

Student Name:
Matrikel-Nr:

Numerical Methods in Informatics - Exercise 1 -

Assignment date: 05.10.2023 - Due date: 18.10.2023 (End of day)

Please upload your solutions to the OLAT system.

Theory

1.1 Floating Point Numbers

a) (5 Min, 1 Points) Errors and Precision

Please explain, why $0.3 + 0.9 \neq 1.2$ in a floating point system (IEEE-754 floating point standard with 32 bit). Which errors occur?

Hints:

- Use the website <http://weitz.de/ieee/> to do the calculation (make sure you select 32 bit);
- Use the website <https://www.h-schmidt.net/FloatConverter/IEEE754.html> to check the numbers.

Floating point values are always an approximation to the real number. When you add two floating point numbers together, there will be a slight deviation due to the limited precision of a 32 bit floating point number. This will lead to rounding errors.

b) (1 Min, 1 Points) Comparison

Given that you have some written programming-code and you want to test:

$0.3 + 1.1 == 1.4$

How would you do this?

1.2 Linear System

a) (15 Min, 7 Points) Solving Linear Systems

Please solve the following linear system $A\vec{x} = \vec{b}$ using elementary row operations and **applying partial pivoting**. Reach echelon form and then do back-substitution to get to the final result

$$A = \begin{pmatrix} 6 & -10 & 14 \\ -9 & 6 & -3 \\ -3 & 3 & -2 \end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix} 16 \\ -6 \\ -3 \end{pmatrix}$$

1.3 Transformation

a) (3 Min, 1 Points) Function

Given the standard matrix

$$A = \begin{pmatrix} 2 & 4 & 0 & 1 & 3 \\ 0 & 1 & -2 & 2 & 1 \\ 5 & 7 & 1 & -5 & 6 \end{pmatrix}$$

describe and write down the linear transformation function $T(x)$ such that $T(x) = Ax$. Describe the functions how they modify an input vector x .