STAT260 Problem Set 3

Due October 24th in class

Regular problems:

1. Prove the following invariance properties of Poincaré distributions:
   - If $X \sim p$ is σ-Poincaré and $f : \mathbb{R}^n \to \mathbb{R}^m$ is L-Lipschitz, then $f(X)$ is $(L\sigma)$-Poincaré.
   - If $X \in \mathbb{R}^n$, $Y \in \mathbb{R}^m$ are independent and both σ-Poincaré, then the ordered pair $(X, Y) \in \mathbb{R}^{n+m}$ is σ-Poincaré.

2. For a distribution p on $\mathbb{R}^d$, let $p_k(v) = \mathbb{E}_{x\sim p}[\langle x, v\rangle^k]$ and observe that p_k is a polynomial in v. Give sum-of-squares proofs for the following facts:
   - If $1 \le a \le b$ are both odd then $p_a(v)p_b(v) \preceq_{\text{sos}} p_{a-1}(v)p_{b+1}(v)$.
   - If $2 \le a \le b$ are both even then $p_a(v)p_b(v) \preceq_{\text{sos}} p_{a-2}(v)p_{b+2}(v)$.

   [Hint: Start by showing that for every even r, $x_r - x^{r-1}y - xy^{r-1} + y^r \succeq_{\text{sos}} 0$ (as a polynomial in x and y), and similarly $x^r - x^{r-2}y^2 - x^2y^{r-2} + y^r \succeq_{\text{sos}} 0$. This may require some clever factorization/algebra.]

3. In class we defined sets $\mathcal{G}^{\downarrow}$ and $\mathcal{G}^{\uparrow}$ for an arbitrary loss $L(p, \theta)$. Here we consider the following generalized construction, that also incorporates a *bridge function* $B(p, \theta)$:

$$\mathcal{G}^{\downarrow}(\rho_1, \epsilon) \stackrel{\text{def}}{=} \{p \mid B(r, \theta^*(p)) \le \rho_1 \text{ whenever } r \le \frac{p}{1-\epsilon}\}, \tag{1}$$

$$\mathcal{G}^{\uparrow}(\rho_1, \rho_2, \epsilon) \stackrel{\text{def}}{=} \{p \mid L(p, \theta) \le \rho_2 \text{ whenever } B(r, \theta) \le \rho_1 \text{ for some } r \le \frac{p}{1-\epsilon}, \theta\}. \tag{2}$$

   Show that the modulus of continuity of $\mathcal{G}^{\downarrow}(\rho_1, \epsilon) \cap \mathcal{G}^{\uparrow}(\rho_1, \rho_2, \epsilon)$ under L is at most ρ_2.

4. Here we bound the modulus of continuity for linear classification. Given $(x, y) \sim p$ where $x \in \mathbb{R}^d$ and $y \in \{\pm 1\}$, Define $L(p, \theta) = \mathbb{P}_{(x,y)\sim p}[y \ne \text{sign}(\langle x, \theta)]$ and $B(p, \theta) = \mathbb{E}_{(x,y)\sim p}[\max(0, 1 - y\langle x, \theta\rangle)]$, and let $\theta^*(p)$ be the minimizer of $B(p, \theta)$. Suppose that p satisfies the following two properties:
   - $\mathbb{E}_{(x,y)\sim p}[\max(0, 1 - y\langle x, \theta^*(p)\rangle)] \le (1 - \epsilon)\rho_1$
   - Whenever $\mathbb{P}_{(x,y)\sim p}[y\langle x, \theta\rangle \le \frac{1}{2}] \le \epsilon + 2(1-\epsilon)\rho_1$ for some θ, then $\mathbb{P}_{(x,y)\sim p}[y\langle x, \theta\rangle \le 0] \le \rho_2$.

   Show that $p \in \mathcal{G}^{\downarrow}(\rho_1, \epsilon) \cap \mathcal{G}^{\uparrow}(\rho_2, \epsilon)$.

   [Remark: Note that the second condition is a type of tail bound, where in every direction where we are somewhat unlikely to be close to the boundary, we are very unlikely to cross the boundary entirely.]

5. Define the norm $\|x\|_{\mathcal{S}_k} = \max_{\|v\|_2 \le 1, \|v\|_0 \le k} \langle x, v\rangle$.
   (a) What is the dual norm to $\|\cdot\|_{\mathcal{S}_k}$?
   (b) Show that if x and y both have at most k non-zero entries, then $\|x - y\|_{\mathcal{S}_{2k}} = \|x - y\|_2$.
   (c) Show that any distribution that is (ρ, ϵ)-resilient in the ℓ_2-norm is also (ρ, ϵ)-resilient in the $\mathcal{S}_k$-norm.

Challenge problems (turn in as a separate document typeset in LaTeX):

6. Construct a distribution that is $(\sqrt{\epsilon}, \epsilon)$-resilient in the $\mathcal{S}_k$-norm for all $\epsilon < 1/4$, but not $(\rho, 1/10)$-resilient in the ℓ_2-norm for any $\rho < \Omega(k^{0.1})$.

   [The constants $1/4$*,* $1/10$*,* 0.1 *are all arbitrarily chosen, the point is to show a polynomial separation between* $\mathcal{S}_k$ *and* ℓ_2 *for some distribution. Note that your construction will likely need to have* d/k *going to* ∞ *as* $k \to \infty$*.]*

7. Suppose that $X \sim p$ is σ-Poincaré and f is L-Lipschitz. Show that $f(X)$ is sub-exponential with parameter $\mathcal{O}(\sigma L)$.

   [Hint: Define $u(\lambda) = \mathbb{E}[e^{\lambda f(x)}]$ *and compare to* $\mathsf{Var}[e^{\lambda f(x)/2}]$*.]*