Contents

1. Introduction

Numerical differentiation is a crucial component of numerical solution of ordinary and partial differential equations. It is also an example of a simple and visual numerical operation that allows us to illustrate some ideas such as Richardson's extrapolation used in other contexts. Finally, it is a good problem on which a trade-off between the truncation error and the roundoff can be demonstrated.

2. Finite difference approximations for derivatives

We consider some commonly used estimators for the first derivative of f at the point x:

$$D^1_+[f](x,h) := \frac{f(x+h)-f(x)}{h} \tag{1}$$

$$D^1_-[f](x,h) := \frac{f(x)-f(x-h)}{h} \tag{2}$$

$$D^1_0[f](x,h) := \frac{f(x+h)-f(x-h)}{2h} \tag{3}$$

Error estimates for these first derivative estimators are obtained by using Taylor expansions. For $D^1_+[f](x,h)$ we calculate

$$\begin{aligned} D^1_+[f](x,h) &= \frac{f(x+h)-f(x)}{h} = \frac{1}{h}\left(f(x)+hf'(x)+\frac{h^2}{2}f''(\zeta)-f(x)\right) \\ &= f'(x)+\frac{h}{2}f''(\zeta), \text{ where } \zeta\in(x,x+h) \end{aligned} \tag{4}$$

Since the error is proportional to h we say that it is a first order accurate estimator. In general, if the error estimate is proportional to h^p we say that the estimator is pth order accurate. For $D_0^1[f](x,h)$ we calculate

$$\begin{aligned} D_0^1[f](x,h) &= \frac{f(x+h)-f(x-h)}{2h} \\ &= \frac{1}{2h}\left(f(x)+hf'(x)+\frac{h^2}{2}f''(x)+\frac{h^3}{6}f'''(\zeta_1)-\left[f(x)-hf'(x)+\frac{h^2}{2}f''(x)-\frac{h^3}{6}f'''(\zeta_2)\right]\right) \\ &= f'(x)+\frac{h^2}{6}f''(\zeta), \text{ where } \zeta\in(x-h,x+h). \end{aligned} \tag{5}$$

Since the error is proportional to h^2, it is a second order accurate estimator.

Exercise Obtain an error estimate for $D_-^1[f](x,h)$.

The central difference estimator for f'' at the point x is given by

$$D_0^2[f](x,h) := \frac{f(x+h)-2f(x)+f(x-h)}{h^2}. \tag{6}$$

An error estimate for is obtained by using Taylor expansions:

$$\begin{aligned} D_0^2[f](x,h) &= \frac{f(x+h)-2f(x)+f(x-h)}{h^2} \\ &= \frac{1}{h^2}\left(2f(x)+2\frac{h^2}{2}f''(x)+\frac{h^4}{24}f^{(4)}(\zeta_1)+\frac{h^4}{24}f^{(4)}(\zeta_2)-2f(x)\right) \\ &= f''(x)+\frac{h^2}{12}f^{(4)}(\zeta) \text{ where } \zeta\in(x-h,x+h). \end{aligned} \tag{7}$$

This estimator is second order accurate.

There are several ways to obtain one-sided second-order estimators for the first derivative. One of them is the method of undetermined coefficients. We will look for a second order estimator for $f'(x)$ of the form

$$D_{2+}^1[f](x,h) = \frac{1}{h}\left(af(x)+bf(x+h)+cf(x+2h)\right), \tag{8}$$

where a, b and c are to be determined. First we write out the Taylor expansions of $f(x+h)$ and $f(x+2h)$.

$$f(x+h) = f(x)+hf'(x)+\frac{h^2}{2}f''(x)+\frac{h^3}{6}f'''(x)+O(h^4), \tag{9}$$

$$f(x+2h) = f(x)+2hf'(x)+2h^2f''(x)+\frac{4h^3}{3}f'''(x)+O(h^4) \tag{10}$$

Plugging in these expansions into Eq. (11) and grouping the terms according to the powers of h in the enumerator we get

$$h^0 : (a+b+c)f(x)$$
$$h^1 : (b+2c)f'(x)$$
$$h^2 : \left(\frac{b}{2}+2c\right)f''(x)$$
$$h^3 : \left(\frac{b}{6}+\frac{4c}{3}\right)f'''(x)$$

