Question:	1	2	3	4	5	Total
Points:	15	25	10	40	10	100

Floating point numbers

1. Floating point numbers typically represented in computers in the following binary form:

$$\pm\left(1+\frac{b_1}{2}+\frac{b_2}{2^2}+\ldots+\frac{b_d}{2^d}\right)\times 2^E$$

(a) (5 points) What is the (approximate) value of machine epsilon for a microprocessor that uses $d = 8$? Briefly explain.

Machine epsilon, ε, is the separation between 1 and the next number that is larger than 1, i.e. $1+1/2^d \longrightarrow \varepsilon = 2^{-d} = 2^{-8} = 1/256 \approx 4\cdot 10^{-3} \sim 10^{-3}$

(b) (5 points) For the same microprocessor, how many floating point numbers x, such that $4 \le x < 5$ are there? Briefly explain.

Floating point numbers are equidistant on the interval between 4 and 8. Hence, $N(4\le x<5) = \frac{1}{4}(4\le x<8)$. We know that $N(4\le x<8) = 2^d$. Thus, $N(4\le x<5) = \frac{2^8}{4} = 2^6 = 64 \sim 10^2$

(c) (5 points) For the same microprocessor, assuming that the smallest value of E is -16, what is (approximately) the smallest positive floating point number? Briefly explain.

x_{min} = (smallest possible mantissa) x (smallest exponent)

$= 1\times 2^{-16} = 2^{-10}\cdot 2^{-6} \approx 10^{-3}\frac{1}{64} \approx \frac{10^{-3}}{50} \approx 2\cdot 10^{-5}$

Systems of linear equations

2. The chemical equation

$$x_1\,[Ca(OH)_2] + x_2\,[HNO_3] \to x_3\,[Ca(NO_3)_2] + 2\,[H_2O]$$

indicates that x_1 molecules of calcium hydroxide $Ca(OH)_2$ combine with x_2 molecules of nitric acid HNO_3 to yield x_3 molecules of calcium nitrate $Ca(NO_3)_2$ and 2 molecules of water H_2O.

Since atoms are not destroyed or created in chemical reactions, the balance of oxygen atoms requires that

$$2x_1 + 3x_2 = 6x_3 + 2.$$

The balance of hydrogen atoms requires that

$$2x_1 + x_2 = 4.$$

The balance for nitrogen atoms requires that

$$x_2 = 2x_3$$

(a) (5 points) Rewrite the balance equations above in matrix form $Ax = b$:

$$\begin{cases} 2x_1 + 3x_2 - 6x_3 = 2 \\ 2x_1 + x_2 = 4 \\ x_2 - 2x_3 = 0 \end{cases} \qquad A = \begin{pmatrix} 2 & 3 & -6 \\ 2 & 1 & 0 \\ 0 & 1 & -2 \end{pmatrix}; \; b = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}$$

(b) (5 points) Verify that the following two matrices are indeed the results of LU-factorization of A:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 2 & 3 & -6 \\ 0 & -2 & 6 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$L \cdot U = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & -1/2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 & -6 \\ 0 & -2 & 6 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 & -6 \\ 2 & 1 & 0 \\ 0 & 1 & -2 \end{pmatrix} \equiv A$$

(c) (5 points) Use L and U to calculate the determinant of matrix A. Write you calculations below:

$$\det(A) = \det(L \cdot U) = \underbrace{\det(L)}_{1} \cdot \det(U) =$$

$$= 2 \cdot (-2) \cdot 1 = -4$$

(d) (5 points) Use the forward substitution to solve the equation $Ly = b$. Write you calculations below:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}$$

$1 \cdot y_1 = 2 \rightarrow y_1 = 2$

$1 \cdot y_1 + 1 \cdot y_2 = 4 \rightarrow y_2 = 4 - y_1 = 2$

$-\frac{1}{2} y_2 + y_3 = 0 \rightarrow y_3 = \frac{y_2}{2} = 1$

(e) (5 points) Use the backward substitution to solve the equation $Ux = y$. Verify by direct substitution that x is the solution of $Ax = b$. Write you calculations below:

$$\begin{pmatrix} 2 & 3 & -6 \\ 0 & -2 & 6 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

$1 \cdot x_3 = 1 \rightarrow x_3 = 1$

$-2 \cdot x_2 + 6 \cdot x_3 = 2 \rightarrow x_2 = 3x_3 - 1 = 2$

$2 \cdot x_1 + 3 \cdot x_2 - 6 \cdot x_3 = 2 \rightarrow$

$x_1 = 1 - \frac{3 \cdot 2}{2} + \frac{6 \cdot 1}{2} = 1$

$x = (1\ 2\ 1)^T$

3. (10 points) You wrote your own function to solve a system of linear equations. It takes about 10 seconds (on a slow computer) to solve the system of 100 equations with 100 unknowns. **Estimate** how long it would take to solve a system of 200 linear equations with 200 unknowns if your code implements LU-factorization method to solve the equations. Present your answer and explain your reasoning in the gitlab's README.md file.

Matlab

4. (40 points) TBA

Git and Gitlab

5. (10 points) Upload all the code you wrote/used for this exam:

   1. Create a new gitlab project called **midterm1-sample** (the name must be exactly as shown)
   2. Add *README.md* file to your project and edit it to add some meaningful content
   3. Upload your matlab code to your project
   4. Grant the access to your project (with the permission of the *Reporter*) to the instructor.