# Exercise 1 - Math Refresher

Winter term 2019/2020

student1, student2, student3

---


## General information

The assignments are voluntary. All students who choose to participate have to form groups comprising three to four students (not more and not less). The groups do not have to be static, you may form new groups for each assignment. You have **two weeks** to answer the questions and to submit your work. The solutions are going to be presented and discussed after the submission deadline. Sample solutions will **not** be uploaded. However, you are free to share correct solutions with your colleagues **after they have been graded.**

## Formal requirements for submissions

Please submit your solutions via Moodle (as a .zip file) as well as in printed form. The .zip file must contain one .pdf file for the pen-and-paper tasks as well as one .py file per programming task. Only pen-and-paper tasks have to be printed, you do not have to print the source code. Only one member of the group has to submit the solutions. Please make sure to specify the matriculation numbers (**not the names!**) of all group members so that all participants receive the points they deserve!

Please refrain from submitting hand-written solutions or images of solutions (*.png* / *.jpg* files). Rather use proper type-setting software like LaTeX or other comparable programs. If you choose to use LaTeX, you may want to use the template files provided.

Code assignments have to be done in Python. Please submit *.py* files (**no jupyter notebooks**). The following packages are allowed for code submissions: `numpy`, `pandas` and `scipy`. Please ask **beforehand**, if you want to use a specific package not mentioned here. Finally, do not use already implemented models (e.g. from `scikit-learn`).

## Grading details

Your homework is going to be corrected and given back to you. Correct solutions are rewarded with a bonus for the exam which amounts to at most ten percent of the exam, if all solutions submitted by you are correct (this corresponds to at most six points in the exam). It is still possible to achieve full points in the exam, even if you choose not to participate in the assignments (it is additional). The function which is used to compute the bonus is given by:

$$b(a) = \min\left(B, \left\lceil \frac{B}{A^2} \cdot a^2 \right\rceil\right) \tag{1}$$

- $b$ denotes the number of bonus points you get for the exam (this is up to you)
- $B$ refers to the maximum attainable bonus points for the exam (six points)
- $A$ denotes the maximum attainable points in the assignments (40 points)
- $a$ is the score you achieved in the assignments (this is up to you)

**Please note:** You have to pass the exam **without the bonus points!** This means that it is not possible to turn a failing grade ($= 5.0$) into a passing grade ($\leq 4.0$). The bonus points will be taken into account in case you have to repeat the exam (i. e. they do not expire if you fail the first attempt).

## Important!

**The solutions have to be your own work. If you plagiarize, you will lose all bonus points!**

# 1 Linear Algebra Refresher

a) Matrix Operations (1 point)

A fellow student suggests that matrix addition and multiplication are very similar to scalar addition and multiplication, i. e. commutative, associative and distributive. Is this a correct statement? Prove it mathematically or disprove it by providing at least one counter example per property (commutativity, associativity, distributivity).

**Solution:**

b) Matrix Inverse (1 point)

What is a matrix inverse? How can you build the inverse of a non-square matrix? You would like to invert a matrix $\boldsymbol{M} \in \mathbb{R}^{2\times 3}$, write down the equation for computing it and specify the dimensionality of the matrices after each single operation (e. g. multiplication, inverse).

**Solution:**

c) Eigenvectors and Eigenvalues (1 point)

Explain what eigenvectors and eigenvalues of a matrix $\boldsymbol{A}$ are. Why are they relevant in machine learning?

**Solution:**

## 2 Statistics Refresher

a) Terminology (1 point)

What is a random variable? What is a probability density function (PDF)? What is a probability mass function (PMF)? What do a PDF and a PMF tell us about a random variable?

**Solution:**

b) Expectation and Variance (1 point)

State the general definition of expectation and variance for the probability density $f : \Omega \to \mathbb{R}$ of a continuous random variable. What do expectation and variance express?

**Solution:**

# 3 Optimization

a) Numerical Optimization - Gradient Descent (5 points)

Implement a simple gradient descent algorithm for finding a minimum of the Rosenbrock function with $n = 2$ using Python and NumPy:

$$f(\boldsymbol{x}) = \sum_{i=1}^{n-1} \left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right]$$

Submit your code and a plot of the learning curve for the best run of your gradient descent implementation. Which learning rate worked best? (Hint: You need to find the first derivative(s) of $f(\boldsymbol{x})$ for $n = 2$ and iteratively evaluate them during gradient descent. Automatic differentiation tools are not allowed for this exercise. Choose a random starting point for the parameters, for example $\boldsymbol{x} \in [-2, 2]$.)

**Solution:**