10/21/24

Lecture 13: Statistical Query Model (Basics)

Parity w/ noise (LPN):

- Unknown $S \subseteq [n]$
- Given examples $(x_1, y_1), \ldots, (x_N, y_N)$ s.t.

$$x_i \sim \mathrm{Unif}(\{\pm 1\}^d)$$
$$y_i = \begin{cases} x_S & \text{w.p. } 1-\eta \\ -x_S & \text{o.w.} \end{cases}$$

Thm [Kearns et al. '98]: Even when $\eta = 0$, any SQ algorithm for LPN requires tolerance $2^{-\Omega(d)}$ or must make $2^{\Omega(d)}$ queries.

Pf: Consider any CSQ $\phi : \{\pm 1\}^d \to [-1, 1]$

Define $\phi_S \triangleq \mathbb{E}_x [x_S \cdot \phi(x)]$.

Claim: For uniformly random S,

$$\mathrm{Var}_S[\phi_S] \leq 2^{-\Omega(n)}$$

Pf: $\mathrm{Var}_S[\phi_S]$: $\quad$ 1, -1

$$\mathbb{E}_S[\phi_S^2] - \mathbb{E}_S[\phi_S]^2$$

$$= \mathbb{E}_S \mathbb{E}_{x,x'}[x_S x'_S \, \phi(x)\phi(x')]$$

$$- \mathbb{E}_{S,S'} \mathbb{E}_{x,x'}[x_S x'_{S'} \, \phi(x)\phi(x')]$$

$$= \mathbb{E}_{x,x'}\Big[\phi(x)\phi(x') \underbrace{\mathbb{E}_{S,S'}[x_S x'_S - x_S x'_{S'}]}\Big]$$

Claim: $= 0$ if $x \neq x'$

Pf: For $z \neq \vec{1}$, $\mathbb{E}_S[z_S] = 0$

So if $x \neq x'$, $\mathbb{E}_S[x_S x'_S] = \mathbb{E}[z_S] = 0$

and $\mathbb{E}_{S,S'}[x_S x'_{S'}] = \mathbb{E}_S[x_S]\,\mathbb{E}_{S'}[x'_{S'}] = 0$

b/c at most one of $x, x' = \vec{1}$.

"parity function is pairwise-independent hash function."

$$= \mathbb{E}_{x,x'}\left[\phi(x)\,\phi(x') \cdot \mathbb{1}[x = x'] \right]$$

$$= \frac{1}{2^n} \mathbb{E}_x \left[\phi(x)^2 \right] \leq \frac{1}{2^n}.$$

So by Chebyshev's,

$$\Pr_S \left[\left| \phi_S - \mathbb{E}_S[\phi_S] \right| \geq \tau \right] \leq \frac{\mathrm{Var}_S(\phi_S)}{\tau^2}$$

$$\leq \frac{1}{2^n \tau^2}.$$

i.e. to answer CSQ ϕ, just output $\mathbb{E}_S[\phi_S]$. If tolerance $= \tau$, this is accurate for $\frac{1}{2^n \tau^2}$ fraction of parity functions. i.e. each CSQ only rules out at most $\frac{1}{\tau^2}$ many S's, so we need $2^n \tau^2$ queries. Taking $\tau = 2^{-n/3}$ completes the proof. $\square$

SQ dimension (recipe for supervised problems)

Def: Class of functions $\mathcal{F} = \{f : \mathbb{R} \to [-1,1]\}$ has SQ dimension $\geq D$ w.r.t q if $\exists$ $f_1, \ldots, f_D \in \mathcal{F}$ s.t. $\forall$ $i \neq j$

$$\left| \mathbb{E}_{x \sim q}\left[f_i(x) \, f_j(x) \right] \right| \leq \frac{1}{D}$$

Thm: If $\mathcal{F}$ has SQ dimension D, then CSQ learning requires tolerance τ or $\Omega(D\tau^2)$ queries

Pf: For convenience, define $\langle f, g \rangle \triangleq \mathbb{E}[f(x) g(x)]$.

WLOG let $\mathcal{F} = \{f_1, \ldots, f_D\}$.

Given query $\phi : \mathbb{R}^d \to [-1, 1]$, let

$$A^+ \triangleq \{ f \in \mathcal{F} : \langle f, \phi \rangle \geq \tau \}$$

By Cauchy-Schwartz:

$$\left\langle \phi, \sum_{f \in A^+} f \right\rangle^2 \leq \underbrace{\|\phi\|^2}_{\leq 1} \cdot \left\| \sum_{f \in A^+} f \right\|^2$$

$$\leq \sum_{f \in A^+} \|f\|^2 + \frac{|A^+| \cdot (|A^+| - 1)}{D}$$

$$\leq \frac{|A^+|^2}{D} + |A^+|$$

But $\langle \phi, \sum_{f \in A^+} f \rangle \geq \tau |A^+|$ by defn., so

$$\tau^2 |A^+|^2 \leq \frac{|A^+|^2}{D} + |A^+|$$

$$\Rightarrow |A^+| \leq \frac{D}{D\tau^2 - 1} \leq O\left(\frac{1}{\tau^2}\right)$$

Similarly, for $A^- \triangleq \{ f \in \mathcal{F} : \langle f, \phi \rangle \leq -\tau \}$, $|A^-| \leq O\left(\frac{1}{\tau^2}\right)$.

So regardless of ϕ, all but $O\left(\frac{1}{\tau^2}\right)$ many f's consistent w/ the answer 0, so need $\Omega(D\tau^2)$ queries. $\square$

: