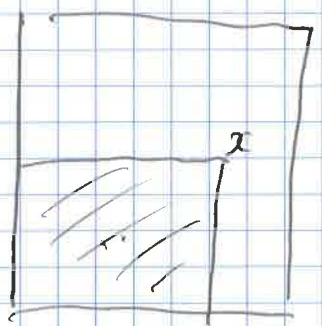
$$\mathrm{disc}^*(\mathcal{P}_n) = \sup \left| \Delta_{\mathcal{P}_n}(x) \right|$$
$$= \sup_x \left| \frac{1}{n} \# \mathcal{P}_n \cap B_x - |B_x| \right|$$



general discrepancy

X set, $\mathcal{C} \subseteq 2^X$, $\mathbb{P}$ on X

$$\mathrm{disc}^{\mathcal{C}}(\mathcal{P}_n) = \sup_{C \in \mathcal{C}} \left| \Delta_{\mathcal{P}_n}(C) \right|$$
$$= \sup_{C \in \mathcal{C}} \left| \frac{1}{n} \# \mathcal{P}_n \cap C - \mathbb{P}(C) \right|$$

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

$$\mathrm{disc}^{\mathcal{C}}(n,d) = \inf_{\# \mathcal{P}_n = n} \mathrm{disc}^*(\mathcal{P}_n)$$
$$n^{\mathcal{C}}(\varepsilon, d) = \min \{ n : \mathrm{disc}^*(\mathcal{P}_n) \leq \varepsilon \}$$

$$\min\left(\varepsilon_0, \frac{cV}{n}\right) \leq \mathrm{disc}^{\mathcal{C}}(n,V) \leq c \sqrt{\frac{V}{n}}$$

yesterday $\quad\quad$ today.

$V = vc(\mathcal{C})$.

assumptions.

$$d_{\mathbb{P}}(B_x, B_y) = |B_x \Delta B_y| \geq \frac{1}{2}\|x-y\|_1$$

$$N(\mathcal{B}, d_{\mathbb{P}}, 2\varepsilon) \geq N(M, \|\cdot\|_1, \varepsilon)$$

$N(M, d, \varepsilon)$ = minimal number of ε-balls covering M

$d_1 \geq d_2$ distances on M

$$N(M, d_1, \varepsilon) \geq N(M, d_2, \varepsilon)$$

$$\{y \in M : d_1(x,y) \leq \varepsilon\} \subseteq \{y \in M : d_2(x,y) \leq \varepsilon\}$$

③ Concentration Inequalities – Slightly weaker upper bound.

$$\mathrm{disc}^{\mathcal{C}}_{T}(n, V) \leq C \cdot \sqrt{\frac{V}{n}}\,(\log V + \log n)^{1/2}$$

$\mathcal{C} = \mathcal{B}$ $\quad V = d$

Have to find $\mathcal{P}_n \subseteq [0,1]^d$

$$\mathrm{disc}^{*}(\mathcal{P}_n) \leq \cancel{\quad n \quad}$$

Probabilistic Method

$P_n = \{t_1, .., t_n\} \subseteq [0,1]^d$

$t_1, .., t_n$ independent and uniformly distributed in $[0,1]^d$.

Show: $\mathbb{P}(\mathrm{disc}^*(P_n) < 2\varepsilon) > 0$

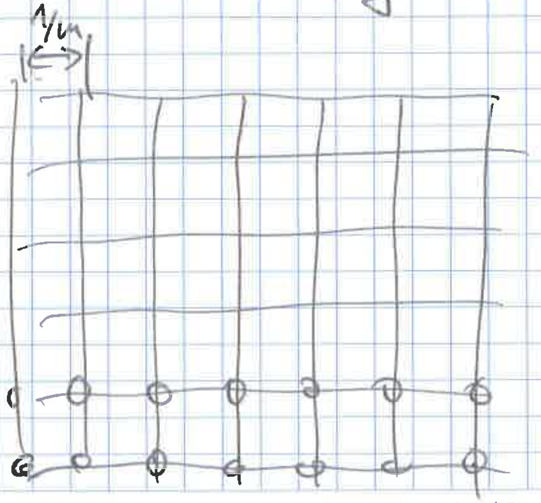
for $\varepsilon = \sqrt{\frac{d}{n}} (\log d + \log n)^{1/2}$

