The errors in numerical differentiation

In practice, we always get two kinds of errors: the **theoretical error** as well as a **rounding error** (due to the fact that we have to round all involved quantities to, say, double precision). In the past, we have been able to mostly ignore the rounding error. However, we cannot afford to do so in the case of numerical differentiation. The reason for the trouble is that our finite difference approximations always subtract nearly equal quantities (that's in the nature of differentiation!) which can lead to a devastating loss of precision.

Comment. The theoretical error is often also called **truncation error**. Here, truncation is meant in the sense of, for instance, having a function $f(x)$ and truncating its Taylor series to get an approximation of $f(x)$.

Example 119. Analyze the overall error in using the approximation $f'(x) \approx \frac{1}{h}[f(x+h) - f(x)]$.

Solution. As we worked out in Example 111, the theoretical error is $\frac{1}{h}[f(x+h) - f(x)] = f'(x) + \frac{h}{2}f''(\xi)$.

In practice, all quantities including $f(x+h)$ and $f(x)$ are slightly rounded and only accurate to within some ε (the machine precision). In the sequel, we assume double precision, in which case $\varepsilon \approx 2^{-52} \approx 2.2 \cdot 10^{-16}$. Our approximation therefore is (ε_1 and ε_2 are the rounding errors for $f(x+h)$ and $f(x)$ respectively) roughly:

$$
\begin{aligned}
\frac{1}{h}[(f(x+h) + \varepsilon_1) - (f(x) + \varepsilon_2)] &= \frac{1}{h}[f(x) - f(x-h)] + \frac{\varepsilon_1 - \varepsilon_2}{h} \\
&= f'(x) + \underbrace{\frac{1}{2}f''(\xi)h}_{\text{theoretical error}} + \underbrace{\frac{\varepsilon_1 - \varepsilon_2}{h}}_{\text{rounding error}}
\end{aligned}
$$

Writing $M = \frac{1}{2}|f''(\xi)|$ (note that $M \approx \frac{1}{2}|f''(x)|$ if h is small) and observing that $\left|\frac{\varepsilon_1 - \varepsilon_2}{h}\right| \leqslant \frac{2\varepsilon}{h}$, we see that the overall error is bounded by

$$
E(h) = Mh + \frac{2\varepsilon}{h}.
$$

Plot the function $E(h)$ to see that this bound becomes bad as h gets too small (that's because of the rounding error $2\varepsilon / h$). In particular, we can see that there will be a "best" choice for h which minimizes this bound for the error. To find this best value $h^\star$, we compute $E'(h) = M - \frac{2\varepsilon}{h^2} = 0$ and solve for h. We find

$$
h^\star = \sqrt{\frac{2\varepsilon}{M}} \approx M^{-1/2} \cdot 2.1 \cdot 10^{-8}.
$$

The corresponding bound for the error is $E(h^\star) \approx Mh^\star + \frac{2\varepsilon}{h^\star} \approx M^{1/2} \cdot 4.2 \cdot 10^{-8}$.

Comment. Note that our estimates match the values we observed in Example 118 fairly well. In that example, we had $f(x) = 2^x$ so that $f''(x) = (\ln 2)^2 \, 2^x$ and, therefore, $M = \frac{1}{2}|f''(\xi)| \approx \frac{1}{2}|f''(1)| = (\ln 2)^2$. In particular, $M^{1/2} \approx \ln 2 \approx 0.6931$ and $E(h^\star) \approx M^{1/2} \cdot 4.2 \cdot 10^{-8} \approx 2.9 \cdot 10^{-8}$.

Example 120. Analyze the overall error in approximating $f'(x) \approx \frac{1}{2h}[f(x+h) - f(x-h)]$.

Solution. The theoretical error is $\frac{1}{2h}[f(x+h) - f(x-h)] = f'(x) + \frac{h^2}{6}f'''(x) + O(h^3)$ (see Example 112).

In practice, proceeding as in the previous example, our approximation is roughly:

$$\frac{1}{2h}[(f(x+h) + \varepsilon_1) - (f(x-h) + \varepsilon_2)] = \frac{1}{2h}[f(x+h) - f(x-h)] + \frac{\varepsilon_1 - \varepsilon_2}{2h}$$
$$= f'(x) + \underbrace{\frac{1}{6}f'''(\xi)h^2}_{\text{theoretical error}} + \underbrace{\frac{\varepsilon_1 - \varepsilon_2}{2h}}_{\text{rounding error}}$$

Note that $\left|\frac{\varepsilon_1 - \varepsilon_2}{2h}\right| \leqslant \frac{\varepsilon}{h}$. Writing $M = \frac{1}{6}|f'''(\xi)|$, the overall error is bounded by

$$E(h) = Mh^2 + \frac{\varepsilon}{h}.$$

Again, plotting the function $E(h)$ we see that this bound becomes bad as h gets too small and that there is a "best" value h^* which minimizes this bound for the error. We compute $E'(h) = 2Mh - \frac{\varepsilon}{h^2} = 0$ and solve for h. We find

$$h^* = \sqrt[3]{\frac{\varepsilon}{2M}} \approx M^{-1/3} \cdot 4.8 \cdot 10^{-6}.$$

The corresponding bound for the error is $E(h^*) \approx M(h^*)^2 + \frac{\varepsilon}{h^*} \approx M^{1/3} \cdot 6.9 \cdot 10^{-11}$.

