Machine Learning 1 Prof. Schmidt-Thieme, Randolf Scholz

Deadline: Fr. January 10th , 15:00 Drop your printed or legible handwritten submissions into the boxes at Samelsonplatz. Alternatively upload a .pdf file via LearnWeb. (e.g. exported Jupyter notebook)

1. SVM practice

A. [6p] To train an SVM we have to solve the following optimization problem (primal form)

$$\min_{\beta_0,\beta,\xi} \frac{1}{2} \|\beta\|^2 + \gamma \sum_{i=1}^n \xi_i \quad \text{s.t.} \quad \begin{array}{c} y_i \left(\beta_0 + \langle \beta | x_i \rangle\right) \ge 1 - \xi_i \\ \xi_i \ge 0 \end{array} \quad \text{for all } i \tag{1}$$

Or equivalently the dual form (and recover $\beta = \sum_{n=1}^{N} \alpha_n y_n x_n$, $\beta_0 = \frac{1}{|\{n:\alpha_n \neq 0\}|} \sum_{n:\alpha_n \neq 0} (y_n - \langle \beta | x_n \rangle)$

$$\max_{\alpha} -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \langle x_i | x_j \rangle + \sum_{i=1}^{n} \alpha_i \quad \text{s.t.} \quad \sum_{i=1}^{n} \alpha_i y_i = 0$$

$$0 \le \alpha_i \le \gamma$$
(2)

Rewrite both of these problems as an inequality constrained QPs, i.e. optimization problems of the form:

$$\min_{z} \frac{1}{2} z^{\mathsf{T}} C z + c^{\mathsf{T}} z \quad \text{s.t.} \quad \begin{array}{c} A z = a \\ B z < b \end{array}$$
(3)

By explicitly constructing the matrices/vectors A, a, B, b, C, c. Note that z should be the concatenation of all variables.

B. [2p] Let $\gamma = 1$. Explicitly construct the matrices/vectors (A, a, B, b, C, c) of the **primal form** given the data from Table 1.

C. [2p] Let $\gamma = 1$. Explicitly construct the matrices/vectors (A, a, B, b, C, c) of the dual form given the data from Table 1.

D. [2p] Explain why the Active Set Algorithm is generally not applicable to the primal form. Is it always applicable to the dual form?

x_1	x_2	y
0	0	-1
-1	-1	1
1	0	1
0	1	1

Table 1: toy data

(12 points)