Machine Learning (911.236)

Exercise sheet **D**

Linear predictors

Exercise 1. 10 P.

Show how to write the following problem

$$\min_{\mathbf{w}\in\mathbb{R}^d} \sum_{i=1}^{m} |\langle \mathbf{w}, \mathbf{x}_i\rangle - y_i|$$

as a **linear program** (see slides on implementing the ERM paradigm for halfspaces as a linear program). This problem is known as linear regression with absolute value loss. It is an example of implementing the ERM paradigm for regression.

Hint(s): Think about what it means to minimize over the sum of the absolute values. An idea would be to let $|\langle \mathbf{w}, \mathbf{x}_i\rangle - y_i| = s_i$ and then formulate the objective function and the constraints.

Exercise 2. 15 P.

Consider the class of closed balls in $\mathbb{R}^d$, i.e.,

$$\mathcal{B}_d = \{B_{\mathbf{v},r} : \mathbf{v} \in \mathbb{R}^d, r > 0\}$$

with

$$B_{\mathbf{v},r}(\mathbf{x}) = \begin{cases} 1 & \text{, if } \|\mathbf{x}-\mathbf{v}\| \leq r \\ 0 & \text{, else} \end{cases}$$

Now, consider the mapping

$$\phi : \mathbb{R}^d \to \mathbb{R}^{d+1}$$

defined by

$$\phi(\mathbf{x}) = (\mathbf{x}, \|\mathbf{x}\|^2)$$

Show that if $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are shattered by $\mathcal{B}_d$, then $\phi(\mathbf{x}_1), \ldots, \phi(\mathbf{x}_m)$ is shattered by the class of halfspaces in $\mathbb{R}^{d+1}$. Also, what does this say about the VC dimension of $\mathcal{B}_d$? Assume, in this example, that $\mathrm{sign}(0) = 1$.

Total #points: 25 P.