In order to obtain

$$D^1_{2+}[f](x,h) = \frac{af(x)+bf(x+h)+cf(x+2h)}{h} = f'(x)+O(h^2), \tag{11}$$

we set

$$a+b+c=0$$
$$b+2c=1$$
$$\frac{b}{2}+2c=0$$

Solving this system we find $c=-\frac{1}{2}$, $b=2$, $a=-\frac{3}{2}$. The error estimate coefficient is $b/6+4c/3=-1/3$. Therefore,

$$D^1_{2+}[f](x,h) = \frac{-3f(x)+4f(x+h)-f(x+2h)}{2h} = f'(x)-\frac{1}{3}h^2f''(\zeta),\ \zeta\in(x,x+2h). \tag{12}$$

In a similar manner we can obtain the other one-sided second-order accurate estimator for the first derivative:

$$D^1_{2-}[f](x,h) = \frac{3f(x)-4f(x-h)+f(x-2h)}{2h}. \tag{13}$$

3. Trade-off between the truncation and the roundoff errors

Consider the following problem. Students got a task to evaluate df/dx at $x=1$ numerically. The function f is not given analytically but can be evaluated in MATLAB with a relative error not exceeding $10\epsilon_{mach}$, where $\epsilon_{mach}\approx 1.1\cdot 10^{-16}$. Furthermore, it is known that $1<f(x)\le 2$ and $f'''(x)\le 100$ for $0<x<2$. Student A chose the formula

$$\frac{f(x+h)-f(x-h)}{2h} \quad \text{with} \quad h=1e-14. \tag{14}$$

Student B used the same formula with $h=0.01$.

(1) Whose estimate of df/dx at x = 1 do you expect to be more accurate? Provide upper bounds for the absolute errors of their results.
(2) Estimate the optimal choice of h.

Solution In the exact arithmetic we have:

$$\frac{f(x+h)-f(x-h)}{2h} = f'(x) + \frac{h^2}{6}f'''(x) + O(h^4).$$

In the floating point arithmetic, we evaluate f with a relative error:

$$\hat{f} = f(1+\eta), \quad \text{where} \quad |\eta| \le 10\epsilon_{mach}.$$

Therefore, evaluating (14) in the floating point arithmetic we get:

$$\frac{f(x+h)(1+\eta_1)-f(x-h)(1+\eta_2)}{2h} = f'(x)+\frac{h^2}{6}f'''(x)+O(h^4)+\frac{f(x+h)\eta_1 - f(x-h)\eta_2}{2h}.$$

The last term is due to roundoff. Thus, there are two components of error. The truncation error is bounded by

$$|E_{tr}| := \left|\frac{h^2}{6}f'''(x)+O(h^4)\right| \lesssim \frac{100h^2}{6},$$

while the roundoff error is bounded by

$$|E_{roff}| := \left|\frac{f(x+h)\eta_1 - f(x-h)\eta_2}{2h}\right| \lesssim \frac{4\cdot 10\epsilon_{mach}}{2h} = \frac{2.2\cdot 10^{-15}}{h}.$$

(1) For student A we have:

$$|E_{tr}| \lesssim \frac{10^{-26}}{6}, \quad |E_{roff}| \lesssim 0.22.$$

For student B:

$$|E_{tr}| \lesssim \frac{10^{-2}}{6}, \quad |E_{roff}| \lesssim 22\cdot 10^{-14}.$$

Therefore, student B made a better choice of h than student A.

