VC dimension

X – ground set

$\mathcal{R} \subseteq \mathcal{P}(X)$ "range"

V – Vapnik
C – Chervonenkis

examples

① $X = \mathbb{R}^2$
$\mathcal{R} = \{ \text{lines in } \mathbb{R}^2 \}$

② $X = \mathbb{R}^3$
$\mathcal{R} = \{ \text{balls in } \mathbb{R}^3 \}$

③ $X = \{0,1\}^n$

$\mathcal{R} = \{\text{subcubes}\}$

$\downarrow$

$\{x \in \{0,1\}^n \text{ s.t } x|_A = b\}$

where $A \subseteq [n]$

$b \in \{0,1\}^A$.

Defn The VC-dimension of $(X, \mathcal{R})$ is the largest d s.t.

$\exists\ A \subseteq X$ with $|A| = d$ and

$$\{A \cap R : R \in \mathcal{R}\} = \mathcal{P}(A).$$

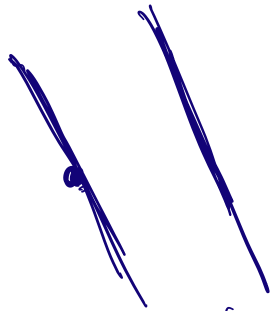
Example $(\mathbb{R}^2,\ \text{lines.})$

Say that

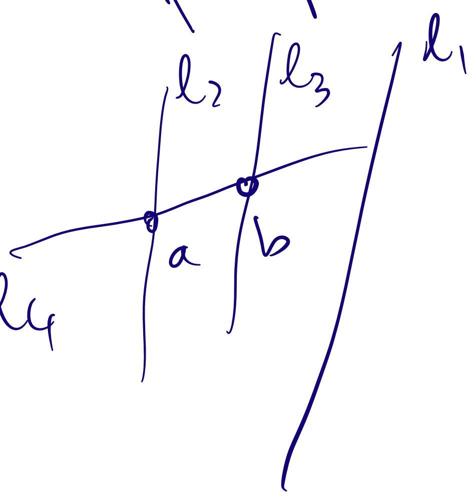
VC dim?

A is "shattered" by R.

1 ok

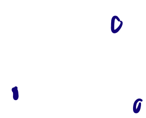


2 OK



ϕ?	l_1	✓
$\{a\}$	l_2	✓
$\{b\}$	l_3	✓
$\{a,b\}$	l_4	✓

3?



either collinear or not

if collinear, then no line hits ~~all 3~~ exactly 2

if not collinear, no line hits all 3.

VC dim = 2

Defn

Shatter function of $(X, \mathcal{R})$

$f : \mathbb{N} \to \mathbb{N}$

$$f(m) = \max_{\substack{A \subseteq X \\ |A| = m}} \left| \{ A \cap R : R \in \mathcal{R} \} \right|$$

Q Express VC dim in terms of shatter function.

A the **largest** m s.t. $f(m) = 2^m$.

Shatter fn for lines in $\mathbb{R}^2$?

Suggestion: m pts on a circle

$f(m) \geq 1 + m + \binom{m}{2}$

Actually equal.

Thm VC

For any range space, exactly one of the following happens.

(1) $f(m) = 2^m$ for all m.

(2) $\exists\, d$ s.t. $f(m) \leq O(m^d)$ $\forall m$.

Sauer Shelah lemma

If (X, R) has VC dim $\leq d$, then the shatter function

$$f(m) \leq 1 + \binom{m}{1} + \binom{m}{2} + \cdots \binom{m}{d}.$$

Sauer-Shelah lemma v2

Suppose $B \subseteq \{0,1\}^m$ with

$$|B| > 1 + \binom{m}{1} + \binom{m}{2} + \dots + \binom{m}{d}$$

Then there exists some $I \subseteq [m]$ with $|I| = d+1$ and

$$B|_I = \{0,1\}^I$$

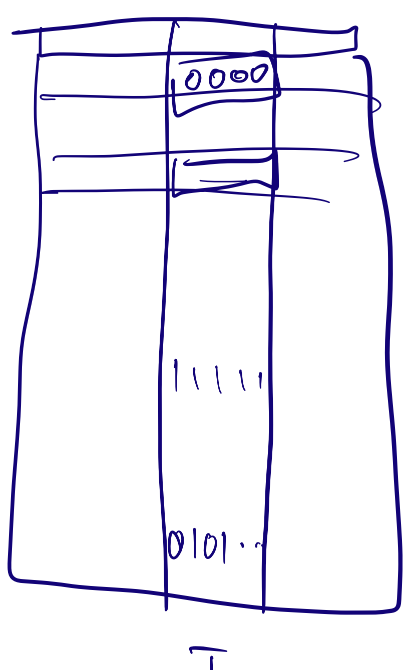
(where $B|_I \triangleq \{x|_I : x \in B\}$).