$\Longleftrightarrow \mathbb{P}(|\Delta_{P_n}(x)| \geq 2\varepsilon \text{ for some } x) < 1$

$|\Delta_{P_n}(x)| \geq 2\varepsilon$ for some x

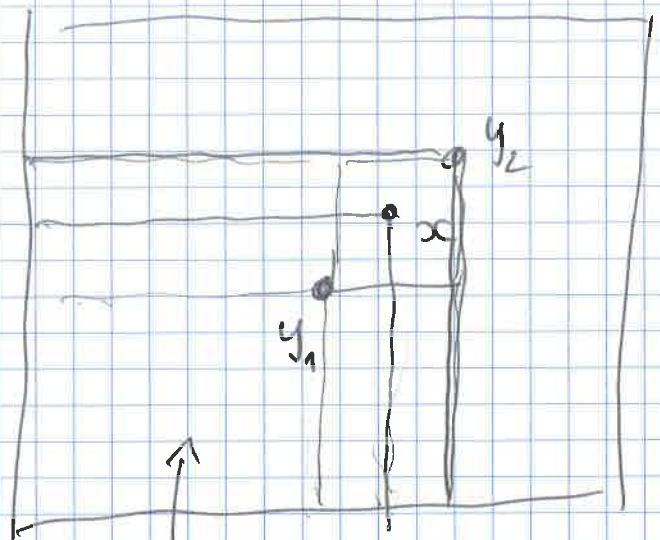
$\Longrightarrow |\Delta_{P_n}(y)| \geq \varepsilon$ for some y

for some $y \in \Gamma_m = \{$points with coordinates $\frac{k_i}{m}$, $k_i \in \{0,1,..,m\}\}$



$\# \Gamma_m = (m+1)^d$

$m = \lceil \frac{d}{\varepsilon} \rceil$



too many path in B_x

$\Rightarrow$ too many path in B_{y_2}.

$$\mathbb{P}\left(|\Delta_{P_n}(y)| \geq \varepsilon \text{ for some } y \in \Gamma_m\right) < 1$$

$$\wedge\!\!|$$

union bound $\quad \sum_{y \in \Gamma_m} \mathbb{P}\left(|\Delta_{P_n}(y)| \geq \varepsilon\right) \overset{\text{to show}}{<} 1$

for each y: $\quad \mathbb{P}\left(|\Delta_{P_n}(y)| \geq \varepsilon\right) < \frac{1}{\#\Gamma_m} = (m+1)^{-d}$

~~to show~~

$$n\,\Delta_{\mathcal{P}_n}(y) = \sum_{i=1}^{n} \left(\mathbb{1}_{B_y}(t_i) - |B_y| \right)$$

$$X_i = \mathbb{1}_{B_y}(t_i) - |B_y|$$

$$\mathbb{E} X_i = \mathbb{E}\,\mathbb{1}_{B_y}(t_i) - |B_y|$$

$$= |B_y| - |B_y| = 0.$$

$$|X_i| \leq 1$$

$$n\,\Delta_{\mathcal{P}_n}(y) = \underbrace{\sum_{i=1}^{n} X_i} \qquad X_i \text{ iid.}$$

<u>HOEFFDING's Inequality</u>

$$\mathbb{P}(|n\,\Delta_{\mathcal{P}_n}(y)| > t) = \mathbb{P}\left(\left|\sum_1^n X_i\right| > t\right)$$

$$\leq 2 \cdot e^{-t^2/2n}$$

$\underline{t = n \cdot \varepsilon}$ $\quad \mathbb{P}(|\Delta_{\mathcal{P}_n}(y)| > \varepsilon) \leq 2 \cdot e^{-\frac{n^2\varepsilon^2}{2n}}$

$$= 2 \cdot e^{-n\varepsilon^2/2}$$

this is true for our choice

<u>to show</u> $< (m+1)^{-d}$

④ Empirical Processes – true upper bound |

$X_1, \dots, X_n$ i.i.d $\sim \mathbb{P}$ on some set X

$\mathcal{F}$ class of measurable fcts.

$$\alpha_n(f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(f(X_i) - \mathbb{E}(f(X_i)) \right)$$

empirical process to $\mathcal{F}$.

