Exercise 1 - Math Refresher

Winter term 2019/2020

name1, name2, name3

Important

Please solve the assignments in groups of 3 to 4 students. 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**. Please submit your solutions via Moodle **and** in printed form. Only one member of the group has to submit the solutions. Therefore, make sure to specify the names of all group members. Please do not submit hand-written solutions, rather use proper type-setting software like LaTeX or other comparable programs.

Your homework will be corrected and given back to you. Correct solutions are rewarded with a bonus for the exam (max. 10 percent, if all solutions submitted are correct). **Please note:** You have to pass the exam **without the bonus points**! *(i.e. it is not possible to turn 5.0 into 4.0)* The solutions have to be your own work. If you plagiarize, you will lose all bonus points!

Further remarks:

- Code assignments have to be done in `Python`
- The following packages are allowed: `numpy`, `pandas`
  (please ask, if you want to use a specific package not mentioned here)
- **Do not use already implemented models** (e. g. from `scikit-learn`)

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: