

PS2 due tonight

- We will send Peer Evaluation assignments via Courseworks "mail"
- Print out on paper & write feedback by hand (in pen)
- Due 10/27 in lecture

PS3 due 10/29

Midterm 11/5

- Check out the exercises in the lecture notes

Last time:

Uniform convergence for finite hypothesis classes

"Behaviors" of $FS \{0, 1\}^Z$ on finite set

"Shattering"

VC dimension

Today: Sauer's lemma

Recall:

Let $\mathcal{F} \subseteq \{0,1\}^{\mathbb{Z}}$ be a class of $\{0,1\}$ -valued functions on domain $\mathbb{Z}$

For a set of n points $T = \{z_1, \dots, z_n\} \subseteq \mathbb{Z}$ the set of behaviors of $\mathcal{F}$ on T is

$$\mathcal{F}[T] = \{ (f(z_1), \dots, f(z_n)) \mid f \in \mathcal{F} \} \subseteq \{0,1\}^n$$

$\mathcal{F}$ shatters T if $\mathcal{F}[T] = \{0,1\}^{|T|}$ (all behaviors)

VC dimension of $\mathcal{F}$: largest n s.t. there is a set $T \subseteq \mathbb{Z}$ with $|T|=n$ that is shattered by $\mathcal{F}$.
(or ∞ if arbitrarily large sets shattered by $\mathcal{F}$)

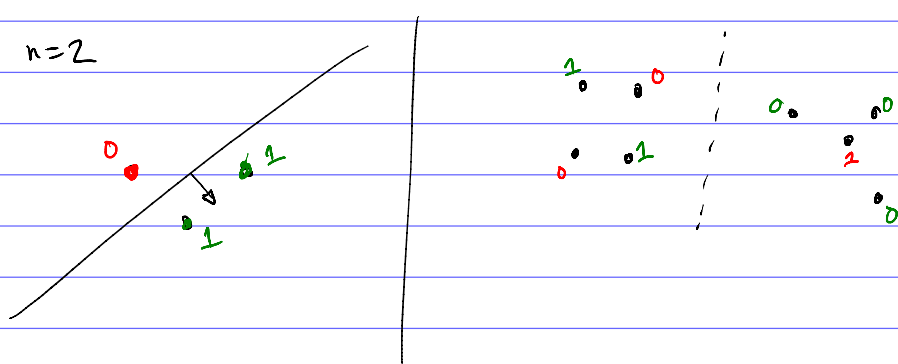
Ex $\mathcal{F} = \{f_V \mid V \subseteq [0,1], |V| < \infty\} \quad \mathbb{Z} = [0,1]$
 $f_V(x) = 1 \text{ if } x \in V$
 $\text{vc dim}(\mathcal{F}) = \infty$

Pick any $z_1, \dots, z_{1000} \in [0,1]$

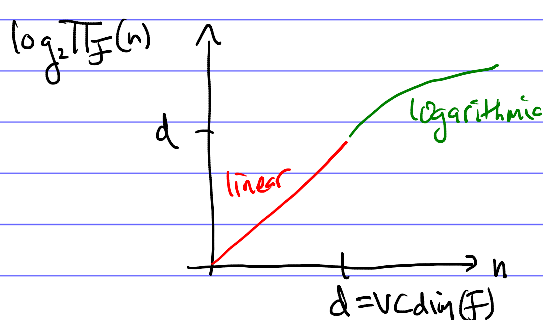
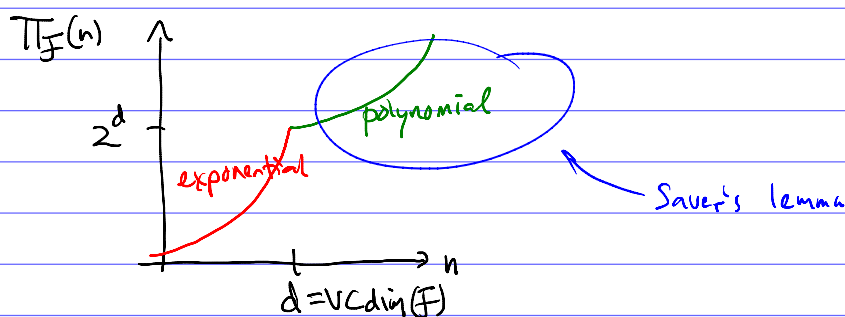
$\underbrace{1 \dots 1 \ 0 \dots 0}_{\text{put } z_1, \dots, z_{500} \text{ in } V}$