Says:

m

B

$$> \binom{m}{0} + \binom{m}{1} + \dots + \binom{m}{d}$$



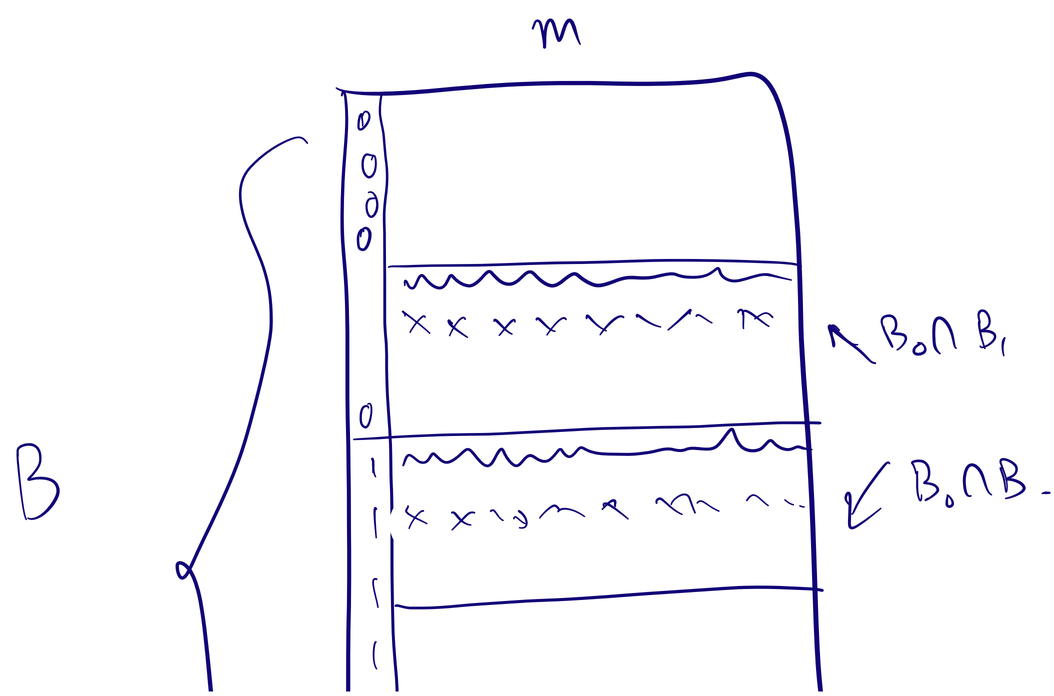
I

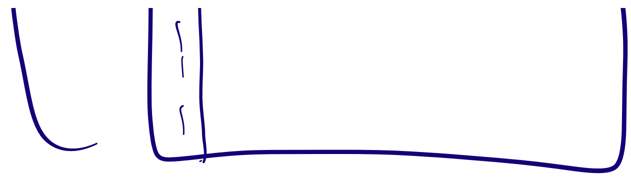
<u>Example</u> Take $B = \{x \in \{0,1\}^m$ with $\le d$ 1's$\}$

$$|B| = \binom{m}{0} + \cdots \binom{m}{d}$$

For any I of size $d+1$, we do not see the $(1,1,\ldots 1) \in \{0,1\}^{d+1}$ inside $B|_I$.

$B \subseteq \{0,1\}^m$ $\quad |B| > \binom{m}{0} + \cdots + \binom{m}{d}$.





Proof By induction on m, d.

$B_0 = \{ y \in \{0,1\}^{m-1} \text{ s.t. } 0y \in B \}$

$B_1 = \{ y \in \{0,1\}^{m-1} \text{ s.t. } 1y \in B \}$.

$B_0 \cap B_1$

$B_0 \cup B_1$

$|B| = |B_0| + |B_1| = |B_0 \cap B_1| + |B_0 \cup B_1|$

define $g(m,d) = \binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{d}$.

