Mathematics of Learning – Worksheet 1 - – Discussion on October 19/20th, 2023

- The exercise sheets will be uploaded every Monday. Solution sketches will be uploaded one week later.
- You can hand in your own solutions via StudOn and we correct them - this is not mandatory. Please hand in in small groups of 2-3 students.
- For questions, please use the forum on StudOn since other students may have similar questions. If you have a more personal question about the exercises please send an email to ehsan.waiezi@fau.de or lars.weidner@fau.de respectively.

Basics [Solving linear equation systems.][1]
Solve the linear equation system $Ax = b$. A and b are given as

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 9 \\ 1 & 8 & 27 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 2 \\ 7 \end{pmatrix}.$$

Control yourself, if your solution is right. If you need some practice, generate some random linear equation systems and solve them.

Solution. We apply the gaussian elimination algorithm.
1. Subtract line 1 from lines 2 and 3, 2. Subtract 3 times line 2 from line 3, 3. divide through diagonal elements:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 1 & 4 & 9 & 2 \\ 1 & 8 & 27 & 7 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 2 & 6 & -1 \\ 0 & 6 & 24 & 4 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 2 & 6 & -1 \\ 0 & 0 & 6 & 7 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 3 & -0.5 \\ 0 & 0 & 1 & \frac{7}{6} \end{array}\right)$$

4. Insert backwards.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 3 & -0.5 \\ 0 & 0 & 1 & \frac{7}{6} \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 3 & 3 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & \frac{7}{6} \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & 7.5 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & \frac{7}{6} \end{array}\right).$$

Basics [Norms.]
A mapping $\|\cdot\|$ from any (real) vector space V to the non-negative real numbers $\mathbb{R}$ (bonus exercise: find a mapping on the real numbers, which has the three properties and takes some negative values, or show that it is impossible) is called a norm, whenever

$$\|v+w\| \leq \|v\| + \|w\|, \quad \|v\| = 0 \implies v = 0_V, \quad \|\lambda v\| = |\lambda|\|v\| \text{ for all } \lambda \in \mathbb{R}, v, w \in V.$$

Prove for the following statements if they are true or false.

[1] There are lots of nice tutorial books for linear algebra and analysis available in our library. For a less formal introduction, you can, e.g., also consult wikipedia ;)

1. Let $V = \mathbb{R}^n$ for some $n \in \mathbb{N}$. The euclidean norm

$$\|v\|_2 := \sqrt{\sum_{i=1}^{n} v_i^2}$$

is a norm.

Solution. This is a norm, since we can check that the three properties hold.
Property 1. Let $v, w \in \mathbb{R}^n$ be arbitrary vectors. Since both sides of property 1 are positive, it suffices to check if

$$\|v+w\|_2^2 \leq (\|v\|_2 + \|w\|_2)^2 = \|v\|_2^2 + \|w\|_2^2 + 2 \cdot \|v\|_2 \|w\|_2$$

The left hand side is just

$$\sum_{i=1}^{n} v_i^2 + \sum_{i=1}^{n} w_i^2 + 2 \sum_{i=1}^{n} v_i w_i = \|v\|_2^2 + \|w\|_2^2 + 2 \sum_{i=1}^{n} v_i w_i,$$

the latter term being the scalar product of v and w (which induces the euclidean norm). Hence the inequality is equivalent to

$$\langle v, w \rangle \leq \|v\|_2 \|w\|_2,$$

which is the well-known Cauchy-Schwarz Inequality, and thus holds true.
Property 2. We assume that $\|v\|_2 = 0$. Then it holds, that

$$\|v\|_2 = 0 \implies \sqrt{\sum_{i=1}^{n} v_i^2} = 0 \implies \sum_{i=1}^{n} v_i^2 = 0 \implies v_i^2 = 0 \quad \forall i \in [n] \implies v = 0.$$

Property 3. It holds for all $\lambda \in \mathbb{R}$ that

$$\|\lambda v\|_2 = \sqrt{\sum_{i=1}^{n} (\lambda v_i)^2} = \sqrt{\sum_{i=1}^{n} \lambda^2 v_i^2} = \sqrt{\lambda^2 \sum_{i=1}^{n} v_i^2} = |\lambda| \|v\|_2.$$

2. Let $V = \mathbb{R}^n$ for some $n \in \mathbb{N}$. The mapping

$$\|v\|_{\frac{1}{2}} := (\sum_{i=1}^{n} \sqrt{|v_i|})^2$$

is a norm.

