Homework 1

due January 28 2016, 12:00

Task 1: Machine Epsilon

The following code can be used in MATLAB to determine the machine accuracy ε.

```
numprec=double(1.0); % Define 1.0 with double precision
numprec=single(1.0); % Define 1.0 with single precision
while(1 < 1 + numprec)
    numprec=numprec*0.5;
end
numprec=numprec*2
```

a) Determine ε using the above program, both for single and double precision. Note that a double precision number uses 8 bytes of storage, whereas a single precision number only occupies 4 bytes.

   Note: The implementation of single/double precision arithmetics differs between versions of MATLAB. If runs with both single and double precision give the same answer, please try another computer/version of MATLAB if possible. Otherwise, write down your MATLAB version and move on. The above code is working properly on release 2009a to 2012b on Linux, for instance.

b) Give a definition of the machine accuracy based on the code above. Try to use words and not mathematical expressions.

Task 2: Round-off Error

In this exercise, the errors involved in numerically calculating derivatives are examined. The derivative of a function $f(x)$ can be approximated with central differences:

$$f'_{\text{num}}(x) = \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x} \tag{1}$$

a) Compute, numerically, the relative discretisation error of the derivative of the function $f(x) = \dfrac{1}{1+x} + x$ when using the central difference scheme defined above. The relative discretisation error is given by:

$$\xi_d = \frac{|f'(x) - f'_{\text{num}}(x)|}{|f'(x)|}, \tag{2}$$

   where $f'(x)$ is the analytical derivative of $f(x)$. Compute ξ_d for $x = 2$ and use stepsizes $\Delta x = 10^{-20}, \dots, 10^{0}$. Use both single and double precision for the calculation and present the results in a double logarithmic plot (ξ_d vs. Δx). In MATLAB double logarithmic plots are obtained by the function `loglog()`. Remember that *all* variables used here should be defined as double or single precision as in Task 1.

b) The general formula for the propagation error, for a function $h(X_j)$ with n variables X_j is given by:

$$\xi_h = \sum_{j=1}^{n} \left| \frac{X_j}{h} \frac{\partial h}{\partial X_j} \right| \varepsilon_{X_j}, \tag{3}$$

where ε_{X_j} is the accuracy on the quantity X_j. Based on this, show that the propagation error, ξ_h, of the sum of two numbers X_1 and X_2 is given by:

$$\xi_h = \frac{|X_1|}{|X_1 + X_2|}\varepsilon_{X_1} + \frac{|X_2|}{|X_1 + X_2|}\varepsilon_{X_2}$$

where ε_{X_1} and ε_{X_2} are the corresponding accuracies of each number.

c) Show that, when using the proposed central differences scheme, the relative discretisation error (equation (2)) is given by:

$$\xi_d = \frac{\Delta x^2 |f'''(x)|}{6|f'(x)|}$$

(Hint: Taylor expansion)
and that the propagation error (equation (3)) is given by:

$$\xi_h = \frac{|f(x)|\varepsilon}{|f'(x)|\Delta x}$$

where ε is the machine accuracy. Find, analytically, the value of Δx that minimises the total error

$$\xi_g = \xi_d + \xi_h.$$

Plot ξ_d, ξ_h and ξ_g together with the results from part a).

Task 3 : Integration of differential equation

In this problem the stability and convergence order of three integration methods is examined. The first order, ordinary, linear differential equation with constant coefficient is considered (Dahlquist equation)

$$\frac{\mathrm{d}u}{\mathrm{d}t} = A(u) = \lambda u, \qquad u(0) = 1$$

where $0 \le t \le T$ and $\lambda = \lambda_{\Re} + i\lambda_{\Im} = \text{const} \in \mathbb{C}$. The time interval $[0, T]$ is split into N parts with the same length Δt. The following integration methods should be used:

- explicit Euler
$$u^{n+1} - u^n = \Delta t A(u^n)$$
- implicit Euler
$$u^{n+1} - u^n = \Delta t A(u^{n+1})$$
- Crank-Nicolson
$$u^{n+1} - u^n = \frac{1}{2}\Delta t(A(u^{n+1}) + A(u^n))$$

where $n = 0, \ldots, N$. Compute the solution until $T = 10$ and use $N = 20, 40, 50, 100, 200$ steps.

a) Derive the analytical solution u_{ex}.

b) For $\lambda = -\sqrt{3}/2 + \pi i$, calculate the numerical solution with the given discretisations N and the three integration methods. Plot the real part of the analytical solution and the three numerical solutions for each value of N.

c) As $\lambda_{\Re} \to -\infty$ the problem becomes more and more stiff. Derive, for the three considered numerical schemes, the expression of the amplification factor, $G(z)$, where $z = \lambda \Delta t \in \mathbb{C}$. Compute the limit of $G(z)$ for $z_{\Re} \to -\infty$ and plot the real part of $G(z)$ against the real part of z for the three schemes and the analytical solution for $z \in [-10, 0.5]$. Which of the schemes provides a better approximation of the exact (analytical) amplification for one time step? Why are the imaginary parts of λ and z irrelevant?

d) For $\lambda = -\sqrt{3}/2 + i$, do as in b) and calculate the numerical and analytical solutions. Show also the error $|u_{ex} - u_{num}|$ at the time $t = 3$ as a function of N in a double logarithmic plot. Explain the differences between the methods.

e) Based on the two examples considered above, discuss the usefulness, stability and accuracy of the methods.