(2) The optimal h is found by equating the truncation and the roundoff errors:

$$\frac{100h^2}{6} = \frac{20\epsilon_{mach}}{h} \Rightarrow h = \sqrt[3]{1.3\cdot 10^{-16}} \approx 2\cdot 10^{-5}.$$

4. Richardson's Extrapolation

Reference:

- Germund Dahlquist, Ake Bjoekr, Numerical Methods in Scientific Computing, Volume I. SIAM, 2008, see Section 3.4.6, page 302

Example We illustrate the idea of Richardson's extrapolation on the following example. We are going to design a high order estimator for the derivative $f'(x)$. We choose the finite difference estimator

$$F(h) = \frac{f(x+h)-f(x-h)}{2h}. \tag{15}$$

Using Taylor expansion we get the following series expansion for $F(h)$:

$$F(h) = \sum_{n=0}^{\infty}\frac{f^{(2n+1)}(x)}{(2n+1)!}h^{2n} = f'(x) + \frac{f'''(x)}{3!}h^2 + \frac{f^{(5)}(x)}{5!}h^4 + O(h^6).$$

Let us apply the estimator (15) using steps h, $2h$ and $4h$. We obtain:

$$F(h) = \frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{f'''(x)}{3!}h^2 + \frac{f^{(5)}(x)}{5!}h^4 + O(h^6), \tag{16}$$

$$F(2h) = \frac{f(x+2h) - f(x-2h)}{4h} = f'(x) + \frac{f'''(x)}{3!}(2h)^2 + \frac{f^{(5)}(x)}{5!}(2h)^4 + O(h^6)$$
$$= f'(x) + \frac{f'''(x)}{3!}4h^2 + \frac{f^{(5)}(x)}{5!}16h^4 + O(h^6), \tag{17}$$

$$F(4h) = \frac{f(x+4h) - f(x-4h)}{8h} = f'(x) + \frac{f'''(x)}{3!}(4h)^2 + \frac{f^{(5)}(x)}{5!}(4h)^4 + O(h^6)$$
$$= f'(x) + \frac{f'''(x)}{3!}16h^2 + \frac{f^{(5)}(x)}{5!}256h^4 + O(h^6), \tag{18}$$

Now we would like to obtain an estimator for $f'(x)$ with error proportional to h^4. Hence we need to get rid of the error term proportional to h^2. To do it, we multiply Eq. (16) by 4 and subtract Eq. (17) from it. Another estimator comes from multiplying Eq. (17) by 4 and subtracting Eq. (18) from it. As a result, we get

$$4F(h) - F(2h) = 3f'(x) - 12\frac{f^{(5)}(x)}{5!}h^4 + O(h^6), \tag{19}$$

$$4F(2h) - F(4h) = 3f'(x) - 192\frac{f^{(5)}(x)}{5!}h^4 + O(h^6). \tag{20}$$

Hence we can define the estimators with error proportional to h^4 as

$$F(h, 2h) := \tfrac{1}{3}(4F(h) - F(2h)) = F(h) + \tfrac{1}{3}(F(h) - F(2h))$$
$$= f'(x) - 4\frac{f^{(5)}(x)}{5!}h^4 + O(h^6), \tag{21}$$

$$F(2h, 4h) := \tfrac{1}{3}(4F(2h) - F(4h)) = F(2h) + \tfrac{1}{3}(F(2h) - F(4h))$$
$$= f'(x) - 64\frac{f^{(5)}(x)}{5!}h^4 + O(h^6). \tag{22}$$

In order to obtain an estimator of $f'(x)$ with error proportional to h^6 we need to get rid of the terms proportional to h^4. Therefore, we multiply Eq. (21) by 16 and subtract Eq. (22) from it. We obtain:

$$16F(h, 2h) - F(2h, 4h) = 15f'(x) + O(h^6).$$

Finally, dividing the last equation by 15 we get a 6th order accurate estimator for $f'(x)$:

$$\tfrac{1}{15}(16F(h, 2h) - F(2h, 4h)) = F(h, 2h) + \tfrac{1}{15}(F(h, 2h) - F(2h, 4h)) = f'(x) + O(h^6). \tag{23}$$

