Last time:

- finish pf that properly learning 3-term DNF is computat. hard, by completing pf of (2) about the transformation f
- start VC dimension unit (KV Chap. 3) / unit on:

sample cxity of PAC learning

- lower bound on #: $\Omega(\mathrm{VCDIM}(\mathcal{C})/\varepsilon)$
- start upper bound on #: (roughly) $O(\mathrm{VCDIM}(\mathcal{C})/\varepsilon)$
  - setup
  - combinatorics arg.
  - learning arg.

Today:

we'll finish

① Setup: Ponder $\mathrm{VCDIM}(\mathcal{C})$. (a #)
We'll consider a function "growth function of $\mathcal{C}$" $\equiv \Pi_{\mathcal{C}}(m)$
to get more info.

+ we'll do

② Combinatorics arg: AMAZING THEOREM about how the growth fn behaves, for any $\mathcal{C}$.

③ Learning argument: a lot like CHF thm, but using AMAZING THM to control # ways conc in $\mathcal{C}$ can label a data set.

Questions?

S = a set of examples

Recall • $\Pi_{\mathcal{C}}(S) = \{c \cap S : c \in \mathcal{C}\}$ = a collection of sets
"shatter function"

m = a #

• $\Pi_{\mathcal{C}}(m) = \max\limits_{|S|=m} |\Pi_{\mathcal{C}}(S)|$ = a number
"growth func"

So $\mathrm{VCDIM}(\mathcal{C})$ = max value m st $\Pi_{\mathcal{C}}(m) = 2^m$.

Ex: $\mathcal{C}$ = intervals of $\mathbb{R}$ VCDIM=2

$|S| = m$: • • • • • • •

$\Pi_{\mathcal{C}}(1) = 2$ [ ] [ ]

$\Pi_{\mathcal{C}}(2) = 4$ • • all 4 ways

$\Pi_{\mathcal{C}}(3) = 7 < 8$ + − +

$\Pi_{\mathcal{C}}(m) = \binom{m}{0}$ all −

$+ \binom{m}{1}$ one +

$+ \binom{m}{2}$ left + right +

AMAZING THM (VC Theorem Vap. + Cherv. Perles Sauer/Shelah)

Let $\mathcal{C}$ be a concept class w/ VCDIM$(\mathcal{C}) = d$.
Then

- if $m \le d$, $\Pi_{\mathcal{C}}(m) = 2^m$; ✓ (expon.)
- if $m > d$, $\Pi_{\mathcal{C}}(m) \le \binom{m}{0} + \binom{m}{1} + \dots + \binom{m}{d}$ ✪

  $\le \left(\frac{em}{d}\right)^d$ ✪ (bin. coeff. manip's)

For now, forget $\mathcal{C}$.

Def: $\Phi_d(m)$, for $d, m \ge 0$, is:

- $\Phi_0(m) = \Phi_d(0) = 1$ $\forall\, d, m \ge 0$;

- For $d, m \geq 1$: $\Phi_d(m) = \Phi_d(m-1) + \Phi_{d-1}(m-1)$.

$\Phi_d(m)$

d \ m	0	1	2	3	4	
	1	2	4	8		$+$
4	1	2	4	8	16	
3	1	2	4	8	15	
2	1	2	4	7	11	
1	1	2	3	4	5	
$d =$ 0	1	1	1	1	1	1 ...

$m =$

Fact: $\boxed{\Phi_d(m) = \sum_{i=0}^{d} \binom{m}{i}.}$

Pf: Induc. on $m + d$.

Base cases: $\Phi_d(0) \overset{(\text{def})}{=} 1 \overset{?\checkmark}{=} \sum_{i=0}^{d} \binom{0}{i} = \binom{0}{0} = 1$:)

$\Phi_0(m) \overset{(\text{def})}{=} 1 \overset{?}{=} \sum_{i=0}^{0} \binom{m}{i} = \binom{m}{0} = 1$:)

Induc. step:

$$\Phi_d(m) \overset{(\text{def})}{=} \Phi_d(m-1) + \Phi_{d-1}(m-1)$$

$$= \sum_{i=0}^{d} \binom{m-1}{i} + \sum_{i=0}^{d-1} \binom{m-1}{i} \qquad (\text{ind. hyp.})$$

$$= \sum_{i=0}^{d} \underline{\binom{m-1}{i} + \binom{m-1}{i-1}} \qquad \left(\text{b/c } \binom{m-1}{-1} = 0\right)$$

$$= \sum_{i=0}^{d} \underline{\binom{m}{i}} \qquad \blacksquare$$

b/c $\boxed{\binom{m}{i} = \binom{m-1}{i-1} + \binom{m-1}{i}}$

if include item m, there are $\binom{m-1}{i-1}$ ways to complete selection;

if don't include item m, there are $\binom{m-1}{i}$ ways to complete sel.

