Last time: 2) Amazing Theorem (combinatorics) about $\Pi_{\mathcal{C}}(m)$, for any $\mathcal{C}$:

Amazing Theorem: Fix any $\mathcal{C}$. Let $d = \text{VCDIM}(\mathcal{C})$. Have

- $\Pi_{\mathcal{C}}(m) = 2^m$ for $m \le d$

(!) • $\Pi_{\mathcal{C}}(m) \le \binom{m}{0} + \binom{m}{1} + \dots + \binom{m}{d} \le \left(\frac{em}{d}\right)^d$ for $m > d$

And started 3) Learning-style argument: our new CHF thm.

Today: • finish (+ whole proof of sample cxity u.b. based on VCDIM) (and prove the inequality)

- Application: PAC learning LTFs over $\mathbb{R}^n$ via linear programming
- Start unit on confidence + accuracy boosting

Questions?

Admin: ps3 due Tues; Rocco OH ~~this week~~ today 4-6pm, CSB 480

Recall our new CHF thm:

Key to pf: union bd over possible labelings of a 2m-elt set.

New Thm Fix any ~~finite~~ $\mathcal{C}$, any dist $\mathcal{D}$, any target concept $c \in \mathcal{C}$. Given data set m examples drawn from $EX(c, \mathcal{D})$, where

(★) $\boxed{m \ge \frac{2}{\varepsilon}\left(\log\left(\Pi_{\mathcal{C}}(2m)\right) + \log\frac{2}{\delta}\right)}$, $m \ge 8/\varepsilon$,

w.p. $\ge 1-\delta$, all bad $h \in \mathcal{C}$ are inconsistent with data set.

$\star$ is great, b/c of AT: AT says

$$\log\left(\Pi_e(2m)\right) \le \underline{\log\left(\left(\frac{2em}{d}\right)^d\right)}, \text{ so}$$

$\star$ becomes

$$\boxed{m \geq \underbrace{\frac{2}{\varepsilon}\, d \log\left(\frac{2em}{d}\right)}_{A} + \underbrace{\frac{2}{\varepsilon} \log \frac{2}{\delta}}_{B}} \quad \star\star$$

$\star\star$ is sat. provided

K = const. like 8

$$\boxed{m \geq \underline{K} \cdot \left(\frac{d}{\varepsilon} \log \frac{1}{\varepsilon} + \underline{\frac{1}{\varepsilon} \log \frac{1}{\delta}}\right)}$$

$\star\star \Rightarrow \star$ sketch:

to show $m \geq A + B$, enough to show $m \geq 2A$, $m \geq 2B$

$m \geq 2B$: easy

$m \geq 2A$: $\frac{m}{d} \geq \frac{2}{\varepsilon} \log\left(2e\frac{m}{d}\right)$

$C = \frac{2}{\varepsilon}$ $\quad m' = \frac{m}{d}$: $\boxed{m' \geq C \cdot \log(m')}$

we know $m' \geq 2C \log C$ suffices,

i.e $\frac{m}{d} \geq 2 \cdot \left(\frac{2}{\varepsilon}\right) \log\left(\frac{2}{\varepsilon}\right)$ suffices.

So, remains to prove **New Thm**.

Pf: Consider following prob. experiment:

- Draw $2m$ samples from $EX(c, \mathcal{D})$.

  Call 1st m S_1, next m S_2.

Consider events: $(err_{\mathcal{D}}[h,c] > \varepsilon)$

- A: some bad $h \in \mathcal{C}$ gets all of S_1 right.

  To prove thm: must show $\Pr[A] \le \delta$.

- B: some $h \in \mathcal{C}$ gets all of S_1 right and is wrong on $\ge \frac{\varepsilon}{2} \cdot m$ elts of S_2.

Claim: For $m \ge \frac{8}{\varepsilon}$, $\Pr[A] \le 2\Pr[B]$.

Pf:
$$\Pr[B] \ge \Pr[A+B] = \Pr[A] \cdot \Pr[B \mid A].$$

To show: $\Pr[B \mid A] \ge \frac{1}{2}$.