Now we generalize the description of Richardson's extrapolation. Suppose we would like to estimate a certain quantity. Its estimator $F(h)$ depends on a small parameter h. The error has a series expansion in h. The exact quantity would be $F(0)$. For example,

$$f'(x) \approx F(h) := \frac{f(x+h) - f(x-h)}{2h},$$

$$f''(x) \approx F(h) := \frac{f(x+h) - 2f(x) + f(x-h)}{h^2},$$

$$\int_a^b f(x)dx \approx F(h) := T_h(f) = \tfrac{1}{2}h[f(a) + f(b)] + h\sum_{j=1}^{n-1} f_j,$$

In practice, we cannot set h very small. there are two reasons for it. First, the effects of the roundoff error set a practical bound of how small h can be. Second, the CPU time typically increases sharply as $h \to 0$.

However, if we know that the truncation error $F(h) - F(0)$ has an asymptotic error expansion in powers of h, i.e.,

$$F(h) = F(0) + a_1 h^p + O(h^r), \quad \text{as} \quad h \to 0,\ r > p,$$

we can design more accurate estimator. We define $a_0 := F(0)$. We compute $F(h)$ and $F(qh)$, $(q > 1)$:

$$F(h) = a_0 + a_1 h^p + O(h^r),$$
$$F(qh) = a_0 + a_1 (qh)^p + O(h^r).$$

Then multiplying the first equation by q^p and subtracting the second one from it we get

$$(q^p - 1)a_0 = q^p F(h) - F(qh) + O(h^r).$$

Hence

$$a_0 = F(0) = F(h) + \frac{F(h) - F(qh)}{q^p - 1} + O(h^r).$$

This process is called **Richardson extrapolation**, or the deferred approach to the limit.

Suppose that the form of the expansion of $F(h)$ in powers of h is known. Then one can, even if the values of the coefficients in the expansion are unknown, repeat the use of Richardson extrapolation in a way described below. This process is, in many numerical problems – especially in numerical treatment of integrals and differential equations – the simplest way to get results which have negligible truncation error. The application of this process becomes especially simple when the step lengths form a geometric series:

$$h,\ qh,\ q^2h,\ \ldots.$$

Theorem 1. *Suppose that*

$$F(h) = a_0 + a_1 h^{p_1} + a_2 h^{p_2} + a_3 h^{p_3} + \ldots, \tag{24}$$

where $p_1 < p_2 < p_3 < \ldots$, *and set*

$$F_1(h) = F(h), \quad F_{k+1}(h) = F_k(h) + \frac{F_k(h) - F_k(qh)}{q^{p_k} - 1}. \tag{25}$$

Then $F_n(h)$ has an expansion of the form

$$F_n(h) = a_0 + a_n^{(n)} h^{p_n} + a_{n+1}^{(n)} h^{p_{n+1}} + \dots. \tag{26}$$

Proof. The proof is by induction. Eq. (26) holds for $n = 1$. Suppose Eq. (26) holds for $n = k$, i.e.,

$$F_k(h) = a_0 + a_k^{(k)} h^{p_k} + a_{k+1}^{(k)} h^{p_{k+1}} + \dots.$$

Then

$$F_k(qh) = a_0 + a_k^{(k)} (qh)^{p_k} + a_{k+1}^{(k)} (qh)^{p_{k+1}} + \dots.$$

Hence

$$q^{p_k} F_k(h) - F_k(qh) = a_0(q^{p_k} - 1) + b h^{p_{k+1}} + \dots.$$

Therefore, dividing this equation by $q^{p_k} - 1$ we get

$$F_{k+1}(h) := F_k(h) + \frac{F_k(h) - F_k(qh)}{q^{p_k} - 1} = a_0 + a_{k+1}^{(k+1)} (qh)^{p_{k+1}} + \dots$$

