Interesting Numerical Differentiation Tidbits

Skip Thompson
Department of Mathematics & Statistics
Radford University
Radford, VA 24142
thompson@radford.edu

Solutions and Comments

1. `Total Error Equation.` Taylor's Theorem shows that

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

for some α in the interval $(x, x+h)$. Roughly speaking, the mathematical error is thus bounded by $A\,h$ where A is any bound for $\frac{|f''|}{2}$ on the interval $(x, x+h)$. (In our example $A = \frac{1}{2}$.) Let's see what happens when we compute numerically. Denote by $f'_c(x)$ the computed finite difference approximation and by E_1 an E_2 the roundoff errors in the computed values of $f(x+h)$ and $f(x)$. Then

$$f'_c(x) = \frac{f(x+h) + E_1 - f(x) - E_2}{h} \tag{1}$$

$$= \frac{f(x+h) - f(x)}{h} + \frac{E_1 - E_2}{h}. \tag{2}$$

E_1 and E_2 are bounded by a few multiples of unit roundoff where unit roundoff is the smallest positive number ϵ for which $1 + \epsilon = 1$ (approximately 10^{-16} for our example). Roughly speaking, the roundoff error is thus bounded by a term $\frac{B}{h}$ where $B = O(\epsilon)$. The total error then roughly has the form

$$E(h) = A\,h + \frac{B}{h}.$$

Visualize what the graph of this equation looks like as a function of h. When $h \sim 0$, the hyperbolic term dominates the error. When $h >> 0$, the linear term dominates the error. The total error is minimized when $E'(x) = 0$, that is, when $h = \sqrt{\frac{B}{A}}$. For this "optimal" value of h, the total error is $2\sqrt{A\,B}$. To summarize: below the optimal value of h the error is dominated by roundoff error; for values larger than the optimal h the error is dominated by the mathematical truncation error of the difference quotient used to approximate $f'(x)$. We note that similar analyses can be used to analyze the total error for other derivative approximations (for example, second order accurate centered differences). The interested reader will want to study the optimal values of h for other common approximations.

2. `Estimation of the Constants` A `and` B. Of course, we don't know the values of A and B. We can estimate them using a least squares regression. We wish to determine the coefficients A and B that minimize the quantity

$$r(A, B) = \sum_{i=1}^{n} \left(Ah_i + \frac{B}{h_i} - e_i \right)^2 .$$

Calculating $\dfrac{\partial r}{\partial A}$ and $\dfrac{\partial r}{\partial B}$ and setting them to 0 leads to the following two linear equations.

$$\sum_{i=1}^{n} \left(Ah_i + \frac{B}{h_i} \right) h_i = 0 \tag{3}$$

$$\sum_{i=1}^{n} \left(Ah_i + \frac{B}{h_i} \right) \frac{1}{h_i} = 0. \tag{4}$$

The determinant of this system is equal to

$$D = \left(\sum_{i=1}^{n} h_i^2 \right) \left(\sum_{i=1}^{n} \frac{1}{h_i^2} \right) - n^2.$$

When the two sums are expanded and multiplied, we find that

$$\left(\sum_{i=1}^{n} h_i^2 \right) \left(\sum_{i=1}^{n} \frac{1}{h_i^2} \right) > n + 2 \left(\frac{n^2 - n}{2} \right) = n^2.$$

The first term reflects the fact that n of the terms are $\dfrac{h_i^2}{h_i^2} = 1$. The remaining $n^2 - n$ terms can be grouped in pairs of the form $\dfrac{h_i^2}{h_j^2} + \dfrac{h_j^2}{h_i^2}$ where $h_i \neq h_j$. Each such pair has the form $x + \dfrac{1}{x}$ where $x = \dfrac{h_i^2}{h_j^2}$. The minimum of such a function is 2 obtained for $x = 1$; but $x = 1$ is not permissable since $h_i \neq h_j$. This shows that $D > 0$.