Plan: 1 Vapnik–Chervonenkis Dimension
2 Lower Bound
3 Concentration Inequ. → Upper Bound
4 Empirical Process → Upper — " —

1 VC – Dimension

X set $\quad \mathcal{C} \subseteq 2^X$ collection of subsets

Ex: $X = [0,1]^d$, $\mathcal{C} = \mathcal{B}$

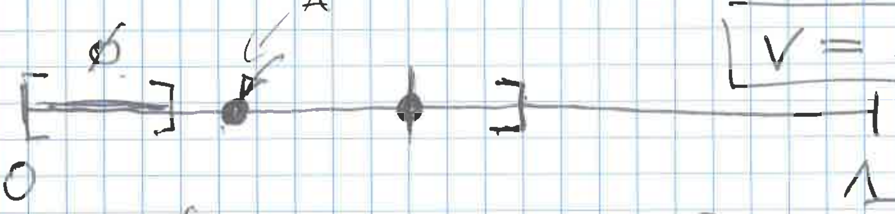
$A \subseteq X$ finite is shattered by $\mathcal{C}$

$\Longleftrightarrow A \cap \mathcal{C} = \{ A \cap C : C \in \mathcal{C} \} = 2^A$

VC – Dimension of $\mathcal{C}$:

$$v = v(\mathcal{C}) = \max\{ \# A : A \text{ is shattered by } \mathcal{C} \}$$

Ex: $\boxed{v = 1} = d$



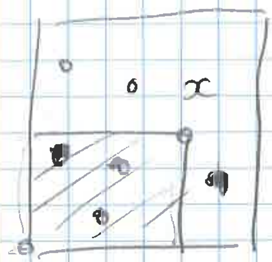
$\exists\ \mathcal{C} = \{ [0,x] : x \in [0,1] \}$

~~all~~ 1-pt sets are shattered, ~~thate~~ 2 pt-sets not shattered

Star Discrepancy in High Dimension

The problem:

$$P_n \subseteq [0,1]^d \,, \qquad \# P_n = n$$



$$\mathcal{B} = \{ B_x = [0,x] : x \in [0,1]^d$$

$$\Delta_{P_n}(x) = \frac{1}{n} \# P_n \cap B_x - |B_x|$$

$$= \frac{1}{n} \sum_{i=1}^{n} 1_{[0,x)}(t_i) - x_1 \cdots x_d$$

$$\mathrm{disc}^*(P_n) = \sup_x |\Delta_{P_n}(x)|$$

$$\mathrm{disc}^*(n,d) = \inf_{\# P_n = n} \mathrm{disc}^*(P_n)$$

$$n^*(\varepsilon,d) = \min\{n : \exists P_n \ \mathrm{disc}^*(P_n) \le \varepsilon\}$$

AIM: $\min\left(\varepsilon_0, \frac{cd}{n}\right) \le \mathrm{disc}^*(n,d) \le C \sqrt{\frac{d}{n}}$

$C \le 10$ 100

$\varepsilon_0, \ c, C > 0$

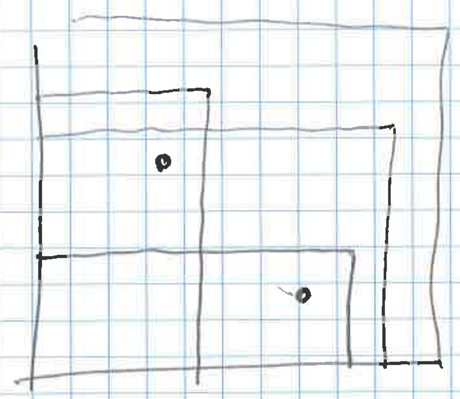
$$c \cdot \frac{d}{\varepsilon} \le n^*(\varepsilon, d) \le C \cdot \frac{d}{\varepsilon^2} \qquad 0 < \varepsilon \le \varepsilon_0$$

