Sample Problems 4

To be discussed on 07.06.2024

Sample Problem 4.1: Agnostic PAC Learnability

Let $\mathcal{H} \subset \{\pm\}^{\mathcal{X}}$ be a finite hypothesis class. We know that $\mathcal{H}$ is agnostic PAC learnable with respect to 0-1 loss since $\mathrm{VCdim}(\mathcal{H}) \leq \log_2(|\mathcal{H}|) < \infty$.

Now consider the following loss function $\ell^{c_1,c_2} : \{\pm\} \times \{\pm 1\} \to \{0, c_1, c_2\}$ such that

$$\ell^{c_1,c_2}(h(x), y) = \begin{cases} 0 & \text{if } h(x) = y, \\ c_1 & \text{if } h(x) = -1, y = +1, \\ c_2 & \text{if } h(x) = +1, y = -1 \end{cases}$$

ℓ^{c_1,c_2} is called a label-dependent loss, and useful for applications where incorrectly predicting $+1$ is more (or less) harmful than predicting -1, for instance, using face detection to provide access, or in credit scoring models.

Assume $c_1, c_2 < \infty$. Show that $\mathcal{H}$ is agnostic learnable with respect to ℓ^{c_1,c_2}.

Hint: To prove this, you need consider ERM for above loss, and adapt the uniform convergence bound for finite $\mathcal{H}$ to this case.

Hint: Hoeffding's inequality works even when we are not dealing with the 0-1-loss.

Sample Problem 4.2: On-average-replace-one Stability for ERM

Let $\mathcal{H}$ be a hypothesis class with $\mathrm{VCdim}(\mathcal{H}) = d$, let ℓ be the 0-1-loss and let $\mathcal{A}$ be the ERM learner for $\mathcal{H}$. Use uniform convergence results to derive an upper bound β_m for

$$\mathbb{E}_{S\sim\mathcal{D}^m,(x',y')\sim\mathcal{D},i\sim\mathrm{Unif}(m)}\left[\ell\big(\mathcal{A}_{S^i}(x_i), y_i\big) - \ell\big(\mathcal{A}_S(x_i), y_i\big)\right]$$

Show that $\lim_{m\to\infty} \beta_m = 0$.

Sample Problem 4.3: Rademacher Complexity of Sets

The Rademacher complexity of a subset $X \subset \mathbb{R}^m$ is defined as

$$\mathcal{R}_m(X) = \frac{1}{m}\mathbb{E}_\sigma\left[\sup_{x\in X}\langle\sigma, x\rangle\right]$$

where the expectation is with respect to m independent Rademacher random variables $\sigma = (\sigma_1, \ldots, \sigma_m) \in \{\pm 1\}^m$. Furthermore, we define the convex hull of a set X as

$$conv(X) = \left\{ \sum_{i=1}^{N} \lambda_i x_i \middle| x_i \in X, \lambda_i \geq 0, \sum_{i=1}^{N} \lambda_i = 1, N > 0 \right\}$$

Show that $\mathcal{R}_m(X) = \mathcal{R}_m(conv(X))$.