ML Theory — Homework 3

*your NetID here*

Version 1

**Instructions.** (Same as homework 1.)

- Everyone must submit an individual write-up.
- You may discuss with up to 3 other people. State their NetIDs clearly on the first page. Outside of office hours, you should not discuss with anyone but these three.
- Homework is due **Friday, December 22, at 3:30pm**; no late homework accepted.
- Please consider using the provided **\LaTeX** file as a template.
1. Calisthenics.

(a) Let \( k \) real-valued functions \( \mathcal{F} := (f_1, \ldots, f_k) \) be given. Prove

\[
\text{vc}(\text{span}(\mathcal{F})) = \text{vc}\left( \left\{ x \mapsto \text{sgn} \left( b + \sum_i \alpha_i f_i(x) \right) : a \in \mathbb{R}^k, b \in \mathbb{R} \right\} \right) \leq k + 1.
\]

Hint. Use the VC-dimension of linear separators from lecture.

Bonus (ungraded). When is the above an equality?

(b) Let \( \mathcal{F} := \{ x \mapsto \mathbb{1}[\|x - a\|^2_2 \geq b] : a \in \mathbb{R}^d, b \in \mathbb{R} \} \) denote indicators of balls in \( \mathbb{R}^d \). Prove \( \text{vc}(\mathcal{F}) \leq d + 2 \).

Note. Your proof should not get \( d + 1 \) as an upper bound.

(c) Recall from Lecture 23 the ramp loss \( \ell_\gamma \) (where \( \gamma > 0 \)), defined as

\[
\ell_\gamma(r) := \begin{cases} 
0 & x < -\gamma, \\
1 + \frac{1}{\gamma} & x \in [-\gamma, 0], \\
1 & x > 0.
\end{cases}
\]

Prove that for any convex \( \ell : \mathbb{R} \to \mathbb{R} \) with \( \ell(0) > 0 \) and \( \ell'(0) > 0 \),

\[
\ell_\gamma(r) \leq \frac{\ell(r)}{\min \{\ell(0), \ell'(0) \gamma\}}.
\]

(d) Prove the final theorem in Lecture 20, the three-part “core Rademacher theorem”, via the other lemmas and theorems in Lecture 20. (Your main work is in checking the bounded differences condition, and then applying McDiarmid’s inequality to a few other quantities from that lecture.)

Note. The theorem is stated with two typos: the first \( f \) in the first line should be \( \mathcal{F} \), and the right hand side of the third bound should have \( \ln(2/\delta) \) not \( \ln(1/\delta) \).

Solution.

(Your solution here.)
2. Covering non-decreasing functions.

Let $\mathcal{F}$ denote all non-decreasing functions from $\mathbb{R}$ to $[0,1]$. Let a sample $S = (x_1, \ldots, x_n)$ be given, and as usual let $\mathcal{F}|_S \subseteq \mathbb{R}^n$ denote the restriction of $\mathcal{F}$ to the sample $S$.

(a) Prove $\mathcal{N}(\mathcal{F}|_S, \epsilon, \| \cdot \|_2) \leq n^{\sqrt{n}/\epsilon}$.

(b) Using the Pollard bound from Lecture 26, prove

$$\text{Rad}(\mathcal{F}|_S) \leq 16(n \ln(n))^{2/3}.$$ 

(c) Using the Dudley bound from Lecture 26, prove

$$\text{Rad}(\mathcal{F}|_S) \leq 16(n \ln(n))^{1/2}.$$

Solution.

(Your solution here.)
3. Covering linear functions.
Throughout, let $S = (x_1, \ldots, x_n)$ denote a sample of size $n$, and construct matrix $X \in \mathbb{R}^{n \times d}$ with the sample points as rows.

(a) Prove

$$\ln \mathcal{N} \left( \{ x \mapsto \langle x, w \rangle : w \in \Delta_d \} \mid S, \epsilon, \| \cdot \|_2 \} \leq \left\lceil \frac{\| X \|_{2,\infty}^2}{\epsilon^2} \right\rceil \ln(d),$$

where $\Delta_d = \{ \alpha \in \mathbb{R}^d_{\geq 0} : \sum_i \alpha_i = 1 \}$ and $\| X \|_{2,\infty} = \max_i \| X e_i \|_2$.

**Hint.** Use the Maurey Lemma from Lecture 15.

(b) Prove

$$\ln \mathcal{N} \left( \{ x \mapsto \langle x, w \rangle : \| w \|_1 \leq a \} \mid S, \epsilon, \| \cdot \|_2 \} \leq \left\lceil \frac{a^2 \| X \|_{2,\infty}^2}{\epsilon^2} \right\rceil \ln(2d).$$

**Hint.** Use the previous part.

(c) Define

$$\mathcal{F}_2(a) := \{ x \mapsto \langle x, w \rangle : \| w \|_2 \leq a \}.$$

Prove

$$\ln \mathcal{N} \left( \mathcal{F}_2(a) \mid S, \epsilon, \| \cdot \|_2 \} \leq \left\lceil \frac{a^2 \| X \|_2^2}{\epsilon^2} \right\rceil \ln(2d),$$

where $\| X \|_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^d (x_{ij})^2}$ denotes the Frobenius norm.

**Hint.** Use the previous part.

(d) Use the Pollard bound from Lecture 26 to prove

$$\operatorname{Rad}(\mathcal{F}_2(a) \mid S) \leq 16a \| X \|_2 \left( n \ln(2d) \right)^{1/4}.$$

(e) Use the Dudley bound from Lecture 26 to prove

$$\operatorname{Rad}(\mathcal{F}_2(a) \mid S) \leq 16a \| X \|_2 \left( \ln(n) - \ln(a \| X \|_2) \right) \ln(2d).$$

**Remark.** The direct Rademacher proof gave $\operatorname{Rad}(\mathcal{F}_2(a) \mid S) \leq a \| X \|_2$.

**Solution.**

*(Your solution here.)*
4. Are we still friends?

Solution.

(Your solution here.)