$\mathcal{F} = \text{LTF}_{\mathbb{R}^n}$

$\text{vc dim}(\text{LTF}_{\mathbb{R}^n}) = n+1$



Growth function $\Pi_{\mathcal{F}}(n) = \max \{ |\mathcal{F}[T]| \mid T \subseteq \mathbb{Z}, |T|=n \}$



$$\sqrt{\frac{\log(|\mathcal{F}|/\delta)}{n}}$$

if $\log_2 |\mathcal{F}[T]| = O(\log n)$

Sauer's lemma $\Pi_{\mathcal{F}}(n) \leq \binom{n}{\leq d} := \binom{n}{0} + \binom{n}{1} + \dots + \binom{n}{d}$
where $d = \text{VCdim}(\mathcal{F})$

if $n \geq d$, then $\binom{n}{\leq d} \leq \left(\frac{ne}{d}\right)^d$

For any $T \subseteq \mathcal{Z}$, $|T| = n$

$$\mathcal{F}(T) \subseteq \{0,1\}^n$$

We're fundamentally "just" talking about subsets of $\{0,1\}^n$

For $V \subseteq \{0,1\}^n$, and $J \subseteq [n]$

$$V[J] = \{(v_i)_{i \in J} \mid v \in V\} \quad \text{"behaviors of } V \text{ on } J"$$

V shatters J if $V[J] = \{0,1\}^{|J|}$

$\text{vcdim}(V)$ = largest d s.t. $\exists J \subseteq [n]$, $|J| = d$
that is shattered by V .

Can think of $v \in \{0,1\}^n$ as characteristic vector for a
subset $\{i \in [n] \mid v_i = 1\} \subseteq [n]$

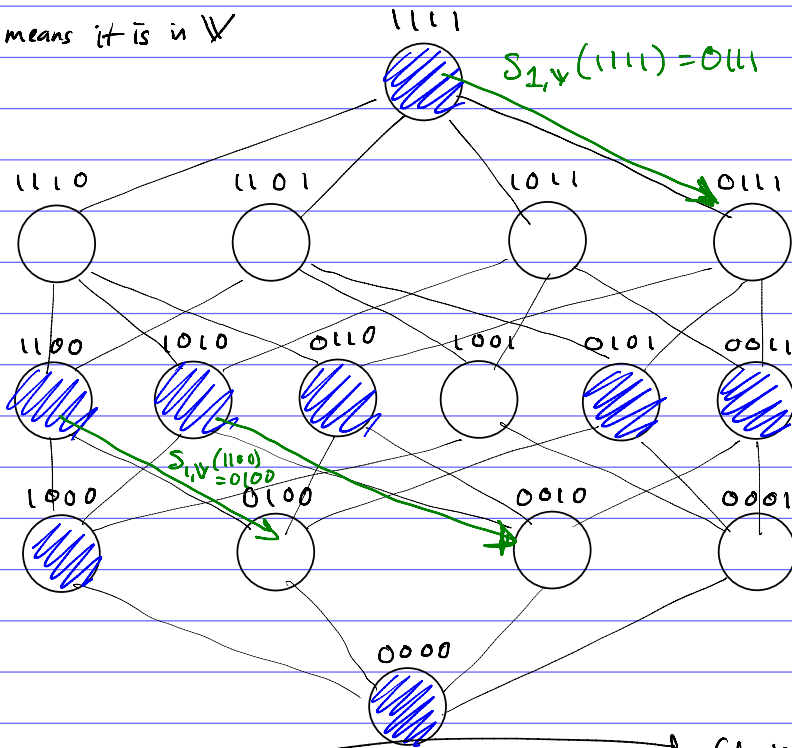
Behavior: $V[J] = \{v \cap J \mid v \in V\}$