that is of desired form. □

Example Let us evaluate the derivative of $f(x) = \log x$ at $x = 3$. We choose the finite difference estimate

$$F(h) = \frac{f(x+h) - f(x-h)}{2h}.$$

Using Taylor expansion we get the following series expansion for $F(h)$:

$$F(h) = \sum_{n=0}^{\infty} \frac{f^{(2n+1)}(x)}{(2n+1)!} h^{2n} = f'(x) + \frac{f'''(x)}{3!} h^2 + \frac{f^{(5)}(x)}{5!} + \dots.$$

We pick $q = 2$ and steps 0.8, 0.4 and 0.2 and write

$$F_{00} = \frac{\log(3.8) - \log(2.2)}{1.6} = a_0 + a_1 \cdot (2^2)^2 h^2 + a_2 \cdot (2^2)^4 h^4 + \dots = 0.341589816480044,$$

$$F_{10} = \frac{\log(3.4) - \log(2.6)}{0.8} = a_0 + a_1 \cdot 2^2 h^2 + a_2 \cdot 2^4 h^4 + \dots = 0.335329983243349,$$

$$F_{20} = \frac{\log(3.2) - \log(2.8)}{0.4} = a_0 + a_1 \cdot h^2 + a_2 \cdot h^4 + \dots = 0.333828481561307.$$

Then we calculate taking into account that $p_1 = 2$, $p_2 = 4$, ...,

$$F_{11} = F_{10} + \frac{F_{10} - F_{00}}{2^{p_1} - 1} = \tfrac{4}{3} F_{10} - \tfrac{1}{3} F_{00} = 0.333243372164451,$$

$$F_{21} = F_{20} + \frac{F_{20} - F_{10}}{2^{p_1} - 1} = \tfrac{4}{3} F_{20} - \tfrac{1}{3} F_{10} = 0.333327981000626,$$

$$F_{22} = F_{21} + \frac{F_{21} - F_{11}}{2^{p_2} - 1} = \tfrac{16}{15} F_{21} - \tfrac{1}{15} F_{11} = 0.333333621589704.$$

The absolute error of the result is $2.8e-7$. For comparison, if we take $h = 0.005$, the error of of the estimate is $3.1e-7$.

Example Let us design a one-sided second-order estimator for $f'(x)$. We set

$$F_{10} := \frac{f(x+h)-f(x)}{h} = f'(x) + \tfrac{1}{2}f''(x)h + \tfrac{1}{6}f'''(x)h^2 + O(h^3), \tag{27}$$

$$F_{00} := \frac{f(x+2h)-f(x)}{2h} = f'(x) + \tfrac{1}{2}f''(x)2h + \tfrac{1}{6}f'''(x)4h^2 + O(h^3). \tag{28}$$

Multiplying Eq. (27) by 2 and subtracting Eq. (28) from it we get the desired estimator

$$F_{11} := 2F_{10} - F_{00} = f'(x) - \tfrac{1}{3}f'''(x)h^2 + O(h^3). \tag{29}$$

Now we can express F_{11} via $f(x)$, $f(x+h)$, and $f(x+2h)$:

$$\begin{aligned} F_{11} : &= 2F_{10} - F_{00} = 2\frac{f(x+h)-f(x)}{h} - \frac{f(x+2h)-f(x)}{2h} \\ &= \frac{-3f(x)+4f(x+h)-f(x+2h)}{2h} = f'(x) - \tfrac{1}{3}f'''(x)h^2 + O(h^3). \end{aligned} \tag{30}$$

5. Differentiation matrices

Suppose we have a uniform grid x_0, x_1, ..., x_n, and we need to obtain a derivative of $f(x)$ at the interior grid points. We will denote the values of f at the grid points by f_j's: $f_j \equiv f(x_j)$, $j = 0, \dots, n$. Using the central difference (3), we write:

$$f'(x_j) \approx \frac{f_{j+1}-f_{j-1}}{2h}, \quad \text{where} \quad h = \frac{x_n - x_0}{n}, \quad j = 1, \dots, n-1. \tag{31}$$

