**Deadline:** Fr. Mai 17, 14:00 Drop your printed or legible handwritten submissions into the boxes at Samelsonplatz, or a .pdf file via LearnWeb.

**NOTE:** Make sure to use the 2019 slides, there are some errors in the 2018 version

### 1 SVM training

In all tasks below, use a regularization constant of  $\lambda = 1$  and initial parameters  $\beta = (0, 0, 0)$  (including bias) and  $\alpha = (0, 0, 0, 0)$  respectively. For simplicity always take the samples in order.

Given the dataset from Table 1,

- **A.** [5p] Perform 2 iterations of primal Gradient Descent, with learn-rate  $\eta = \frac{1}{2}$
- **B.** [5p] Perform 2 iterations of PEGASOS with batch size K = 2
- C. [5p] Perform 2 iterations of dual coordinate descent (don't forget clipping!)
  - Table 1: dataset  $x_1$  $x_2$ y20 +120 -122+11 1  $^{-1}$

### 2 SVM theory

**A.** [5p] Compare SVM with a linear kernel to logistic regression. Both try to learn a separating hyperplane. How are they different?

# 1/1

## (5 points)

## (15 points)