Exercise session on error definitions and the floating point approximation of real numbers

September 20 2021

Exercise 1

For a generic real number a consider the function $f(x) = a\exp(x)$ and its Taylor series approximation close to $x_0 = 0$ given by

$$a\left(1 + x + \frac{x^2}{2}\right).$$

1. Compute absolute and relative error for the approximation in $x = 0.2$ in the case $a = 2$.
2. Repeat the computation for $a = 10^5$.
3. Repeat the computation in $x = 4$ for $a = 10^3$.
4. Repeat the computation in all the previous cases for the approximation

$$a\left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}\right).$$

Exercise 2

Write a MATLAB script to compute $(1 + 10^{-k}) - 1$ for $k = 0, \ldots, 20$. Compute in each case the relative error of the result, using the formula

$$\frac{|[(1 + 10^{-k}) - 1] - 10^{-k}|}{10^{-k}}.$$

Repeat the error computation using the formula

$$\frac{|[1 + 10^{-k} - 10^{-k} - 1|}{10^{-k}}.$$

Repeat the computations for $(10^m + 10^{-k}) - 10^m$ for the same values of k and $m = 1, \ldots, 10$. In both cases, plot the error as a function of k using a logarithmic scale on the y axis (command `semilogy`). Explain the error behaviour based on the theory of the floating point representation of real numbers.

Exercise 3

Given real numbers $a > 0$ and x_0, define $x_n,\ n \geq 1$ iteratively by

$$x_{n+1} = \frac{1}{2}\Big(x_n + \frac{a}{x_n}\Big).$$

If x_0 is not too far from $\sqrt{a}$, one has $\lim_{n\to\infty} x_n = \sqrt{a}$. Write a MATLAB script to check this fact using a `for` or a `while` cycle and interrupting the computation when the relative error is smaller that some quantity ϵ. Run the script for different values of a, x_0, ϵ.

Exercise 4

Given real numbers $a > 0$ and x_0, define $x_n,\ n \geq 1$ iteratively by

$$x_{n+1} = \frac{x_n}{2}(3 - a x_n^2).$$

If x_0 is not too far from $1/\sqrt{a}$, one has $\lim_{n\to\infty} x_n = 1/\sqrt{a}$. Write a MATLAB script to check this fact using a `for` or a `while` cycle and interrupting the computation when the relative error is smaller that some quantity ϵ. Run the script for different values of a, x_0, ϵ.

Exercise 5

It is known that

$$1 = \lim_{h\to 0} \frac{\exp{(h)} - 1}{h}.$$

Write a MATLAB script to compute $\frac{\exp{(h)}-1}{h}$ for $h = 10^{-k}, k = 0, \ldots, 20$. Compute in each case the absolute error of the result. Plot the error as a function of k using a logarithmic scale on the y axis (command `semilogy`). Explain the error behaviour based on the theory of the floating point representation of real numbers.