Is $\Pr[B \mid A] \ge \frac{1}{2}$? Yes: since A holds, some bad hyp in $\mathcal{C}$ (call it h^*) gets S_1 right. $\mathbb{E}[\#\text{ mist of } h^* \text{ on } S_2] \ge \varepsilon m$; by mult. CB $(\gamma = \frac{1}{2})$,

$$\Pr[h^* \text{ makes fewer than } \tfrac{\varepsilon}{2} m \text{ mist. on } S_2] \le e^{-\varepsilon m \gamma^2/2}$$
$$= e^{-m\varepsilon/8}.$$

$m = 8/\varepsilon$ makes this $< \frac{1}{2}$!!.

To finish pf of New Thm, suff to show $\Pr[B] \le \frac{\delta}{2}$.

Idea: upper bound $\Pr[B]$ by

- considering a fixed lab. of pts in $S_1 + S_2$ (achievable by $\mathcal{C}$)

  by considering randomness in div. of S into $S_1 +$

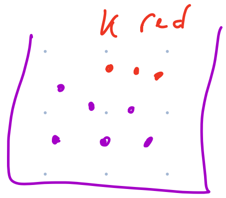
S_2, getting a good u.b. for <u>that</u> labeling;
- doing <u>union bd</u> over all labelings in $\Pi_e(S)$ (control # via $\Pi_e(m)$).

- Equiv. way to view B: Draw S $2m$ ex. <u>Randomly</u> split into S_1, S_2 each of size m.
  B: there's some $h \in \mathcal{C}$ <u>cons. with S_1</u> but wrong on $\geq \frac{\varepsilon}{2} m$ in S_2.

- Fix a lab. of the ($h \neq c$) $2m$ pts in S that's in $\Pi_e(S)$.
  Can suppose it mislabels <u>$k \geq \frac{\varepsilon}{2} m$</u> of the $2m$ pts.

Barrel
$2m$ balls

k red



Draw rand. subset (S_2) of $2m$ of balls out of barrel.

Pr[all k red balls are among the m drawn for S_2]

$$= \binom{m}{k} \Big/ \binom{2m}{k} = \frac{m!}{k!(m-k)!} \Big/ \frac{(2m)!}{k!(2m-k)!}$$

$$= \frac{m(m-1)\cdots(m-k+1)}{2m(2m-1)\cdots(2m-k+1)} \leq \frac{1}{2^k}$$

$1/2$ $\quad < \frac{1}{2}$

So for this <u>fixed</u> lab. of S

$$\Pr_{\substack{\text{rand.} \\ \text{split} \\ \text{of } S \\ \text{into } S_1, S_2}}\left[\text{all } k \geq \tfrac{\varepsilon}{2} m \text{ mist. are in } S_2\right] \leq \frac{1}{2^k} \leq \frac{1}{2^{\frac{\varepsilon m}{2}}}.$$

UB over all poss lab. of S ($\leq \Pi_e(2m)$ many):

$$\Pr[B] \leq \underbrace{\Pi_e(2m) \cdot \frac{1}{2^{\frac{\varepsilon m}{2}}}}_{\leq \frac{\delta}{2},} .$$

If $\leq \frac{\delta}{2}$, done.

$$\Pi_e(2m) \cdot \frac{1}{2^{\frac{\varepsilon m}{2}}} \leq \frac{\delta}{2} \quad \text{means}$$

$$\log \Pi_e(2m) - \frac{\varepsilon m}{2} \leq \log \frac{\delta}{2} \quad ; \quad \text{rearrange + get}$$

$$\boxed{\frac{2}{\varepsilon}\left(\log \Pi_e(2m) + \log \frac{2}{\delta}\right) \leq m}$$

, giving (*): done!

