Exercise for Machine Learning (SS 20)

Assignment 0: Introduction (Theory)

Prof. Dr. Steffen Staab, steffen.staab@ipvs.uni-stuttgart.de

Alex Baier, alex.baier@ipvs.uni-stuttgart.de

Janik Hager, janik-manel.hager@ipvs.uni-stuttgart.de

Analytic Computing, IPVS, University of Stuttgart

Submit your solution in Ilias as either PDF for theory assignments or Jupyter notebook for practical assignments.
Mention the names of all group members and their Email addresses in the file.
Submission is possible until the following Monday, 27.04.2020, at 14:00.

This assignment does not count towards the admission. Nonetheless you should try to upload your solution as PDF to Ilias. This will help you and us in testing the platform and identifying potential problems.
The first two tasks were taken from "Exercise 1" of the previous semester held by Prof. Dr. Marc Toussaint.

1 Matrix equations

1. Let X, YA be arbitrary matrices, A invertible. Solve for X:

   $$XA + A^\top = \mathbf{I}$$

2. Let X, A, B be arbitrary matrices, $(C - 2A)^\top$ invertible. Solve for X:

   $$X^\top C = [2A(X + B)]^\top$$

3. Let $x \in \mathbb{R}^n$, $y \in \mathbb{R}^d$, $A \in \mathbb{R}^{d \times n}$, $A^\top A$ invertible. Solve for x:

   $$(Ax - y)^\top A = \mathbf{0}_n^\top$$

4. As above, additionally $B \in \mathbb{R}^{n \times n}$, B positive-definite. Solve for x:

   $$(Ax - y)^\top A + x^\top B = \mathbf{0}_n^\top$$

2 Vector derivatives

Let $x \in \mathbb{R}^n$, $y \in \mathbb{R}^d$, $A \in \mathbb{R}^{d \times n}$.

1. What is $\frac{\partial}{\partial x} x$? (Of what type/dimension is this thing?)
2. What is $\frac{\partial}{\partial x} [x^\top x]$?
3. Let B be symmetric and positive definite. What is the minimum of $(Ax-y)^\top(Ax-y) + x^\top Bx$ w.r.t. x?

3 Error Measures

Let $y, \hat{y} \in \mathbb{R}^n$ be n true and predicted values of a regression problem.

1. Formally define the error measures *Mean Squared Error* (MSE) and *Mean Absolute Error* between y and $\hat{y}$.

2. How would choosing MAE or MSE as objective function for a regression problem impact the resulting prediction model? How do MSE and MAE differ?

3. Let $y, \hat{y} \in \{0, 1\}^n$ be n true and predicted labels of a binary classification problem. What would the MSE and MAE calculate in this case?