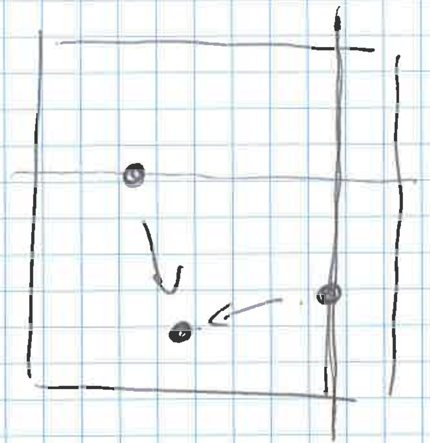
H. 2004

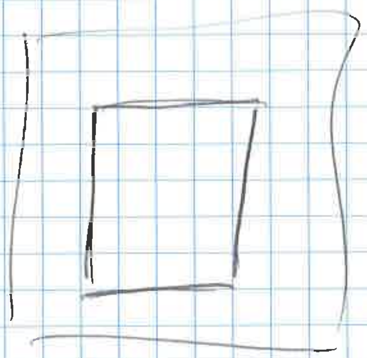
Heinrich, Novak, Wasilkowski, Wozniakowski 2001

Aistleitner 2011

$d \geq 2$ $\quad n = d \quad \Rightarrow \quad vc(\mathcal{B}) = d$







$vc = 2d$

$vc(\mathcal{B}^{\pi}) = \,?\, \geq 2d+1$

Conjecture $3d$ $\quad d=2$

$vc(\mathcal{C}) = v$

THM: $m = \# A$. Then

$$\# A \cap \mathcal{C} \leq \sum_{i=0}^{v} \binom{m}{i} \leq \left(\frac{em}{v}\right)^{v}$$

Sauer '72, Shelah '72, VC '71

equivalent: $\# \mathcal{C} > \sum_{i=1}^{v-1} \binom{m}{i}$, $\mathcal{C} \subseteq 2^A$

$\Longrightarrow$ $\mathcal{C}$ shatters a set of size v in A.

Strengthening: $\mathcal{C}$ finite shatters at least $\# \mathcal{C}$ sets.

Pajor '84

Pajor $\Rightarrow$ Sauer: $\# \mathcal{C} > \sum_{i=1}^{v-1} \binom{m}{i} = \# \binom{A}{< v}$

$\Longrightarrow$ $\mathcal{C}$ shatters a set of size v

Proof Pajor Induction.

$\# \mathcal{C} = 1 \Longrightarrow \mathcal{C}$ shatters $\emptyset$

$\# \mathcal{C} > 1$: $x \in A$ x is not contained in all $C \in \mathcal{C}$.

$\mathcal{C}_1 = \{C \in \mathcal{C} : x \in C\}$ $\quad \mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$

$\mathcal{C}_2 = \{C \in \mathcal{C} : x \notin C\}$

$\# \mathcal{C} = \# \mathcal{C}_1 + \# \mathcal{C}_2$

$\mathcal{C}_1$ shatters $\geq \# \mathcal{C}_1$ sets

$\mathcal{C}_2$ — " — $\geq \# \mathcal{C}_2$ sets

Problem: Some sets maybe shattered by $\mathcal{C}_1$ and $\mathcal{C}_2$.

Assume: $B = A \setminus \{x\}$ shattered by $\mathcal{C}_1$ & $\mathcal{C}_2$.

$\Longrightarrow B \cup \{x\}$ is shattered by $\mathcal{C}_1 \cup \mathcal{C}_2 = \mathcal{C}$.

$\Longrightarrow \mathcal{C}$ shatters at least $\# \mathcal{C}_1 + \# \mathcal{C}_2 = \# \mathcal{C}$ sets ☺

General Discrepancy

$(X, \mathbb{P})$ prob. space, $\mathcal{C} \subseteq 2^X$

(Countable) collection of measurable sets.

$\mathcal{P}_n \subseteq X$ $\quad \Delta_{\mathcal{P}_n}(C) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_C(t_i) - \mathbb{P}(C)$

$\mathcal{P}_n = \{t_1, \dots, t_n\}$

$$\operatorname{disc}^{\mathcal{C}}(\mathcal{P}_n) = \sup_{C \in \mathcal{C}} |\Delta_{\mathcal{P}_n}(C)|$$

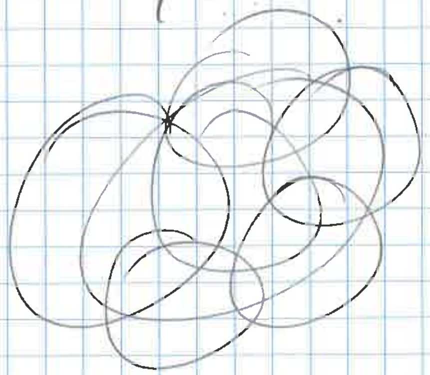
Ex: $\operatorname{disc}^{\mathcal{B}} = \operatorname{disc}^*$