IOU: $\binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{d}$ is $\begin{cases} = 2^m & \text{for } m \leq d \\ \leq \left(\frac{e\,m}{d}\right)^d & \text{for } m > d. \end{cases}$

Pf: if $m \leq d$, $\binom{m}{0}+\binom{m}{1}+\cdots+\binom{m}{d} = \underbrace{\binom{m}{0} + \cdots + \binom{m}{m}}_{\#\text{subsets of } \{1,\dots,m\} = 2^m} + \underbrace{\binom{m}{m+1} + \cdots \binom{m}{d}}_{0}$

If $m > d$: have $\left(\frac{d}{m} < 1\right)$

$$\left(\frac{d}{m}\right)^d \cdot \left[\binom{m}{0}+\binom{m}{1}+\cdots+\binom{m}{d}\right]$$

$$\leq \sum_{i=0}^{d} \left(\frac{d}{m}\right)^i \binom{m}{i} \qquad \text{b/c } \frac{d}{m}<1$$

$$\leq \sum_{i=0}^{m} \left(\frac{d}{m}\right)^i \binom{m}{i} \cdot 1^{m-i}$$

$$= \left(1+\frac{d}{m}\right)^m \qquad \text{binom. thm: } (a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}, \quad a=\frac{d}{m}, \; m=n, \; b=1$$

$$\leq \left(e^{d/m}\right)^m \qquad (1+x) \leq e^x$$

$$= e^d.$$

Rearrange: $\left(\frac{d}{m}\right)^d \cdot \left[\binom{m}{0}+\binom{m}{1}+\cdots+\binom{m}{d}\right] \leq e^d$

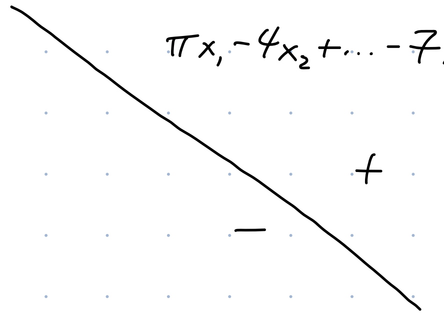
$$\Downarrow$$

$$\binom{m}{0}+\binom{m}{1}+\cdots+\binom{m}{d} \leq e^d \left(\frac{m}{d}\right)^d = \left(\frac{em}{d}\right)^d.$$

Done with UB on sample cxity of PAC learning.

Applic:

$X = \mathbb{R}^n$, $\mathcal{C}$ = all LTFs, e.g.



$\pi x_1 - 4x_2 + \dots - 7x_n \geq 34$.

Sad fact: given $m = n$ ex. over $\{0,1\}^n$ ($\ell = 1$), there are LTFs s.t. "the Perc. alg. δ-param" is $\leq \frac{1}{2^n}$

Fact 1: $VCDIM(\mathcal{C}) = n+1$.

Fact 2: There's a poly-time CHF for $\mathcal{C} = LTFs$ using $\mathcal{H} = LTFs$.

alg which, given m many n-dim ex. with ℓ bits of precision in each coord, runs in poly(m, n, ℓ) time

Let $w_1x_1 + \dots + w_nx_n \geq \theta$ is the unknown target LTF.

Say have pos. ex. $(8, 2, -3, \dots, 7)$; $+$

$\Downarrow$ $\boxed{8w_1 + 2w_2 - 3w_3 \dots + 7w_n \geq \theta}$

Neg ex $(4, 3, \dots 5)$, $-$: $\boxed{4w_1 + 3w_2 \dots + 5w_n < \theta}$

LP Feasibility of Linear Prog: given linear ineq's determine if $\exists$ sol & if so find it.

We have poly time alg's for this :)

Putting pieces together, here's a poly-time PAC learner for LTFs:

- Draw $m = \frac{8}{\varepsilon}(n+1)\log\frac{1}{\varepsilon} + \frac{8}{\varepsilon}\log\frac{1}{\delta}$

ex from $EX(c, \mathcal{D})$;

- Run poly-time <u>LP</u> to find (v, θ') labeling them all correctly;

- Output hyp $\mathbb{1}(v \cdot x \geq \theta)$.

Next time: accuracy + confidence boosting