$X_i = t_i \in [0,1]^d$ unif. distributed.

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

$$\cancel{\alpha_n(f)} \quad \alpha_n(1_{B_x}) = \sqrt{n}\, \Delta_{P_n}(x)$$

$$\mathbb{E} \sup_x |\alpha_n(1_{B_x})| = \sqrt{n}\, \mathbb{E}\, disc^*(P_n) \quad \Big\|$$ Show $< c\sqrt{v}$

Talagrand, Haussler: $\mathcal{F} = \{ 1_C : C \in \mathcal{C} \}$, $vc(\mathcal{C}) = v$

$\mathbb{P}$

$$\mathbb{P}\left(disc^*(P_n) \geq \underbrace{c \frac{1}{\sqrt{n}}}_{\alpha \cdot K \sqrt{\frac{v}{n}}} \right) \leq \frac{1}{c} \left(\frac{Kc^2}{v} \right)^v e^{-2v^2} < 1$$

$$\frac{1}{\sqrt{v}} \quad v^v \qquad e^{-2v^2}$$

Tractability of high dimensional problems - integration

<u>Ex</u> Discrepancy $\mathrm{disc}^*(P_n) = L_\infty$-norm

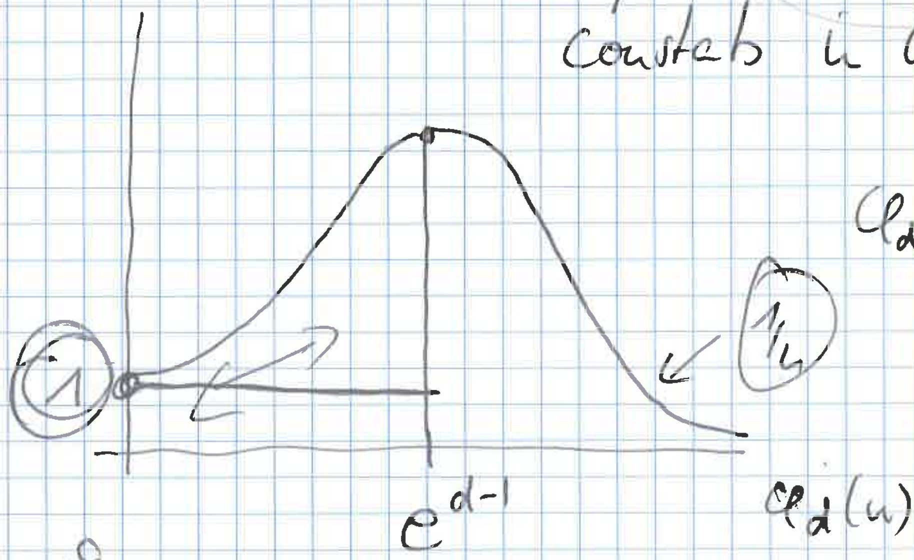
= worst case error of integration of functions from the unit ball of some L_1-Sobolev space with mixed first derivatives.

"classical": $\mathrm{disc}^*(P_n) \lesssim \dfrac{(\log n)^{d-1}}{n} = \varphi_d(n)$

constants in d.

for fixed d.

$$\varphi_d(x) = \frac{(\log x)^{d-1}}{x}$$



$$\boxed{10 \cdot \sqrt{\frac{d}{n}} \ll \frac{(\log n)^{d-1}}{n}}$$

$\cancel{e^{19}}$ $n \le e^{d-1}$ the bound does not tell you anything.

