interpolation is exact only if $f(x)$ is a polynomial of degree at most two.

This results in Simpson's rule being much more accurate than the trapezoidal rule.

Composite Simpson's $\frac{1}{3}$-rule

Let $n \geq 2$, and even, $h = \frac{b-a}{n}$, and $x_j = a + jh$ for $j = 0, 1, 2, \ldots, n$. Then

$$\begin{aligned} I(f) &= \int_a^b f(x)\, dx = \sum_{j=1}^{n/2} \int_{x_{2j-2}}^{x_{2j}} f(x)\, dx \\ &= \sum_{j=1}^{n/2} \left\{ \left(\frac{h}{3}\right) \left[f_{2j-2} + 4f_{2j-1} + f_{2j}\right] - \frac{h^5}{90} f^{(4)}(\eta_j) \right\} \end{aligned}$$

with $x_{2j-2} \leq \eta_j \leq x_{2j}$. Simplifying the first terms in the sum, we obtain the **composite Simpson's $\frac{1}{3}$-rule**,

$$I_n(f) = \frac{h}{3}\Big[f_0 + 4(f_1 + f_3 + \cdots + f_{n-1}) + 2(f_2 + f_4 + \cdots + f_{n-2}) + f_n\Big]$$

with $f(x_j) \equiv f_j$.

Error in Composite Simpson's $\frac{1}{3}$-rule

For the error, as with the trapezoidal rule

$$E_n(f) = I(f) - I_n(f) = \frac{-h^5(n/2)}{90}\ \frac{2}{n}\sum_{j=1}^{n/2} f^{(4)}(\eta_j).$$

Since f is 4-times differentiable, we get

$$E_n(f) = \frac{-h^4(b-a)}{180} f^{(4)}(\eta) \quad \text{where } \eta \in [a, b].$$

Numerical Integration using Newton's Forward Difference Formula

If we are given a set of data points $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ of a function

$$y = f(x)$$

where $f(x)$ is not known explicitly, it is required to compute the value of the definite integral

$$I = \int_a^b y \; dx.$$

Different integration formulae can be obtained depending upon the type of the interpolation formula used.

We derive 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

$$a = x_0 < x_1 < \cdots < x_n = b.$$

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

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

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

From this **general formula**, we can obtain different integration formulae by putting $n = 1, 2, 3, \ldots$.

Some formulae have been already discussed.

Trapezoidal Rule

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

$$\int_{x_0}^{x_1} y \, dx = \frac{h}{2}(y_0 + y_1)$$

with error term

$$E = \frac{-(b-a)^3}{12} y''(\eta) \quad \text{for some } \eta \in [x_0, x_1]$$

which is known as the **trapezoidal rule**.

Simpson's $\frac{1}{3}$-rule

Setting $n = 2$ in the general formula, we obtain

$$\int_{x_0}^{x_2} y\ dx = \frac{h}{3}(y_0 + 4y_1 + y_2)$$

with error term

$$E = \frac{-h^5}{90} y^{(4)}(\eta) \quad \text{where } \eta \in [x_0, x_2]$$

which is known as the **Simpson's $\frac{1}{3}$-rule**.

It should be noted that this rule requires the division of the whole range into an **even number of subintervals** of width h.

Simpson's $\frac{3}{8}$-rule

Setting $n = 3$ in the general formula, we obtain

$$\int_{x_0}^{x_3} y \; dx = \frac{3h}{8}(y_0 + 3y_1 + 3y_2 + y_3)$$

with error term

$$E = \frac{-3h^5}{80} y^{(4)}(\eta) \quad \text{where } \eta \in [x_0, x_3]$$

which is known as the **Simpson's $\frac{3}{8}$-rule**.

It should be noted that this rule requires the number of subintervals will have to be a **multiple of three**.

Boole's Rule

Setting $n = 4$ in the general formula, we obtain

$$\int_{x_0}^{x_4} y \, dx = \frac{2h}{45}(7y_0 + 32y_1 + 12y_2 + 32y_3 + 7y_4)$$

with error term

$$E = \frac{-8h^7}{945} y^{(4)}(\eta) \quad \text{where } \eta \in [x_0, x_4]$$

which is known as the **Boole's-rule**.

It should be noted that this rule requires the number of subintervals will have to be a **multiple of four**.

Weddle's Rule

Setting $n = 6$ in the general formula, we obtain

$$\int_{x_0}^{x_6} y \, dx = \frac{3h}{10}(y_0 + 5y_1 + y_2 + 6y_3 + y_4 + 5y_5 + y_6)$$

with error term

$$E = \frac{-h^7}{140} y^{(6)}(\eta) \quad \text{where } \eta \in [x_0, x_6]$$

which is known as the **Weddle's-rule**.

It should be noted that this rule requires the number of subintervals will have to be a **multiple of six**.

Double Integration

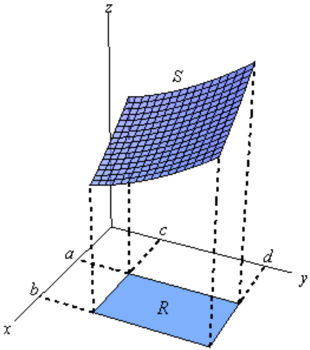
Consider $z = f(x, y)$ over the rectangle

$$R = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}.$$

We can evaluate

$$\int_a^b \int_c^d f(x, y)\ dy\ dx$$

by repeatedly applying the "trapezoidal rule for one dimension case".



Trapezoidal $2D$ Rule

Given that the interval $[a, b]$ is subdivided into m subintervals

$$\Big\{[x_{i-1}, x_i]\Big\}_{i=1}^{m} \text{ of equal width } h = \frac{b-a}{m}$$

by using the equally spaced sample points $x_i = x_0 + ih$ for $i = 0, 1, \ldots, m$.

Also, assume that the interval $[c, d]$ is subdivided into n subintervals

$$\Big\{[y_{j-1}, x_j]\Big\}_{j=1}^{n} \text{ of equal width } k = \frac{d-c}{n}$$

by using the equally spaced sample points $y_j = y_0 + jk$ for $j = 0, 1, \ldots, n$.

Composite Trapezoidal $2D$ Rule

The **composite Trapezoidal rule** is

$$\int_a^b \int_c^d f(x,y)\,dy\,dx \approx \frac{hk}{4}\Big\{ f(a,c) + f(b,c) + f(a,d) + f(b,d) + 2\Big\{ \sum_{i=1}^{m-1} f(x_i,c) + \sum_{i=1}^{m-1} f(x_i,d) + \sum_{j=1}^{n-1} f(a,y_j) + \sum_{j=1}^{n-1} f(b,y_j) \Big\} + \sum_{j=1}^{n-1} \Big(\sum_{i=1}^{m-1} f(x_i,y_j) \Big) \Big\}.$$

In a similar way, **composite Simpson's** $2D$ **rules** for the evaluation of double integral can be derived by repeatedly applying "Simpson's rules for one dimension".

References

- Richard L. Burden and J. Douglas Faires, "*Numerical Analysis – Theory ad Applications*", Cengage Learning, New Delhi, 2005.
- Kendall E. Atkinson, "*An Introduction to Numerical Analysis*", John Wiley & Sons, Delhi, 1989.