Identity $\quad g(m,d) = g(m-1,d) + g(m-1,d-1)$

If $|B| > g(m,d)$, then

Either ① $|B_0 \cap B_1| > g(m-1,d-1)$

or ② $|B_0 \cup B_1| > g(m-1,d)$

In case (2), by IH, we get a subset J of $\{2,3,\dots m\}$ of size $d+1$ s.t. $(B_0 \cup B_1)\big|_J = \{0,1\}^J$.

This J works for B also.

In case (1). by IH, we get a subset J of $\{2,3,\dots m\}$ of size $\underline{\underline{d}}$ s.t. $(B_0 \cap B_1)\big|_J = \{0,1\}^J$.

Then $I = \{1\} \cup J$ is of size $d+1$ and s.t. $B\big|_I = \{0,1\}^I$.

<u>Proof 2</u> (Smolensky).

Suppose for every $I \subseteq [m]$ of size

$d+1$, there is an omitted pattern $y_I \in \{0,1\}^I$. Then let us show that B is small.

View B as a subset of $\mathbb{F}_2^m$.

Consider $V = \{$all $f: B \to \mathbb{F}_2\}$. This is an $\mathbb{F}_2$-vector space. $\dim(V) = |B|$.

We will give a spanning set for V.

$$S = \left\{ \prod_{j \in J} x_j : J \subseteq [m], |J| \leq d \right\}$$

<u>Claim</u> Every $f: B \to \mathbb{F}_2$ can be expressed as

$$f(x) = \sum_{\substack{J \subseteq [m] \\ |J| \leq d}} a_J \cdot \prod_{j \in J} x_j$$

<u>Observation:</u>

Let $|I| = d+1$.

Suppose $B|_I$ misses $(1,1,\ldots,1) \in \{0,1\}^{d+1}$

Then $\prod_{i \in I} x_i = 0$ on B

Obs:

Let $|I| = d+1$

Suppose $B|_I$ misses $(0,0,\ldots,0) \in \{0,1\}^{d+1}$

Then $\prod_{i \in I} (x_i - 1) = 0$ on B.

So $\prod_{i \in I} x_i = \left(\text{some poly of degree} \le d\right)$ on B.

$x_i^2 = x_i$ on B.

For any I of size $d+1$,

(.) = on B

$$\prod_{i \in I} (x_i - (y_I)_i - 1) = 0$$

$$\Rightarrow \prod_{i \in I} x_i = \text{some deg} \le d \text{ poly}$$

Start with any $f : B \to \mathbb{F}_2$.

We first write it as

$$f = \sum_{J \subseteq [m]} a_J \prod_{j \in J} x_j \quad \text{(can always be done)}.$$

Keep replacing x_i^2 with x_i,

and $\prod_{i \in I} x_i$ (for $|I| = d+1$) with

some lower degree poly.

When done, there is no

monomial of degree $d+1$ or more,
and we got a representation
of the exact same fn on B.

So $\dim(V) \leq |S| = g(m,d)$.

Can answer things like this:
How many regions in $\mathbb{R}^2$ can
n circles create?

Many applications in
1. combinatorial geometry
2. computational geometry.

$X = \mathbb{R}^2$

$\mathcal{R}$ = unions of 2 lines.

What is VC-dim of $(X, \mathcal{R})$?

We know that for lines

shatter fn

$f(m) \leq 1 + \binom{m}{1} + \binom{m}{2}$.

So for unions of lines

shatter fn

$$f(m) \leq \binom{1 + m + \binom{m}{2}}{2}$$

Is there an m for which

$f(m) < 2^m$?

<u>Yes</u> $\binom{1+9+10}{2} \cdots \binom{6}{2} = 21$

$m = 10$ works.

~~Lemma~~ If R_1, R_2 are range spaces on X, with VC dim d, then $R_1 \cup R_2$ has VC dim $\leq O(d \log d)$.

<u>Proof</u>:

Shatter functions

f_1 of R_1

f_2 of R_2

f of $R_1 \cup R_2$

Then $f(m) \leq f_1(m) \cdot f_2(m)$

$\leq m^d \cdot m^d$

$= m^{2d}$

Then for which m is

$m^{2d} < 2^m$.

$2^{2d \log m} < 2^m$

i.e. $\frac{m}{\log m} > 2d$.

$m > O(d \log d)$.

So for $m = O(d \log d)$ we

get non-shattering
of some set of size m.