School of Mathematics & Computer Science
Prof. Dr. Dr. h.c. Florian Rupp
Akaki Matcharashvili

Spring Term 2023
1st Mock Quiz

1st Mock Quiz
Introduction to Optimization

Date of Examination:	self study
Start of Examination:	self study
Duration:	1 hour

First Name and Family Name: ______________________

Student ID ______________________

Please, give your answers underneath the questions as indicated. In case you run out of space, extra sheets will be handed to you. Use black or dark blue ink only (pencils are allowed for drawings only). You are allowed to have one page of DIN A4 format of notes with you (both sides can be used). Any electronic devices are prohibited.

Students must at all times abide by KIU's Academic Regulations. In particular, you must refrain from cheating, plagiarizing or any act that may compromise academic integrity.

❀ ❀ ❀ **Good Luck** ❀ ❀ ❀

Please check the number of problems: There are **3 problems** on the **pages 2 to 4**!

question	1	2	3	total
max	(5)	(5)	(5)	(15)
credits				

Note, the real quiz will have a bonus question that is very similar to one of the questions from the section 'Additional/ Central Exercise Problems' of your homework assignment papers and whose solutions are provided to you in written form.

1. **[5 credits] Optimality Conditions Applied:**

   Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ with $f(x_1, x_2) = x_1^2 + x_2^3$.

   a) Show that f has a critical point at $(x_1, x_2) = (0; 0)$ and that its associated Hessian matrix is positive semi-definite.

   b) Is f a convex function and what does this mean for the existence of its minimizers?

   c) Show that $(0; 0)$ is not a local minimizer of f.

   d) Is f coercive?

2. **[5 credits] Gradient Descent:**

Apply the Gradient Descent method on the function $f : \mathbb{R}^2 \to \mathbb{R}$ with $f(x_1, x_2) = \frac{1}{2}x_1^2 + x_2^2$, $x = (x_1, x_2)$, starting at the point $x^0 = (2, 2)$ and using constant step size $\sigma = \frac{1}{2}$.

a) Determine the values of the iterates x^1, x^2, and x^3 generated by this method

b) Determine the global optimum of f. What do you observe in terms of convergence of x^1, x^2, x^3 towards this optimum?

3. **[5 credits] Speed of Convergence:**

Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive definite and $b \in \mathbb{R}^n$. From the lecture, you know that the speed of convergence of the Gradient Descent Method with exact line search applied to the minimization of $f_A(x) = \frac{1}{2}x^T Ax - b^T x$ can be described, for instance, by

$$\|x^{k+1} - \overline{x}\| \;\leq\; \sqrt{\kappa(A)} \left(\frac{\kappa(A) - 1}{\kappa(A) + 1}\right)^k \|x^k - \overline{x}\|\,,$$

where $\overline{x}$ is the minimizer for $f(x)$ and $\kappa(A)$ the condition number of A. Let the following two system matrices be given

$$A \;=\; \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \qquad \text{and} \qquad B \;=\; \begin{pmatrix} 40 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -10 \end{pmatrix}.$$

a) Compute the condition number of each of them.

b) Interpret the speed of convergence for $\min f_A(x)$ and $\min f_B(x)$.

c) How can diagonal scaling speed-up convergence of $\min f_B(x)$? Justify your answer by giving the scaling matrix as well as the condition number of the thus transformed problem.