Solution. This is not a norm, since property 1 does not hold. Consider $\mathbb{R}^2$ and unit vectors $e_1 = (1,0)^T$ and $e_2 = (0,1)^T$. It holds that $\|e_1\|_{\frac{1}{2}} = \|e_2\|_{\frac{1}{2}} = 1$ but on the other hand

$$\|e_1 + e_2\|_{\frac{1}{2}} = (\sqrt{1} + \sqrt{1})^2 = 4 > 2 = \|e_1\|_{\frac{1}{2}} + \|e_2\|_{\frac{1}{2}}.$$

3. Let V be the space of convergent sequences. The mapping

$$\|v\|_{lim} := \lim_{n \to \infty} v_n$$

is a norm.

Solution. This is not a norm, since (e.g.) the sequence $1, 0, 0, \ldots$ violates property 2: The sequence is not the 0-Sequence, but the limit is 0.

Basics [Big-O (Landau) Notation.]

a) Express the relationships of the functions n^{1000}, 2^n, e^n, e^{n^2}, $n!$ and n^n with the help of Landau's notation (i.e., prove, if $f \in \mathcal{O}(g), f \in \Omega(g)$ or $f \in \Theta(g)$ for every pair of the functions above). Prove your statements.

b) An equivalence relation is a homogeneous relation over some set M[2], which is

1. Reflexive: $m_1 \simeq m_1$,
2. Symmetric: $m_1 \simeq m_2$ implies $m_2 \simeq m_1$,
3. Transitive: If $m_1 \simeq m_2$ and $m_2 \simeq m_3$ then $m_1 \simeq m_3$

for all $m_1, m_2, m_3 \in M$. For example, the relation $\simeq$ on $\mathbb{R}^n$, $v \simeq w$ if and only if $v_1 = w_1$ is an equivalence relation.
Prove or disprove, that $\simeq_L$ which we define as

$$f \simeq_L g \text{ if and only if } f \in \Theta(g)$$

is an equivalence relation on the set of mappings from $\mathbb{N}$ to $\mathbb{R}_{>0}$.

Solution.

a) Using transitivity from part b), we show the relation for the sequence

$$n^{1000}, 2^n, e^n, n!, n^n, e^{n^2}.$$

It is never a bad idea to check if we can calculate the limits of the quotients corresponding (positive) functions. The reason is (here exemplarily shown for the first definition; Bonus: transfer the proof to the other two definitions) that you can equivalently define the Big-O sets as follows:

$$\begin{aligned}
\mathcal{O}(g(n)) &:= \left\{ f(n) : \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty \right\}, \\
\Omega(g(n)) &:= \left\{ f(n) : \liminf_{n\to\infty} \frac{f(n)}{g(n)} > 0 \right\}, \\
\Theta(g(n)) &:= \left\{ f(n) : 0 < \liminf_{n\to\infty} \frac{f(n)}{g(n)} \le \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty \right\}.
\end{aligned}$$

Exemplary proof for the first equivalence. If for a finite $c > 0$ it holds that

$$\limsup_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \sup_{m\ge n} \frac{f(m)}{g(m)} = c$$

(Bonus: Why does the limit always exist?) it holds that there is for every $\epsilon > 0$, for example, lets take $\epsilon = 1$ (but any other positive number would also do it), an $n_0 \in \mathbb{N}$ such that it holds that

$$\sup_{m\ge n_0} \left| \frac{f(n)}{g(n)} - c \right| \le \epsilon \implies \sup_{m\ge n_0} \frac{f(n)}{g(n)} \le c + \epsilon = c + 1$$

[2] i.e. $\simeq : M \times M \to \{0,1\}$; one would rather write $m_1 \simeq m_2$ instead of $\simeq (m_1, m_2) = 1$ and $m_1 \not\simeq m_2$ instead of $\simeq (m_1, m_2) = 0$.