Equation (31) can be written in matrix form:

$$\begin{bmatrix} f_1' \\ f_2' \\ \vdots \\ f_{n-2} \\ f_{n-1}' \end{bmatrix} = \frac{1}{h}\begin{bmatrix} 0 & 1/2 & & & \\ -1/2 & 0 & 1/2 & & \\ & \ddots & \ddots & \ddots & \\ & & -1/2 & 0 & 1/2 \\ & & & -1/2 & 0 \end{bmatrix}\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-2} \\ f_{n-1} \end{bmatrix} + \frac{1}{h}\begin{bmatrix} -f_0/2 \\ 0 \\ \vdots \\ 0 \\ f_n/2 \end{bmatrix}. \tag{32}$$

The matrix in this equation is often referred to as the *differentiation matrix*, and the last term is the *boundary term*. This differentiation matrix os tridiagonal. It is customarily to assume that all entries left blank in a matrix are zeros.

If we would have a periodic function f then $f_0 = f_n$, we would have the differentiation matrix $n \times n$ with extra nonzero entries at the top right and bottom left corners:

$$\begin{bmatrix} f_0' \\ f_1' \\ \vdots \\ f_{n-2} \\ f_{n-1}' \end{bmatrix} = \frac{1}{h}\begin{bmatrix} 0 & 1/2 & & & -1/2 \\ -1/2 & 0 & 1/2 & & \\ & \ddots & \ddots & \ddots & \\ & & -1/2 & 0 & 1/2 \\ 1/2 & & & -1/2 & 0 \end{bmatrix}\begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{n-2} \\ f_{n-1} \end{bmatrix}. \tag{33}$$

6. A connection between numerical differentiation and interpolation

Reference: L. N. Trefethen, Spectral Methods in Matlab. Read Chapter 1 "Differentiation matrices", pages 1–8.

Now we will show that the entries in the differentiation matrices can be obtained using Lagrange's interpolation. Let us fix j, $1 \le j \le n-1$, and consider Lagrange's interpolating polynomial through (x_{j-1}, f_{j-1}), (x_j, f_j), and (x_{j+1}, f_{j+1}). This polynomial is given by

$$p(x) = f_{j-1}\frac{x - x_j)(x - x_{j+1})}{2h^2} + f_j\frac{(x - x_{j-1})(x - x_{j+1})}{-h^2} + f_{j+1}\frac{(x - x_{j-1})(x - x_j)}{2h^2}. \tag{34}$$

Now we will differentiate p with respect to x and evaluate the derivative at $x = x_j$:

$$p'(x) = f_{j-1}\frac{2x - x_j - x_{j+1}}{2h^2} + f_j\frac{2x - x_{j-1} - x_{j+1}}{-h^2} + f_{j+1}\frac{2x - x_j - x_{j-1}}{2h^2}, \tag{35}$$

$$p'(x_j) = f_{j-1}\frac{-h}{2h^2} + f_{j+1}\frac{h}{2h^2} = \frac{f_{j+1} - f_{j-1}}{2h}. \tag{36}$$

We see that the right-hand side of (36) coincides with that of (31). Let us show that differentiating a three-point polynomial interpolant with equispaced points and evaluating its derivative at the central point gives a second order accurate approximation for the derivative of the function f. Indeed, let us fix some x and write:

$$f(x) = p(x) + f[x_{j-1}, x_j, x_{j+1}, x]\pi(x), \tag{37}$$

where $p(x)$ is given by (34), $f[x_{j-1}, x_j, x_{j+1}, x]$ is a divided difference, and $\pi(x)$ is the nodal polynomial

$$\pi(x) = (x - x_{j-1})(x - x_j)(x - x_{j+1}).$$