S-114.100 Computational Science / Laskennallinen tiede. Fall 2004.

Assignment 1. Taylor series. Number representation and errors.
Chapters 0 and 1 in Lecture notes.
Due Wed 29.9.2004 (5 problems)

Web page: `www.lce.hut.fi/teaching/S-114.100/`

computer = programming task (C / C++ / Fortran / Java)
pencil and paper = solve on paper

Problem 1. (*computer*)
Study the following central difference formula:

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

as $h \to 0$. Using Taylor's theorem, we obtain that the *truncation error* for this formula is $-h^2/6 f'''(\xi)$ for some ξ in the interval $(x-h, x+h)$. In a computer, the result is also subject to a *rounding error*. The total error is given by the sum of these two errors.

Write and run a code which computes the approximate values of the truncation error and the total error as a function of h for $f(x) = sin(x)$ at $x = 0.5$. On the same graph, plot the results showing the rounding error (ε_r), the truncation error (ε_t) and the total error ($\varepsilon = \varepsilon_r + \varepsilon_t$). Use log-scale, i.e. plot $\log_{10} h$ (x-axis) vs. $\log_{10} |error|$ (y-axis). Analyze the results. (Hint: start with an initial value of h, e.g. $h = 0.5$, and divide the value of h at each step by some factor, e.g. $h \leftarrow h/4$).

Problem 2. (*pencil and paper*)
Suppose that you have a floating-point system in which numbers are represented in the following form

$$x = \pm 0.d_1 d_2 d_3 d_4 \times 10^n$$

where the exponent n is a 2-digit (decimal) integer and the normalized mantissa has four digits $d_1 - d_4$. Use this system to compute the difference $p - q$ where $p = 75640$ and $q = 4$.

Problem 3. (*pencil and paper / computer*)
(a) The harmonic series $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots$ is known to diverge to $+\infty$. The nth partial sum approaches $+\infty$ at the same rate as $\ln(n)$. **Euler's constant** is defined to be

$$\gamma = \lim_{n\to\infty}\left[\sum_{k=1}^{n}\frac{1}{k} - \ln(n)\right] \approx 0.57721$$

Write and test a program that uses a loop of 5000 steps to estimate Euler's constant. Print intermediate answers at every 100 steps. Compare to exact value.

(b) Prove (using pencil and paper) that Euler's constant can also be presented by

$$\gamma = \lim_{m\to\infty}\left[\sum_{k=1}^{m}\frac{1}{k} - \ln(m+\frac{1}{2})\right]$$

Write and test a program that uses $m = 1,2,3,\ldots,5000$ to compute γ by this formula. Compare the convergence to the formula given in part (a). For those interested, see D.W.Temple, "*A quicker convergence to Euler's constant*", Am. Math. Monthly **100**, 468 (1993).

Problem 4. (*pencil and paper*)
Let us have a floating-point representation in which the numbers x are expressed in the following form

$$x = \pm(0.b_1b_2b_3)_2 \times 2^{\pm m} \qquad m, b_i = \{0,1\}$$

List all positive numbers that are representable using this form. (You can form a table of the mantissa and the exponent bits and convert the combination into base 10 numbers). What if the representation is normed (the first bit b_1 is always 1)?

Problem 5. (*computer*)
Using a computer, examine how the computed values of the following functions behave near the point $x = 0$.

$$\text{(a)} \qquad f_1(x) = \frac{e^x - 1}{x}$$

$$\text{(b)} \qquad f_2(x) = \frac{e^x - e^{-x}}{2x}$$

In order to avoid *loss of significance* as a result of subtracting almost equal numbers, expand the exponential functions into Taylor series and use the resulting approximation to calculate the values of the functions f_1 and f_2 near zero. Estimate the error of the approximation by using Taylor's theorem on the series expansion of the exponential function ($f(x) = e^x$):

$$f(x) = f(0) + \sum_{i=1}^{n-1}\frac{x^i}{i!}f^{(i)}(0) + R(n)$$

where the error term is

$$R(n) = \frac{x^n}{n!}f^{(n)}(\xi) \qquad 0 \le \xi \le x$$