$d = 20$ for $n = d^k$

$*$-Discrepancy Problem

tractable meaning "solvable"

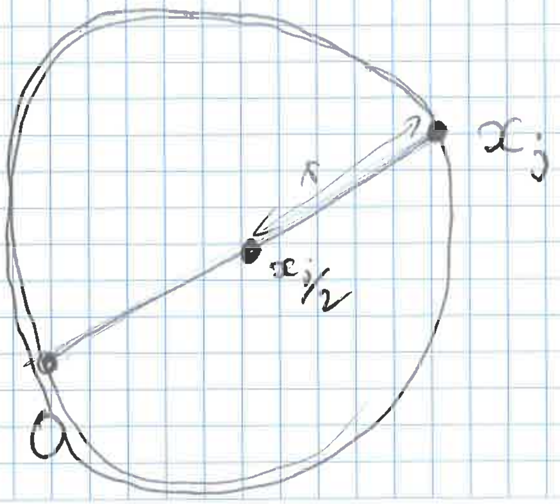
numerically in high dimensions.

Something Completely Else

Lemma: $x_1, .., x_n \in \mathbb{R}^d$, $0 \in \mathbb{R}^d$

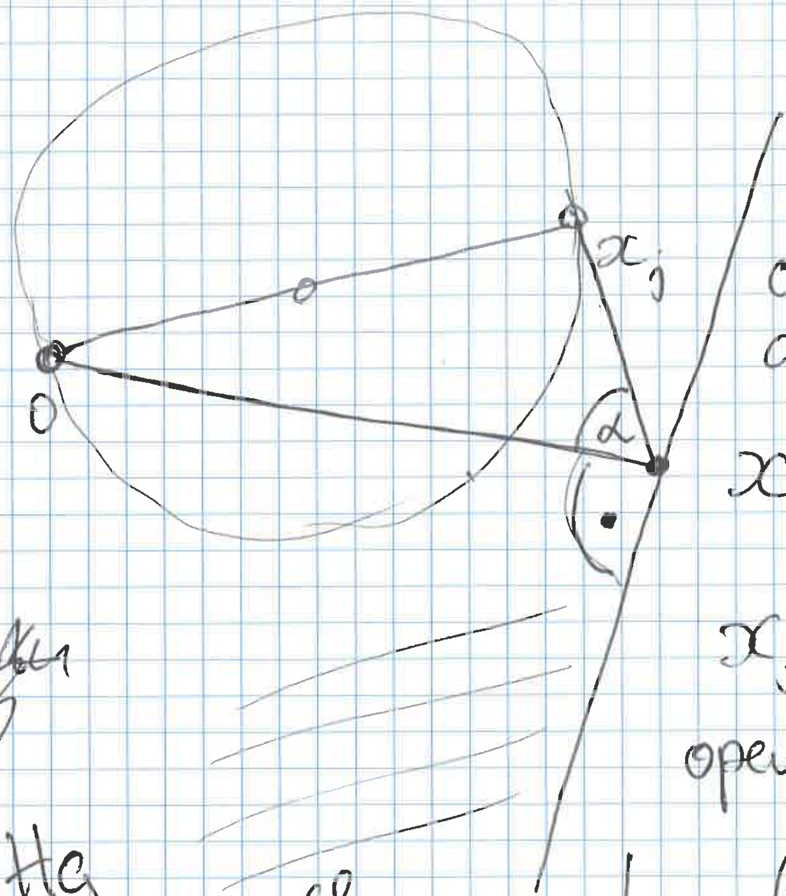
$K = \mathrm{conv}\{x_1, .., x_n\}$.

Then: K is contained in the union of ~~the~~ the balls (euclidean) with centers in $\frac{x_j}{2}$ and radius $\left|\frac{x_j}{2}\right| = \left\|\frac{x_j}{2}\right\|_2$



Pf: $x \in \mathbb{R}^d$ outside of all those balls.

$\alpha < \pi/2$



H hyperplane orthogonal to Ox and passes to x

x_j lies in this open half-space

true for $x_1, x_2, \dots, x_n$

H_a open half space

$\Rightarrow K = \mathrm{Conv}\{x_1, \dots, x_n\}$ also lies in this half space

$\Rightarrow x \notin K$.