V shatters J if $V[J] = 2^J$ = all subsets of J

WTS: if $\text{vcdim}(V) = d$,

$$\text{then } |V| \leq \binom{n}{\leq d}$$

means it is in $\mathbb{V}$



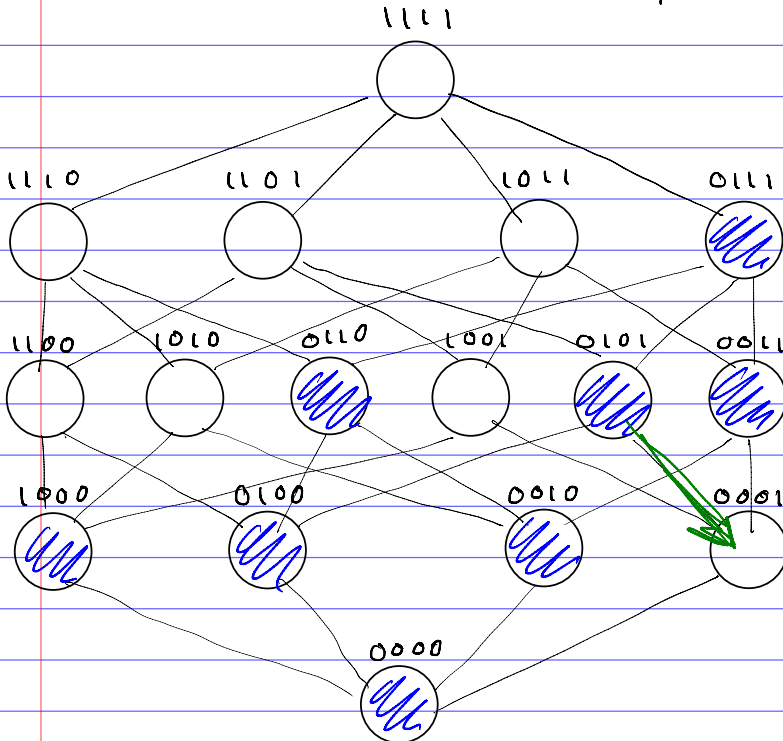
For $i \in [n]$, $v \in \mathbb{V}$

$$S_{i,V}(v) = \begin{cases} v-i & \text{if } i \in v \text{ \& } v-i \in \mathbb{V} \\ v & \text{otherwise} \end{cases}$$

want to transform $\mathbb{V}$ in a way that

- ① doesn't change $|\mathbb{V}|$
- ② no new sets are shattered by $\mathbb{V}$.

(down)shifting



If the shifting operations preserve cardinality
& don't introduce new shattered sets,

then we end up w/ a new $W \subseteq \{0,1\}^n$

with

$$|V| = |W|$$

$$\text{and } d = \text{radim}(V) \geq \text{radim}(W) = d'$$

$$\text{If we show } |W| \leq \binom{n}{\leq d'}$$

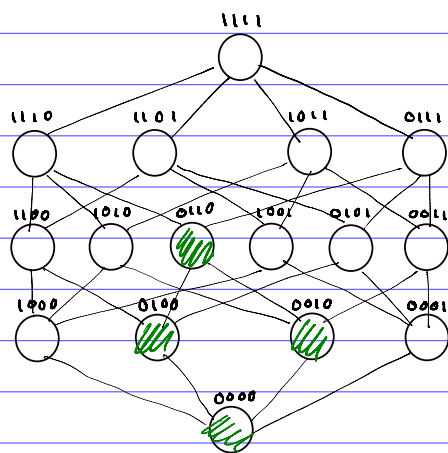
$$\text{then } |V| \leq \binom{n}{\leq d'} \leq \binom{n}{\leq d}$$

If no more shifting is possible,
then

W will be downward-closed

↳ If $v \in W$,

then $\forall u \leq v, u \in W$



WTS: $\mathcal{V}' = \{S_{i,\mathcal{V}}(v) \mid v \in \mathcal{V}\}$

Suppose $J \subseteq [n]$ is shattered by $\mathcal{V}'$

Then J is also shattered by $\mathcal{V}$.

$$S_{i,\mathcal{V}}(v) = \begin{cases} v-i & \text{if } i \in v \\ v & \text{if } i \notin v \end{cases}$$

Case 1: $i \notin J$

$\mathcal{V}'[J] = \mathcal{V}[J]$, so J also shattered by $\mathcal{V}$.

Case 2: $i \in J$

$\forall K \subseteq J, \exists S_{i,\mathcal{V}}(v) \in \mathcal{V}'$ s.t.

$$S_{i,\mathcal{V}}(v) \cap J = K \cup \{i\}$$

can only happen b/c

$$i \in v = S_{i,\mathcal{V}}(v)$$

$$\text{and } v-i \in v$$

works for any
 $K \subseteq J$.

$$v \cap J = K \cup \{i\}$$

$$(v-i) \cap J = K - \{i\}$$

both in $\mathcal{V}$.

So J is shattered by $\mathcal{V}$.