(2) Proof of lower bound

Covering Numbers

(M,d) pseudo-metric space, $\varepsilon>0$

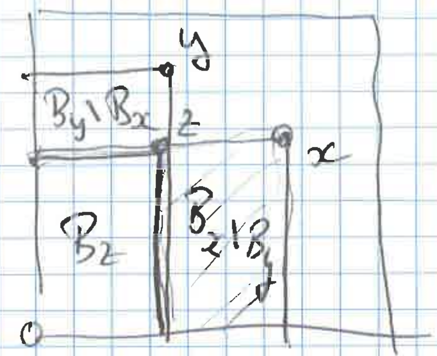
$N(M,d,\varepsilon)$ = minimal number of closed ε-balls $B(x,\varepsilon)=\{y\in M: d(x,y)\le\varepsilon\}$ covering M.



THM:
- $\mathcal{C}\subseteq 2^X$, $v=VC(\mathcal{C})$, closed under intersections $(C_1,C_2\in\mathcal{C}\Rightarrow C_1\cap C_2\in\mathcal{C})$
- $(X,\mathbb{P})$ probability as above, $d_{\mathbb{P}}(C_1,C_2)=\mathbb{P}(C_1\Delta C_2)$
- $N(\mathcal{C},d_{\mathbb{P}},\varepsilon)\ge(K\varepsilon)^{-v}$ for some $K>0$

Then $\quad disc^{\mathcal{C}}(\mathbb{P}_n)\ge\min\left(\varepsilon_0,\frac{cv}{n}\right).$

Ex: $\mathcal{C} = \mathcal{B}$, $X = [0,1]^d$, $\mathbb{P}$ = Lebesgue



$z = x \wedge y$

↑ coordinatewise min.

$B_x \wedge B_y = B_z$.

$$\Longrightarrow \quad \operatorname{disc}^*(P_n) \geq \min\left(\varepsilon_0, \frac{cd}{n}\right)$$

Assumption on $N(\mathcal{C}, d_{\mathbb{P}}, \varepsilon)$.

~~$d_{\mathbb{P}}(B$~~ $\quad d_{\mathbb{P}}(B_x, B_y) = |B_x \Delta B_y|$

Claim: $d(x,y) = d_{\mathbb{P}}(B_x, B_y) = |B_x \Delta B_y|$

$$\geq |B_{x \wedge y}| \, \|x-y\|_1$$

$$\|z\|_1 = \sum_{i=1}^{d} |z_i| .$$

$$|B_x \setminus B_y| = |B_x \setminus B_z| \qquad z = x \wedge y$$

$$= \prod_1^d x_i - \prod_1^d z_i$$

$$= \sum_{i=1}^d x_1 \ldots x_{i-1} \underline{(x_i - z_i)}\, z_{i+1} \ldots z_d$$

$$\geq \prod_{i=1}^d z_i \underbrace{\sum_{i=1}^d (x_i - z_i)}$$

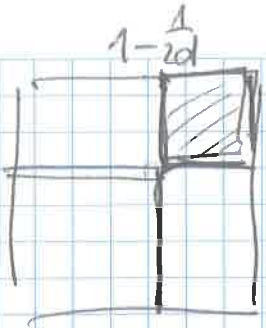
$$= |B_z| \cdot \|x - z\|_1$$

$$|B_y \setminus B_x| \geq |B_z| \cdot \|y - z\|_1$$

$$|B_x \Delta B_y| \geq |B_z| \left(\|x - z\|_1 + \|y - z\|_1\right)$$

$$= |B_z| \sum_{i=1}^d \underbrace{(x_i + y_i - 2z_i)}_{|x_i - y_i|}$$

$$|B_x \Delta B_y| \geq |B_z|\, \|x - y\|_1$$



$$x, y \in \left[1-\frac{1}{2d}, 1\right]^d = M$$

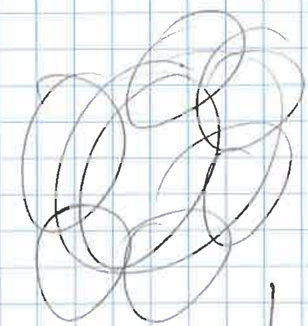
$$|B_z| \geq \left(1-\frac{1}{2d}\right)^d \geq \frac{1}{2}$$