Comment. Again, our estimates match the values we observed in Example 118 rather well.

Richardson extrapolation

Suppose that $A(h)$ is an approximation of a quantity A^* of order n. Our goal is to construct an approximation of A^* of higher order (by combining approximations $A(h)$ for different h).

If $A(h)$ approximates A^* to order n, then we often have $A(h) = A^* + Ch^n + O(h^{n+1})$ for some constant C. This is true, for instance, for all of our numerical differentiation examples.

In that case, we have $A(2h) = A^* + C(2h)^n + O(h^{n+1})$ as well as $2^n A(h) = 2^n A^* + C(2h)^n + O(h^{n+1})$.

Thus the difference is $2^n A(h) - A(2h) = (2^n - 1)A^* + O(h^{n+1})$.

Dividing both sides by $2^n - 1$, we therefore get an approximation of order higher than n.

(Richardson extrapolation) Starting with an approximation $A(h)$ of a quantity A^* of order n, its **Richardson extrapolation** is the approximation

$$R(h) := \frac{2^n A(h) - A(2h)}{2^n - 1}.$$

As we showed above, it typically is an approximation of higher order.

More generally, we get a Richardson extrapolation $R_\lambda(h) = \frac{\lambda^n A(h) - A(\lambda h)}{\lambda^n - 1}$ for any choice of $\lambda > 0$ (the choice $\lambda = 2$ is common but not mathematically special).

Comment. Note that, based on the values at h and $2h$, we are trying to get our hands on A^* which is the value at 0. Because 0 is outside of the interval $[h, 2h]$, this is an extrapolation.

Example 121. Suppose that $A\left(\frac{1}{2}\right) = \frac{3}{8}$ and $A\left(\frac{1}{3}\right) = \frac{5}{12}$ are approximations of order 3 of some quantity A^*. What is the approximation we obtain from using Richardson extrapolation?

Solution. Since $A(h)$ is an approximation of order 3, we expect $A(h) \approx A^* + Ch^3$ for some constant C.

Correspondingly, $A\left(\frac{1}{2}\right) \approx A^* + \frac{1}{8}C$ and $A\left(\frac{1}{3}\right) \approx A^* + \frac{1}{27}C$.

Hence, $27A\left(\frac{1}{3}\right) - 8A\left(\frac{1}{2}\right) \approx (27-8)A^* = 19A^*$. To get an approximation of A^*, we need to divide by 19.

The Richardson extrapolation is $\frac{27}{19}A\left(\frac{1}{3}\right) - \frac{8}{19}A\left(\frac{1}{2}\right) = \frac{27}{19}\frac{5}{12} - \frac{8}{19}\frac{3}{8} = \frac{33}{76}$.

Comment. Note that we did is equivalent to using the formula $R_\lambda(h) = \frac{\lambda^n A(h) - A(\lambda h)}{\lambda^n - 1}$ with $h = \frac{1}{3}$ and $\lambda = \frac{3}{2}$.

Comment. The numbers above are not random. Instead, $A(h) = \frac{1}{2} - \frac{h}{4}$ which is an order 3 approximation of $f(0) = \frac{1}{2}$ for $f(x) = \frac{1}{1+e^x} = \frac{1}{2} - \frac{x}{4} + \frac{x^3}{48} + O(x^5)$. Indeed, $\frac{33}{76} \approx 0.434$ is slightly better than $A\left(\frac{1}{2}\right)$ and $A\left(\frac{1}{3}\right)$.

Example 122. Apply Richardson extrapolation to $f'(x) \approx \frac{1}{h}[f(x+h) - f(x)]$.

Solution. We have seen in Example 111 that the approximation $A(h) = \frac{1}{h}[f(x+h) - f(x)]$ is of order 1. Hence,

$$
\begin{aligned}
\frac{2A(h) - A(2h)}{2-1} &= \frac{2}{h}[f(x+h) - f(x)] - \frac{1}{2h}[f(x+2h) - f(x)] \\
&= \frac{1}{2h}[-f(x+2h) + 4f(x+h) - 3f(x)]
\end{aligned}
$$

is an approximation of higher order (we expect it to be of order 2).

Indeed, this approximation of $f'(x)$ is the same as what we obtained in Example 116 when applying polynomial interpolation to $f(x)$, $f(x+h)$, $f(x+2h)$. As observed there, the error is $-\frac{1}{3}f'''(x)h^2 + O(h^3)$ showing that this is indeed an approximation of order 2.

Example 123. Apply Richardson extrapolation to $f'(x) \approx \frac{1}{2h}[f(x+h) - f(x-h)]$.