This is equivalent to $f(n) \leq (c+1)g(n)$ for all $n \geq n_0$.

We skip the first comparison for the moment.

$$\lim_{n\to\infty} \frac{2^n}{e^n} = 0$$

Hence, $2^n \in \mathcal{O}(e^n)$, $e^n \in \Omega(2^n)$, nothing else (this is always the implication if the limit is 0, so we skip that in the following).

$$\lim_{n\to\infty} \frac{e^n}{n!} \leq \lim_{n\to\infty} e \cdot \frac{e}{1} \cdot \frac{e}{2} \cdot \frac{e^{n-3}}{3 \cdot \dots \cdot (n-1)} \cdot \frac{1}{n} \leq \lim_{n\to\infty} \frac{e^3}{2n} = 0.$$

$$\lim_{n\to\infty} \frac{n!}{n^n} \leq \lim_{n\to\infty} \frac{1}{n} = 0.$$

$$\lim_{n\to\infty} \frac{n^n}{e^{n^2}} = \lim_{n\to\infty} \frac{e^{n\log n}}{e^{n^2}} = \lim_{n\to\infty} \frac{1}{e^{n^2 - n\log n}} = 0.$$

For the first comparison, this is a bit more complicated. But lets analyse the sequence

$$a_n = \frac{n^{1000}}{2^n}.$$

A sufficient criterion for the convergence of a sequence towards 0 is that the ratio of two subsequent sequence elements is bounded from below by a number larger than 1. But this is the case:

$$\lim_{n\to\infty} \frac{a_n}{a_{n+1}} = \lim_{n\to\infty} \frac{\frac{n^{1000}}{2^n}}{\frac{(n+1)^{1000}}{2^{n+1}}} = \lim_{n\to\infty} \frac{2}{(\frac{n+1}{n})^{1000}} = 2,$$

hence, a_n goes to 0 and we get the same result as for the comparisons above.

b) Reflexivity: Obviously, $f \in \Theta(f)$ holds (choose $c_1 = 1, c_2 = 1$, $n_0 = 1$ and check $0 \leq 1 \cdot f(n) \leq f(n) \leq 1 \cdot f(n)$ for all $n \geq n_0$).
Symmetry: Let $f \in \Theta(g)$. Then there exist constants $c_1, c_2 \in \mathbb{R}_{>0}$ and $n_0 \in \mathbb{N}$, such that $0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n)$ for all $n \geq n_0$, but this implies directly $0 \leq \frac{1}{c_2} f(n) \leq g(n) \leq \frac{1}{c_1} f(n)$ for all $n \geq n_0$, so taking $\tilde{c}_1 = \frac{1}{c_2}$, $\tilde{c}_2 = \frac{1}{c_1}$ and $\tilde{n}_0 = n_0$ leads to $g \in \Theta(f)$.
Transitivity. Let f, g, h be functions according to our requirements, and $f \in \Theta(g), g \in \Theta(h)$. Then there are positive constants c_1^f, c_2^f and c_1^g, c_2^g, and natural numbers n_0^f, n_0^g, such that

$$0 \leq c_1^f g(n) \leq f(n) \leq c_2^f g(n) \text{ for all } n \geq n_0^f, 0 \leq c_1^g h(n) \leq g(n) \leq c_2^g h(n) \text{ for all } n \geq n_0^g.$$

Inserting the second inequalities in the first ones, this implies, that

$$0 \leq c_1^f c_1^g h(n) \leq f(n) \leq c_2^f c_2^g h(n) \text{ for all } n \geq \max\{n_0^f, n_0^g\}.$$

This implies (taking $c_1^{\text{trans}} = c_1^f \cdot c_1^g$, $c_2^{\text{trans}} = c_2^f \cdot c_2^g$ and $n_0^{\text{trans}} = \max\{n_0^f, n_0^g\}$) that $f \in \Theta(h)$.