Seminar 1, Statistical Learning Theory

ST510 Foundations of Machine Learning
Lent Term 2022/23

1 PAC learning

Exercise 1.1. Let $\mathcal{X} = \mathbb{R}^d$, $\mathcal{Y} = \{0, 1\}$, $p \in [1, \infty) \cup \{\infty\}$ and let $\mathcal{H}_{\ell_p}$ be the class of balls in ℓ_p-norm, i.e. $\mathcal{H}_{\ell_p} = \{h_r : r \geq 0\}$ where $h_r(x) = \mathbf{1}_{\{||x||_p \leq r\}}$. Prove that $\mathcal{H}_{\ell_p}$ is PAC learnable (assume realizability), and its sample complexity is bounded by

$$m_{\mathcal{H}_{\ell_p}}(\epsilon, \delta) \leq \left\lceil \frac{1}{\epsilon} \log\left(\frac{1}{\delta}\right) \right\rceil .$$

2 Bias-complexity trade-off

Exercise 2.1. We examine the bias-complexity trade-off for predictor functions defined by polynomial functions. That is, $\mathcal{H}_d$ is the set of all functions $h_p : \mathbb{R} \to \{0, 1\}, x \mapsto \mathbf{1}_{\{p(x) \geq 0\}}$ where $p : \mathbb{R} \to \mathbb{R}$ is a polynomial of degree at most d.

Jupyter Notebook: `bias_complexity_tradeoff.ipynb` .

3 Rademacher complexity and growth function

Exercise 3.1. Let $\mathcal{X} = \mathbb{R}^d$, $\mathcal{Y} = \{0, 1\}$, $p \in [1, \infty) \cup \{\infty\}$ and let $\mathcal{H}_{\ell_p}$ be the class of balls in ℓ_p-norm, i.e. $\mathcal{H}_{\ell_p} = \{h_r : r \geq 0\}$ where $h_r(x) = \mathbf{1}_{\{||x||_p \leq r\}}$. Assume 0-1 loss function.

1. Show that the growth function $\Pi_{\mathcal{H}_{\ell_p}}$ satisfies $\Pi_{\mathcal{H}_{\ell_p}}(m) \leq m + 1$ for every integer $m \geq 0$.
2. Prove that for any $h \in \mathcal{H}_{\ell_p}$ and $S = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ i.i.d. examples according to distribution D over $\mathcal{X} \times \mathcal{Y}$, we have for any $\delta \in (0, 1)$, with probability at least $1 - \delta$,

$$L_D(h) \leq L_S(h) + \sqrt{\frac{2 \log(m + 1)}{m}} + 3\sqrt{\frac{\log(2/\delta)}{2m}}.$$

4 VC dimension

Exercise 4.1. What is the VC dimension of $\mathcal{H}_{\ell_p}$, the class of hypotheses we have seen in Exercise 1.1 and Exercise 3.1?

Exercise 4.2. What is the VC dimension of the class of axis-parallel d-dimensional rectangles $\mathcal{H}^d_{\text{rec}}$? (The class of hypotheses $\mathcal{H}^d_{\text{rec}}$ is defined as $\mathcal{H}^d_{\text{rec}} = \left\{h_{a,b} : a \leq b; a, b \in \mathbb{R}^d\right\}$ where $h_{a,b}(x) = \Pi_{i=1}^d \mathbf{1}_{\{x_i \in [a_i, b_i]\}}$. for $x \in \mathbb{R}^d$.)

Exercise 4.3. Revisiting the polynomial hypothesis classes from Section 2: What is the VC dimension of univariate polynomial classifiers of degree d, denoted by $\mathcal{H}_d$?