Solution. We have seen in Example 112 that the approximation $A(h) = \frac{1}{2h}[f(x+h) - f(x-h)]$ is of order 2. Hence,

$$
\begin{aligned}
\frac{2^2 A(h) - A(2h)}{2^2 - 1} &= \frac{2}{3h}[f(x+h) - f(x-h)] - \frac{1}{12h}[f(x+2h) - f(x-2h)] \\
&= \frac{1}{12h}[-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)]
\end{aligned}
$$

is an approximation of $f'(x)$ of higher order. With some more work (do it!), we find that the error is $-\frac{1}{30}f^{(5)}(x)h^4 + O(h^6)$ so that this is an approximation of order 4.

Comment. Note that the above approximations don't change when h is replaced by $-h$ (note the factor of $1/h$). In other words, the approximations are even functions in h. Consequently, their Taylor expansion in h will only have even powers of h. That's the reason why the order of the Richardson extrapolation is 4 rather than order 3 which is what one would otherwise expect when extrapolating an order 2 formula.

Example 124. Apply Richardson extrapolation to $f''(x) \approx \frac{1}{h^2}[f(x+h) - 2f(x) + f(x-h)]$.

Solution. We have seen in Example 115 that the approximation $A(h) = \frac{1}{h^2}[f(x+h) - 2f(x) + f(x-h)]$ is of order 2. Hence,

$$
\frac{2^2 A(h) - A(2h)}{2^2 - 1} = \frac{1}{12h^2}[-f(x+2h) + 16f(x+h) - 30f(x) + 16f(x-h) - f(x-2h)]
$$

is an approximation of $f''(x)$ of higher order. With some more work, we find that the error is $-\frac{1}{90}f^{(6)}(x)h^4 + O(h^6)$ so that this is an approximation of order 4.

In the previous example, we combined $A(h)$ and $A(2h)$ to obtain an approximation of higher order. There is nothing special about $2h$. We can likewise combine $A(h_1)$ and $A(h_2)$ for any h_1, h_2.

Example 125. (homework) $A(h) = \frac{1}{h^2}[f(x+h) - 2f(x) + f(x-h)]$ is an order 2 approximation of $f''(x)$. Apply Richardson extrapolation to $A(h)$ and $A\left(\frac{3}{2}h\right)$ to obtain an approximation of $f''(x)$ of higher order.

Solution. Since $A(h)$ is an approximation of order 2, we expect $A(h) \approx A^* + Ch^2$ for some constant C. Correspondingly, $A\left(\frac{3}{2}h\right) \approx A^* + \frac{9}{4}Ch^2$. Hence, $\frac{9}{4}A(h) - A\left(\frac{3}{2}h\right) \approx \left(\frac{9}{4} - 1\right)A^* = \frac{5}{4}A^*$.

The Richardson extrapolation of $A(h)$ and $A\left(\frac{3}{2}h\right)$ therefore is:

$$\frac{\frac{9}{4}A(h) - A\left(\frac{3}{2}h\right)}{\frac{5}{4}} = \frac{1}{45h^2}\left[-16f\left(x + \frac{3}{2}h\right) + 81f(x+h) - 130f(x) + 81f(x-h) - 16f\left(x - \frac{3}{2}h\right)\right]$$

This is an approximation of $f''(x)$ of higher order. With some more work, we find that the error is $-\frac{1}{160}f^{(6)}(x)h^4 + O(h^6)$ so that this is an approximation of order 4.

Example 126. `Python` Let us now repeat Example 118 with the formula

$$f'(x) \approx \frac{1}{12h}[-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)]$$

that we obtained in Example 123.

```
>>> def central_difference_richardson(f, x, h):
        return (-f(x+2*h)+8*f(x+h)-8*f(x-h)+f(x-2*h))/(12*h)
```

Let us again approximate $f'(1) = 2\ln(2) \approx 1.386$ for $f(x) = 2^x$ at $x = 1$.

```
>>> [central_difference_richardson(f, 1, 10**-n) for n in range(5)]
```

```
[1.375, 1.3862932938249581, 1.3862943610132332, 1.3862943611198109, 1.3862943611187004]
```

Does the error behave as expected?

```
>>> [central_difference_richardson(f, 1, 10**-n) - 2*log(2) for n in range(6)]
```

```
[-0.011294361119890572, -1.072949324330716e-06, -1.0665734961889939e-10, -
7.971401316808624e-14, -1.1901590823981678e-12, 1.5093037930569153e-11]
```

We noted in Example 123 that the approximation is of order 4. Indeed, we can see how the error decreases roughly by $1/10^4$ initially, as expected.

Moreover, we are able to obtain a much better numerical estimate compared to Example 118: this time, our best approximation has error $7.97 \cdot 10^{-14}$, which is decently close to the machine precision of $\varepsilon \approx 2^{-52} \approx 2.2 \cdot 10^{-16}$. This is because the effect of rounding errors becomes devastating as h becomes very small. Using a high-order approximation, we are often able to avoid having to work with very small h.

Indeed, note how we got the best approximation with $h = 10^{-3}$ (whereas we previously needed to a much smaller h for the best approximations).