# Numerical Differentiation and Integration

If the values of a function are given at a few points, say $x_0, x_1, \ldots, x_n$, can that information be used to estimate a derivative $f'(c)$ or an integral $\int_a^b f(x)dx$? The answer is a qualified Yes.

A general approach to numerical differentiation and integration can be based on polynomial interpolation.

♣ **Numerical Differentiation**. Numerical differentiation formulae find their most important application in the numerical solution of differential equations.

To numerically calculate the derivative of $f(x)$, begin by recalling the definition of the derivative:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

This justifies using

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} \equiv D_h f(x)$$

for small values of $h$. $D_h f(x)$ is called a *numerical derivative* of $f(x)$ with stepsize $h$.

If both $f(x+h)$ and $f(x-h)$ are defined, then a superior formula is

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

This is derived from two cases of Taylor's Theorem, namely

$$f(x+h) = f(x) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f^{(3)}(\xi_1)$$

$$f(x-h) = f(x) = f(x) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{6} f^{(3)}(\xi_2)$$

On subtracting one of these from the other, we obtain

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{h^2}{12} \left[ f^{(3)}(\xi_1) + f^{(3)}(\xi_2) \right]$$

This is a more favorable result, because of the $h^2$ term in the error. Notice, however, the presence of $f^{(3)}$ in the error. This error term is applicable if $f^{(3)}$ exists.

Using the Taylor series of higher order for $f(x+h)$ and $f(x-h)$, we obtain

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} - \frac{h^2}{12} \left[ f^{(4)}(\xi) \right]$$

This formula is often used in the numerical solution of second-order differential equations.

♡ **Differentiation Using Interpolation**. Let $P_n(x)$ denote the degree $n$ polynomial that interpolates $f(x)$ at $n+1$ node points $x_0, x_1, \ldots, x_n$. To calculate $f(x)$ at some point $x = c$, use $f'(c) \approx p'_n(c)$. The error for the this formula can be obtained from the interpolation error formula. The main result is given in the following theorem:

**Theorem 1.** *Assume $f(x)$ has $n+2$ continuous derivatives on an interval $[a,b]$. Let $P_n(x)$ denote the degree $n$ polynomial that interpolates $f(x)$ at $n+1$ node points $x_0, x_1, \ldots, x_n$ in $[a,b]$, and let $c$ be an arbitrary given point in $[a,b]$. Then*

$$f'(c) - P_n'(c) = \Phi_n(c)\frac{f^{n+2)}(\xi_1)}{(n+2)!} + \Phi_n'(c)\frac{f^{n+2)}(\xi_2)}{(n+2)!}$$

*with*

$$\Phi_n(c) = (c-x_0)(c-x_1)\cdots(c-x_n)$$

*The constants $\xi_1$ and $\xi_2$ are unknown points of $[a,b]$.*

♣ **Richardson's Extrapolation**. A procedure known as Richardson's Extrapolation is frequently employed to coax more accuracy out of some numerical formulas.

To examine the extrapolation technique, suppose $N_1(h)$ is a numerical formula that produces approximations of order $O(h^2)$ to an unknown value $M$. In addition, assume that the error form for the approximation of $N(h)$ to $M$ can be expressed as

$$M = N(h) + K_1 h^2 + O(h^4), \tag{1}$$

where $K_1$ is a constant. Replacing $N(h)$ by $N_1(h)$ and $h$ by $h/2$ in (1) gives a new, and presumably more accurate, approximation $N_1(h/2)$ that satisfies

$$M = N_1(\frac{h}{2}) + K_1\frac{h^2}{4} + [O(\frac{h^4}{2})]. \tag{2}$$

Multiplying (1) by 4/3 and subtracting 1/3 of (2) gives

$$M = \frac{4N_1(\frac{h}{2}) - N_1(h)}{3} + O(h^4), \tag{3}$$

so

$$N_2(h) = \frac{4N_1(\frac{h}{2}) - N_1(h)}{3}. \tag{4}$$

Using this notation, the approximations to $M$ shown in the following table can be generated.

| | |
|---|---|
| $N_1(h)$ | |
| $N_1(\frac{h}{2})$ | $N_2(h)$ |
| $N_1(\frac{h}{4})$ | $N_2(\frac{h}{2})$ |
| $N_1(\frac{h}{8})$ | $N_2(\frac{h}{4})$. |