$$d_\Pi(B_x, B_y) \geq \frac{1}{2}\|x-y\|_1 .$$

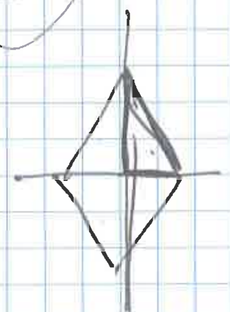
$$N(B, d_\Pi, \varepsilon) \geq N(M, d_\Pi, \varepsilon)$$

$$\geq N(M, \|\cdot\|_1, 2\varepsilon)$$

Packing numbers Maximal set with pairwise distances $\geq \varepsilon$.



$$\geq \frac{|M|}{|\varepsilon B(\ell_1^d)|}$$



$$= \frac{|M|}{\varepsilon^d |B(\ell_1^d)|}$$

$$|M| = \left(\frac{1}{2d}\right)^d \qquad |B(\ell_1^d)| = \frac{2^d}{d!}$$

$$N \geq \frac{\left(\frac{1}{2d}\right)^d}{\varepsilon^d \, 2^d/d!} = \frac{d!}{(4\varepsilon d)^d} \geq \cancel{\frac{1}{(8\varepsilon d)^d}}$$

$$\geq \frac{1}{(8e\,\varepsilon)^d}$$

<u>Proof of the Theorem</u>

(A) $P_n \subseteq X$, $\varepsilon > 0$ st $\sum_{i=0}^{V} \binom{n}{i} < N(\mathcal{C}, d_P, \varepsilon)$

Then there exist $C_1, C_2 \in \mathcal{C}$ such that

$C_1 \neq C_2$ and $\mathbb{P}(C_1 \Delta C_2) > \varepsilon$

~~and disc~~ $\cancel{\in \mathcal{P}_n}$

$$P_n \cap C_1 = P_n \cap C_2$$

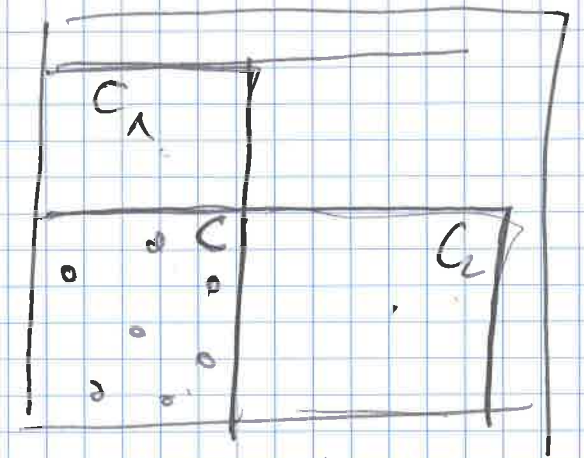
$\mathcal{N} \subseteq \mathcal{C}$ maximal subset with $d_P(C_1, C_2) > \varepsilon$ for $C_1, C_2 \in \mathcal{N}$

$$\# \mathcal{N} \geq N(\mathcal{C}, d_P, \varepsilon) > \sum_{i=0}^{V} \binom{n}{i}$$

Same Lemma: $C_1, C_2 \in \mathcal{N}$

$$P_4 \cap C_1 = P_4 \cap C_2$$

③ $\mathrm{disc}^{\mathcal{E}}(P_4) \geq \frac{\varepsilon}{4}$ //



$C = C_1 \cap C_2$

$P_4 \cap C = P_4 \cap C_1 = P_4 \cap C_2$

$$\mathbb{P}(C_1 \Delta C_2) > \varepsilon$$

④ $\mathrm{disc}^{\mathcal{E}}(P_4) \geq -\Delta_{P_4}(C_1) - \Delta_{P_4}(C_2) + 2\Delta_{P_4}(C)$

$$\geq +\mathbb{P}(C_1) + \mathbb{P}(C_2) - 2\mathbb{P}(C)$$

$$= \mathbb{P}(C_1 \Delta C_2)$$

$$> \varepsilon$$