Otto-von-Guericke-University Magdeburg
Max Planck Institute for Dynamics of Complex Technical Systems
Computational Methods in Systems and Control Theory

Dr. Jens Saak, Dipl.-Math. Martin Köhler
Website: http://www.mpi-magdeburg.mpg.de/csc/teaching/18ws/sc1/

Winterterm 2018/2019

Scientific Computing 1
6th Homework

Handout: 15th Nov. 2018 **Return:** 23th Nov. 2018

Exercise 1: (7 Points)

a.) Consider the following Integral:

$$y_n := \int_0^1 x^n \frac{1}{10+x} \mathrm{d}x \tag{1}$$

Compute y_0 and proof that

$$y_n = -10 y_{n-1} + \frac{1}{n} \tag{2}$$

holds $\forall n > 0$.

b.) Implement the recursion from (2) as a C function. Additionally implement the backward-recursion from y_n to y_{n-1} with the initial value $y_{30} = 0$ as a second function. Use double precision numbers for both functions.

Compute y_i for $0 \leq i \leq 30$ using those two functions. What do you recognize? Figure out possible reasons for this behavior.

Hint: Use $y_{20} \approx 4.34703 \cdot 10^{-3}$ to compare your results.

Exercise 2: (7 Points)

Consider a generic polynomial

$$P_n(x) := \sum_{i=0}^{n} a_i x^i \tag{3}$$

with $a_i \in \mathbb{R}$.

a.) Write a C function which takes the degree n, the coefficients a_i as an array, and x as inputs and evaluates $P_n(x)$ naively using Formula (3).

b.) Write a second C function with the same arguments which evaluates $P_n(x)$ using the Horner-Scheme, i.e.,

$$P_n(x) = (((a_n x + a_{n-1})x + a_{n-2})x + \ldots)x + a_0.$$

c.) Count the number of necessary floating point operations to evaluate $P_n(x)$ for both functions, separately.

Use these two functions to evaluate

$$P(x) = x^6 - 998x^5 - 998x^4 - 998x^3 - 998x^2 - 998x - 998$$

at $x = 999$. Compare and discuss the results of both C functions.

Exercise 3: **(3 Points)**

Determine the absolute and the relative error of 0.5403023059 and π in

a.) $\mathbb{M}(10, 3, -2, 2)$

b.) $\mathbb{M}(2, 3, -2, 3)$

c.) $\mathbb{M}(2, 5, -2, 2)$

Exercise 4: **(4 Points)**

Beside the single and double precison the half precision (16 bit) and the quad precision (128 bit) floating point number types are standardized by the IEEE 754-2008. Compute the smallest and the largest (in magnitude) normalized numbers in these two machine number sets.

Exercise 5: **(2 Points)**

Let x be the exact and $\hat{x}$ the computed solution of a problem. The classic definition of the relative error is $E_{rel}(\hat{x}) = |x - \hat{x}|/|x|$. In practice $\tilde{E}_{rel}(\hat{x}) = |x - \hat{x}|/|\hat{x}|$ is often used as a replacement. Find inequalities to estimate $\tilde{E}_{rel}(\hat{x})$ with respect to $E_{rel}(\hat{x})$. Is the use of $\tilde{E}$ instead of E justifiable?

Exercise 6: **(3 Points)**

Reformulate the following expressions to avoid cancellation:

a.) $\sqrt{1+x} - 1, \quad x \approx 0$

b.) $\frac{1-\cos x}{\sin x}, \quad x \approx 0$

c.) $\frac{1}{1+2x} - \frac{1-x}{1+x}, \quad x \approx 0$

Overall Points: 26