If, in addition, a number $K_2$ exists so that (1) can be expressed as

$$M = N(h) + K_1 h^2 + K_2 h^4 + O(h^6). \tag{5}$$

Then we may increase the accuracy by defining

$$N_3(h) = \frac{4^2 N_2(\frac{h}{2}) - N_2(h)}{4^2 - 1} = \frac{16N_2(\frac{h}{2}) - N_2(h)}{15}.$$

This process can be extended to $m$ such columns provided that the error form for the approximation of $N(h)$ can be expressed as

$$M = N(h) + \sum_{j=1}^{m-1} K_j h^{2j} + O(h^{2m})$$

for some collection of constant $K_j$. The $O(h^{2j})$ approximations are generated recursively by the formula:

$$N_j(h) = \frac{4^{j-1} N_j(\frac{h}{2}) - N_{j-1}(h)}{4^{j-1} - 1}$$

for $j = 2, 3, \ldots, m$.

**Example.** *According to the Taylor's Theorem, if $f(x)$ is continuously differentiable, then*

$$f'(x_0) = \frac{1}{2h}[f(x_0 + h) - f(x_0 - h)] - \frac{h^2}{6} f^{(3)}(x_0) - \frac{h^4}{120} f^{(5)}(x_0) + O(h^6).$$

*Let $f(x) = xe^x$, $x_0 = 2$ and $h = 0.2$. Then using the notation in the preceding discussion, we have*

$$M = f'(2), \quad N(0.2) = \frac{1}{0.4}[f(2.2) - f(1.8)], \quad K_1 = \frac{0.2^2}{6} f^{(3)}(2) \quad \text{and} \quad K_2 = \frac{0.2^4}{120} f^{(5)}(2).$$

*The exact value of $f'(x) = xe^x + e^x$ at $x_0 = 2$ is 22.167168. We construct an extrapolation table for $f(x)$, shown in the next table.*

| | | |
|---|---|---|
| $N_1(0.2) = 22.41416$ | | |
| $N_1(0.1) = 22.22878$ | $N_2(0.2) = \frac{4N_1(0.1) - N_1(0.2)}{3} = 22.16699$ | |
| $N_1(0.05) = 22.18256$ | $N_2(0.1) = \frac{4N_1(0.05) - N_1(0.1)}{3} = 22.16715$ | $N_3(0.2) = \frac{16N_2(0.1) - N_2(0.2)}{15} = 22.16716$ |

♣ **Integration**. The definite integral

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

is defined in the calculus as a limit of what are called *Riemann Sums.* It is known that

$$I(f) = F(b) - F(a),$$

where $F(x)$ is any antiderivative of $f(x)$; this is the *Fundamental Theorem of Calculus.* Many integrals can be evaluated using this formula; nonetheless, some integrands $f(x)$ do not have antiderivatives expressible in terms of elementary functions. Examples of such integrals are

$$\int_0^1 e^{-x^2} dx, \qquad \int_0^\pi \frac{\sin x}{x} dx\,.$$

Other methods such as power series expansion or numerical integration are needed for evaluating such integrals.

A *Composite Rule* is one obtained by applying an integration formula for a single interval to each subinterval of a partitioned interval.

♡ **Integration Using Interpolation**. Suppose we want to evaluate the integral

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

We can select nodes $x_0, x_1, \ldots, x_n$ in $[a,b]$ and set up a Lagrange interpolation process. For $k = 0, 1, \ldots, n$, define

$$L_{n,k} = \prod_{\substack{j=0 \\ j\neq 0}} \frac{x - x_j}{x_k - x_j}$$

The Lagrange polynomial

$$L_n(x) = \sum_{k=0}^{n} f(x_k) L_{n,k}(x)$$

interpolate $f(x)$ at the nodes. Then

$$\int_a^b f(x)dx \approx \int_a^b L_n(x)dx = \sum_{k=0}^{n} A_k f(x_k), \tag{6}$$

where

$$A_k = \int_a^b L_{n,k} dx\,.$$

A formula in the form (6) is called a *Newton-Cotes Formula*, if the nodes are equally spaced.

♠ **Trapezoid Rule**. The simplest case of the Lagrange polynomial results if $n = 1$ and the nodes are $a = x_0,\ x_1 = b$. The polynomials are

$$L_{1,0} = \frac{b - x}{b - a} \qquad L_{1,1} = \frac{x - a}{b - a}$$

Consequently,

$$A_0 = \int_a^b L_{1,0}(x)dx = \frac{1}{2}(b - a) = \int_a^b L_{1,1}(x)dx = A_1$$

The corresponding quadrature formula is

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

This is known as the *Trapezoid Rule*. Its error term is

$$-\frac{1}{12}(b - a)^3 f''(\xi)$$

where $\xi \in (a, b)$. This can be determined by integrating the error term in the polynomial approximation

$$f(x) - L_1(x) = \frac{1}{2} f''(\xi_x)(b - a)(x - b),$$

and employing the Mean-Value Theorem for Integrals.

