University of Salzburg Lecturer: Roland Kwitt

# **Machine Learning** (911.236)

Exercise sheet **F**

**Exercise 1.** 5 P.

We consider the problem of **logistic regression**. Our hypothesis class is given by

$$\mathcal{H} = \mathcal{X} = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \le B, B > 0\} \ .$$

Note that our data space $\mathcal{X}$ is equal to $\mathcal{H}$, so all of our samples lie in a ball of radius $B$. Further, $\mathcal{Y} = \{\pm 1\}$ and our loss function is

$$l : \mathcal{H} \times \mathcal{Z} \to \mathbb{R}, \quad (\mathbf{w}, (\mathbf{x}, y)) \mapsto l(\mathbf{w}, (\mathbf{x}, y)) = \log(1 + \exp(-y\langle \mathbf{w}, \mathbf{x}\rangle))$$

with $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$. Show that the resulting learning problem $(\mathcal{H}, \mathcal{Z}, l)$ is

1. Convex-Lipschitz-Bounded, and
2. Convex-Smooth-Bounded.

Specify the Lipschitz and smoothness constants ($\rho$ and $\beta$, resp.).

**Exercise 2.** 5 P.

In the second exercise, we consider the problem of **learning halfspaces using the hinge-loss**. For this learning problem, the hypothesis class is given by

$$\mathcal{HS}_d = \{\mathbf{x} \mapsto \text{sign}(\langle \mathbf{w}, \mathbf{x}\rangle) : \mathbf{w} \in \mathbb{R}^d\} \ .$$

Assume $\mathcal{Z} = \mathcal{X} \times \{\pm 1\}$ with

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \le R\}$$

and the *hinge*-loss is formally defined as

$$l : \mathcal{H} \times \mathcal{Z} \to \mathbb{R}, \quad (\mathbf{w}, (\mathbf{x}, y)) \mapsto l(\mathbf{w}, (\mathbf{x}, y)) = \max\{0, 1 - y\langle \mathbf{w}, \mathbf{x}\rangle\} \ .$$

We already know (from the lecture) that $l(\cdot, z)$ is convex. Show that it is $R$-Lipschitz, i.e., let $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^d$ and show

$$|\max\{0, 1 - y\langle \mathbf{w}_1, \mathbf{x}\rangle\} - \max\{0, 1 - y\langle \mathbf{w}_2, \mathbf{x}\rangle\}| \le R\|\mathbf{w}_1 - \mathbf{w}_2\|$$

**Total #points:** 10 P.