Fact: $\Phi_d(m) = \sum_{i=0}^{d} \binom{m}{i}$ is $\begin{cases} = 2^m & \text{if } m \le d \\ \le \left(\frac{em}{d}\right)^d & \text{if } \underline{\underline{m > d}}. \end{cases}$

Pf: if $m \le d$,

$$= \underbrace{\binom{m}{0} + \binom{m}{1} + \dots + \binom{m}{m}}_{= 2^m = \#\text{ subsets of } [m]} + \underbrace{\binom{m}{m+1} + \dots + \binom{m}{d}}_{\text{all } 0}$$

$$\binom{m}{0} + \dots + \binom{m}{m} = \sum_{i=0}^{m} \#\text{ subsets of } [m] \text{ of size exactly } i$$

$$= 2^m =$$

If $m > d$: consider $\left(\frac{d}{m}\right)^d \cdot \sum_{i=0}^{d} \binom{m}{i}$: (note $d/m < 1$)

$$\left(\frac{d}{m}\right)^d \cdot \sum_{i=0}^{d} \binom{m}{i} \le \sum_{i=0}^{d} \left(\frac{d}{m}\right)^i \cdot \binom{m}{i}$$ (smaller "penalty" on RHS)

$$\le \sum_{i=0}^{m} \left(\frac{d}{m}\right)^i \binom{m}{i} \cdot 1^{m-i}$$ (added more stuff in)

$$= \left(1 + \frac{d}{m}\right)^m$$ (bin. thm: $(a+b)^n = a^n + a^{n-1}b\binom{n}{1} + \dots$)

$$\le \left(e^{d/m}\right)^m$$ (b/c $1 + x \le e^x$)

$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$

$$= e^d.$$ So

$$\left(\frac{d}{m}\right)^d \sum_{i=0}^{d} \binom{m}{i} \le e^d, \text{ hence } \sum_{i=0}^{d} \binom{m}{i} \le \left(\frac{em}{d}\right)^d. \blacksquare$$

Back to $\mathcal{C}$: Let $\mathcal{C}$ be any conc. class, $d = \mathrm{VCDIM}(\mathcal{C})$.

Thm: $\forall m$, have $\Pi_{\mathcal{C}}(m) \le \Phi_d(m)$.

$\Phi_{d-1}(m-1) + \Phi_d(m-1)$

Pf: induc. on m, d.

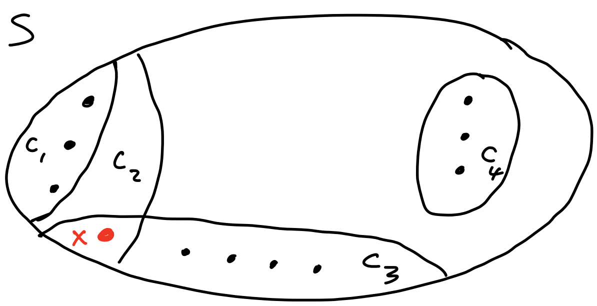
Base case: $m = 0$. $\Pi_{\mathcal{C}}(0) = 1$ (no labels to any ex) $\checkmark = \Phi_d(0)$

$d = 0$. $\Pi_{\mathcal{C}}(m) = ?$ $\mathcal{C}$ has $\mathrm{VCDIM} = 0$:

must have $|\mathcal{E}| \leq 1$, so $\Pi_{\mathcal{E}}(m) \leq 1 \overset{\checkmark}{=} \Phi_{e}(0)$

Induc. step: assume it's true for all m', d' st $m' \leq m$, $d' \leq d$, $m'+d' < m+d$.

Let S be any subset of X with $|S| = m$.
Goal: bound $|\Pi_{\mathcal{E}}(S)|$ by $\Phi_d(m)$.



$x =$ special pt in S.

Claim: $\Pi_{\mathcal{E}}(\overbrace{S-\{x\}}^{m-1}) \leq \Phi_d(m-1)$ by ind. hyp.

Have $\Pi_{\mathcal{E}}(S-\{x\})$ undercounts $\Pi_{\mathcal{E}}(S)$ b/c of pairs of distinct sets in $\Pi_{\mathcal{E}}(S)$ which differ only on x. e.g. c_1 & c_c

"sets like c_1": $\mathcal{E}'$ is a coll of subsets of $S-\{x\}$.

Define $\mathcal{E}' := \{c \in \Pi_{\mathcal{E}}(S) : x \notin c, \ c \cup \{x\} \in \Pi_{\mathcal{E}}(S)\}$

So $|\Pi_{\mathcal{E}}(S)| = |\mathcal{E}'| + \overbrace{|\Pi_{\mathcal{E}}(S-\{x\})|}^{\leq \Phi_d(m-1)}$

Notice: $\mathcal{E}' = \Pi_{\mathcal{E}'}(S-\{x\})$, because each set in $\mathcal{E}'$ is a subset of $S-\{x\}$, &

$\Pi_{e'}(S-\{x\})$ is defined to be $\{ c \cap (S-\{x\}) : c$ ranges over all of $e' \}$

So $|\Pi_e(S)| = |e'| + |\Pi_e(S-\{x\})|$ $\qquad \Phi_d(m) = \Phi_{d-1}(m-1) + \Phi_d(m-1)$

$$= |\Pi_{e'}(S-\{x\})| + \underbrace{|\Pi_e(S-\{x\})|}$$

VCDIM$(e)=d$,
$|S-\{x\}| = m-1$: so
$|\Pi_e(S-\{x\})| \leq \Phi_d(m-1)$.

If have VCDIM$(e') \leq d-1$, then by IH have

$|\Pi_{e'}(S-\{x\})| \leq \Phi_{d-1}(m-1)$, & done!

What's VCDIM(e')?

Sps $S' \subseteq S-\{x\}$, S' is shatt. by e'. (Let S' = largest such set.)

Then $S' \cup \{x\}$ is shattered by e. (twin)

Know VCDIM$(e) = d$, so $|S' \cup \{x\}| \leq d$,

so $|S'| \leq d-1$.

So VCDIM$(e') \leq d-1$. ∎