TECHNISCHE UNIVERSITÄT MÜNCHEN
Zentrum Mathematik

Prof. Dr. M. Wolf
J. Cuesta

Mathematical Foundations of Machine Learning

http://www-m5.ma.tum.de/Allgemeines/MA4801_2018S

Summer 2018
Blatt 4
(9.05.2018)

The following homework will be discussed on May 15th and 18th:

H4.1. (Pac Bound from stability) Consider a loss function with range in $[-c, c]$ and any learning algorithm $S \mapsto h_S$ that is uniformly stable with rate $\epsilon_1 : \mathbb{N} \to \mathbb{R}$. Show that the following holds w.r.to repeated sampling of training data sets of size n. For all $\epsilon > 0$ and all probability measures over $\mathcal{X} \times \mathcal{Y}$:

$$\mathbb{P}_S \left[\left| \hat{R}(h_S) - R(h_S) \right| \geq \epsilon + \epsilon_1(n) \right] \leq 2 \exp \left[\frac{n\epsilon^2}{2(n\epsilon_1(n) + c)^2} \right].$$

Hint: Look up McDiarmid's inequality.

H4.2. Let F be a vector space with a scalar product, $C \subset F$ and $b := \sup_{g \in C} \|g\|$. Show that for every point $x \in \text{conv}(C)$ in the convex hull of C and every interger k one can find points $x_1, \ldots, x_k \in C$ such that

$$\left\| x - \frac{1}{k} \sum_{j=1}^{k} x_j \right\|_2^2 \leq \frac{b^2}{k}. \tag{1}$$

Hint: Use the classical Carathéodory's theorem and consider a suitable random choice for the x_j.

H4.3. (Uniform covering number) Let $\mathcal{X} := \{x \in \mathbb{R}^d \,|\, \|x\|_2 \leq 1\}$ and $\mathcal{F} := \{\mathcal{X} \ni x \mapsto \langle x, f \rangle \,|\, \|f\|_2 \leq 1\}$. Consider a sample $(x_1, \ldots, x_n) \in \mathcal{X}^n$ of size n and let us denote by $X \in \mathbb{R}^{n \times d}$ the matrix with rows $x_j, j = 1, \ldots, n$. Show that

$$\log_2 N_{in}(\epsilon, \mathcal{F}, \|\cdot\|_{2,x}) \leq \left\lceil \frac{4 \|X\|_2^2}{n\epsilon^2} \right\rceil \log_2(2n),$$

where $\|X\|_2^2 = \text{Tr}\, X^T X$.

Hint: The singular value decomposition might be useful. Use the previous exercise.