PHYS 416: Truncation and Rounding Errors

Definition: *Truncation Error* – the difference between the true result and the approximation if exact arithmetic is used.

Example: Finite difference approximation to a derivative of a function f(x). Using a forward difference approximation, the derivative can be approximated as:

$$f'(x) \approx \frac{f(x+h)-f(x)}{h} \tag{1}$$

where h is some small interval. The estimate between the approximation and the exact value can be obtained from Taylor's theorem

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(\theta) \tag{2}$$

where $x \le \theta \le x+h$. If we let $M = |f''(\theta)|$, then (1) becomes

$$\frac{f(x+h)-f(x)}{h} = f'(x) + \frac{h}{2} M \tag{3}$$

so that the truncation error is of order

$$M\frac{h}{2} \tag{4}$$

Definition: *Roundoff Error* – the difference between the true result and approximation if exact arithmetic is with that using finite precision arithmetic.

This is basically the error that results from the inexactness in representing real numbers. From equation (1) this error is approximately

$$\frac{2\varepsilon}{h} \tag{5}$$

In MATLAB, ε is the constant `eps` and is of the order 10^{-16}. The total error (E) is the sum of equations (4) and (5) as

$$E \approx M\frac{h}{2} + \frac{2\varepsilon}{h} \tag{6}$$

This error does not go to 0 as h approaches zero, its smallest value is when

$$h_{\min} \approx 2\sqrt{\varepsilon / M} \tag{7}$$

where the error becomes $E_{\min} = 2\sqrt{\varepsilon}$. The point of this is that a smaller stepsize does not always lead to a smaller error, as roundoff error can become dominant for small enough stepsizes. Figure 1 shows the estimated (from equation (6)) and the actual computed error for the function $f(x) = x^3$ ad x =1, from equation (7), $h_{\min} \approx 10^{-8}$, and $E_{\min} = 10^{-8}$, in the case M was assumed to be 1.

A more accurate truncation error can be obtained using central differences

$$f'(x) \approx \frac{f(x+h)-f(x-h)}{2h} \tag{8}$$

in which case the error becomes

$$\frac{f(x+h)-f(x-h)}{2h} = f'(x) + \frac{h^2}{12} M \tag{9}$$

where $M = |f'''(\theta)|$. The error for this central differences becomes

$$E \approx M\frac{h^2}{6} + \frac{\varepsilon}{h} \tag{10}$$

and

$$h_{\min} \approx \left(3\varepsilon/M\right)^{1/3} \tag{11}$$

and $E_{\min} = \left(3M\varepsilon^2\right)^{1/3}$

Figure 2 shows the resulring plot for equation (10) for $f(x) = x^3$ at x =1.

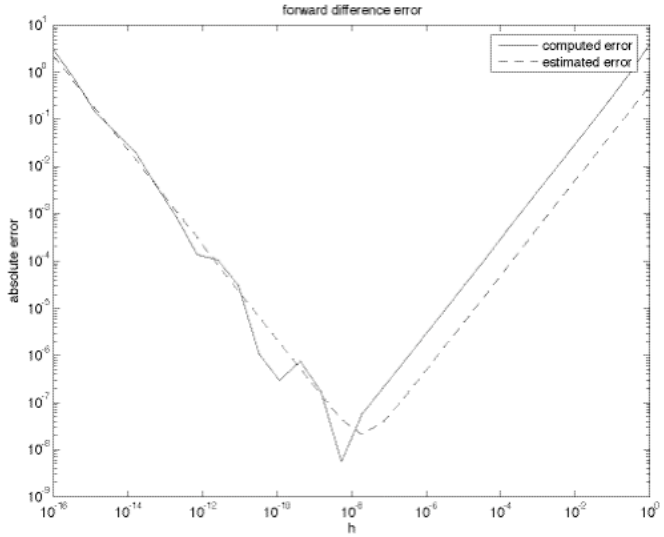

Figure 1: Computed error for $f(x) = x^3$ versus estimated error from equation (7). The differences between the 2 curves for h > 10-8 is due to errors in the value of M.

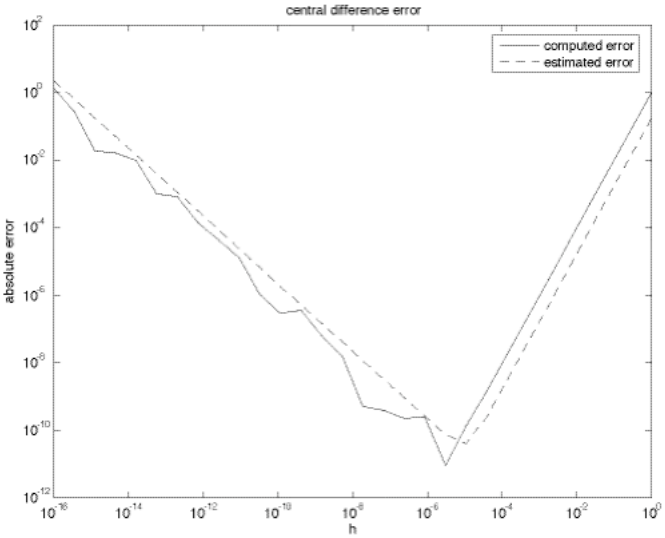

Figure 2: Computed error for $f(x) = x^3$ versus estimated error from equation(10). The differences between the 2 curves for h > 10-8 is due to errors in the value of M.