The *Composite Trapezoid Rule* with uniform spacing (i.e., $h = (b-a)/n$ and $x_j = a + jh$) takes the form

$$\int_a^b f(x)dx \approx \frac{h}{2}\left[f(a) + 2\sum_{j=1}^{n-1} f(a+jh) + f(b)\right]$$

The error term for the composite trapezoid rule is

$$-\frac{1}{2}(b-a)h^2 f''(\xi)$$

where $\xi \in (a,b)$.

♠ **Simpson's Rule**. The *Simpson's Rule* is resulted from interpolating $f(x)$ with the Lagrange polynomial $L_2(x)$ and choosing $n = 2$ and $a = x_0$, $x_1 = (a+b)/2$ and $x_2 = b$. The formula is

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

The error term is

$$-\frac{1}{90}(b-a)^5 f^{(4)}(\xi)$$

where $\xi \in (a,b)$. The *Composite Simpson's Rule* with even number of subintervals and uniform spacing (i.e., $n = 2k$, $h = (b-a)/n$ and $x_j = a + jh$) takes the form

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

The error term for the composite Simpson's rule is

$$-\frac{1}{180}(b-a)h^4 f^{(4)}(\xi)$$

where $\xi \in (a,b)$.

♠ **Romberg Integration**. Although the trapezoid rule is the easiest Newton-Cotes formula to apply, but it lacks the degree of accuracy generally required. *Romberg integration* is a method that uses the trapezoid rule to give preliminary approximation, and then applies the Richardson's Extrapolation process to obtain improvements of the approximations.

**Note.** The second column of the Romberg table may be obtained by using the *Composite Simpson's Rule.*

**Example.** *The Romberg integration scheme*

$$R_{i,j} = \frac{4^{j-1}R_{i,j-1} - Ri - 1, j - 1}{4^{j-1} - 1}$$

*for approximating* $\int_0^\pi \sin x dx$ *with* $n = 6$ *leads to:*

| | | | | | |
|---|---|---|---|---|---|
| $R_{1,1} = 0.00000$ | | | | | |
| $R_{2,1} = 1.57079$ | $R_{2,2} = 2.09439$ | | | | |
| $R_{3,1} = 1.89611$ | $R_{3,2} = 2.00455$ | $R_{3,3} = 1.99857$ | | | |
| $R_{4,1} = 1.97423$ | $R_{4,2} = 2.00026$ | $R_{4,3} = 1.99998$ | $R_{4,4} = 2.00000$ | | |
| $R_{5,1} = 1.99357$ | $R_{5,2} = 2.000016$ | $R_{5,3} = 1.99999$ | $R_{5,4} = 2.00000$ | $R_{5,5} = 2.00000$ | |
| $R_{6,1} = 1.99839$ | $R_{6,2} = 2.000010$ | $R_{6,3} = 1.99857$ | $R_{6,5} = 2.00000$ | $R_{6,5} = 2.00000$ | $R_{6,6} = 2.00000$ |

**Romberg Integration Algorithm**

INPUT Endpoints $a, b$; integer $n$.

OUTPUT An array R ($R_{n,n}$ is the approximation to $\int_a^b f(x)dx$. Computed by rows; only 2 rows saved in storage.)

STEP1 Set $h = b - a$ and $R_{1,1} = h[f(a) + f(b)]/2$;

STEP2 OUTPUT ($R_{1,1}$).

STEP3 For $i = 2, \ldots, n$ do STEP4-8.

STEP4 (Approximation from Trapezoid Rule).

$$\text{Set} \quad R_{2,1} = \frac{1}{2}\left[R_{1,1} + h\sum_{k=1}^{2^i-2} f(a + (k - 0.5)h)\right].$$

STEP5 For $i = 2, \ldots, i$ (use Richardson's Extrapolation formula)

$$\text{Set} \quad R_{i,j} = \frac{4^{j-1}R_{i,j-1} - Ri - 1, j - 1}{4^{j-1} - 1}.$$

STEP6 OUTPUT ($R_{2,j}$ for $j = 1, 2, \ldots, i$ do STEP4-8.

STEP7 Set $h = h/2$.

STEP8 For $j = 1, 2, \ldots, i$ set $R_{1,j} = R_{2,j}$. (Update row 1 of R.)