Chap 2

Class 5 | I

Chap 2

II - Fluctuatio

[3] Proof of Talagrand's inequality

II - [3] Proof of Talagrand's inequality

Recall the definition of the convex distance.

$$d_C(x,A) := d_E(0, \text{Conv}\,\mathcal{S}(-x+A))$$

& $\mathcal{S}(A-x) := \{ s \in \{0,1\}^n \mid \exists y \in A, \ s_i = 0 \Rightarrow y_i = x_i \}$

$\equiv$ " Union of the supports of the $y \in A$ whose coordinates coincide with those x "

$\rightsquigarrow$ Measures the complexity of the approx of x by A

$\equiv$ " Set of boolean vectors (in $\{0,1\}^n$) indicating the coordinates that must be changed in x to reach A "

($\simeq$ Transport plan of element x to A)

The proof is by recurrence over n, & yields an a-dimensional result. $\Omega_n = [0,1]^n$

- $\underline{n=1}$:

$$\mathcal{S}(A-x) = \begin{cases} \{0,1\} & \text{if } x \in A \\ \{1\} & \text{if } x \notin A \end{cases}$$

$$\Longrightarrow \text{Conv}\,\mathcal{S}(A-x) = \begin{cases} [0,1] & \text{if } x \in A \\ 1 & \text{if } x \notin A \end{cases}$$

$$\Rightarrow d_C(x,A) = d_E(0, \text{Conv}\,\mathcal{S}(A-x)) = \mathbb{1}_{\{x \notin A\}}$$

$$\Rightarrow \quad \mathbb{E}\left(e^{\frac{1}{4} d_c(x,A)^2}\right)$$

$$= e^{\frac{1}{4}}\, \mathbb{P}(A^c) + \mathbb{P}(A) = e^{\frac{1}{4}}(1-p) + p \quad \text{for } p \in [0,1]$$

$$\leq \ 1/p$$

$\uparrow$ Prove $\frac{1}{p} - p - e^{1/4}(1-p) \geq 0$

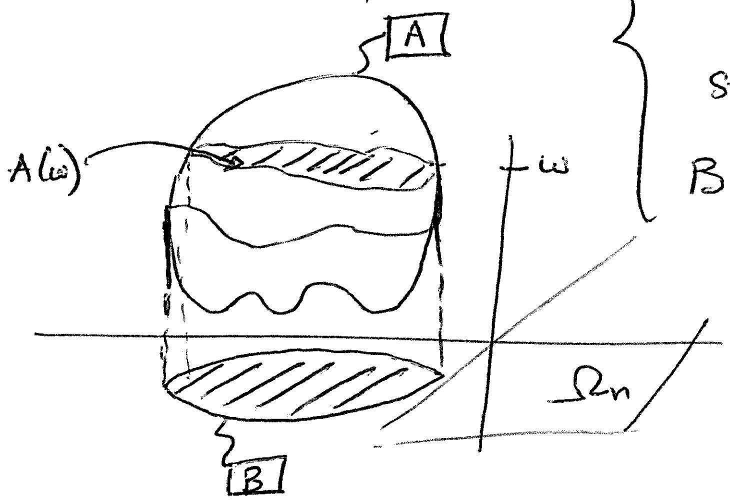
- Recurrence $n \Rightarrow n+1$.

To that endeavor, write

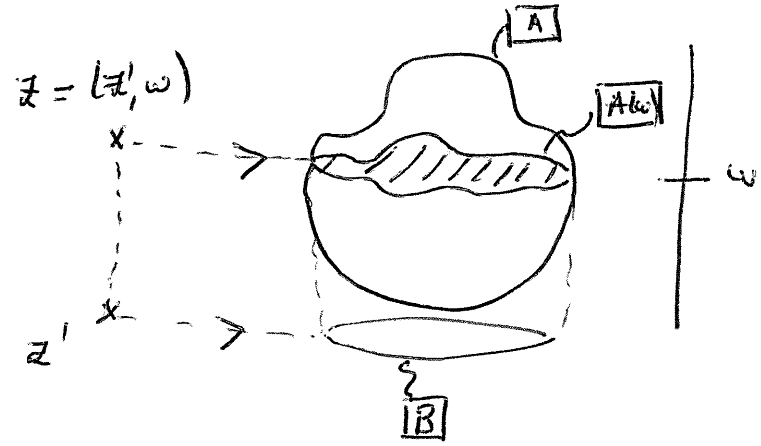
$$\underset{z}{\underset{\cup\!\!\!\,\shortmid}{\Omega_{n+1}}} = \underset{(z',}{\underset{\cup\!\!\!\,\shortmid}{\Omega_n}} \times \underset{\omega)}{\underset{\cup\!\!\!\,\shortmid}{[0,1]}}$$

Idea 1: Control the convex distance $d_c(z, A)$ using $\begin{cases} \text{the projection of } A \\ \text{slices of } A \text{ indexed by } \omega \end{cases}$

Given $A \subset \Omega_{n+1}$, define
$$\begin{cases} A(\omega) := \{ z' \in \Omega_n \mid (z', \omega) \in \Omega_{n+1} \} \subset \Omega_n \\ \text{slice at level } \omega \in [0,1] \\ B = \bigcup_{\omega \in [0,1]} A(\omega) \quad \text{projection} \subset \Omega_n. \end{cases}$$



Now
- $s \in \mathcal{S}(A - x)$ contains 0 at the positions where it is not necessary to change x in order to reach $y \in A$.
- There are two possible "simple" plans to reach $A \subset \Omega_{n+1}$ from z using transport plans of dim n.



* Either we leave ω fixed & use an $s' \in \mathcal{S}(A(\omega) - z')$

* Either move ω to 0 and use a transport plan $s' \in \mathcal{S}(B - z')$

Important remarks

⚠ $B = \bigcup_{\omega \in [0,1]} A(\omega)$ union of slices $\Longrightarrow$ In 2^{nd} case, s' is the transport plan to <u>some</u> slice

⚠ If $A(\omega) = \emptyset$, only the 2^{nd} case is possible

Considering the points reached after transport / coordinate modificato:

$$a \in \mathcal{S}(A(\omega) - z') \Longrightarrow (a,0) \in \mathcal{S}(A - z)$$

(& $(a,1)$ but useless)

$$b \in \mathcal{S}(B - z') \Longrightarrow (b,1) \in \mathcal{S}(A - z)$$

By convexity

$$\begin{cases} a \in \operatorname{Conv} \mathcal{S}(A(\omega) - z') \\ b \in \operatorname{Conv} \mathcal{S}(B - z') \end{cases} \Longrightarrow \mathcal{S}(A - z) \ni \lambda (a,0) + (1-\lambda)(b,1) \quad \forall \lambda \in [0,1]$$

$$\Longrightarrow d_c(z,A)^2 := d_E\left(0, \operatorname{conv} \mathcal{S}(A - z)\right)^2$$

$$\leq \left| \lambda(a,0) + (1-\lambda)(b,1) \right|^2_{E(\Omega_{n+1})}$$

$$= (1-\lambda)^2 + \left| \lambda a + (1-\lambda) b \right|^2_{E(\Omega_n)}$$

$$\leq (1-\lambda)^2 + \lambda |a|^2_{E(\Omega_n)} + (1-\lambda)|b|^2$$

by convexity of the norm squared

Exercise: $\forall \lambda \in [0,1]$, $\forall y \in \mathbb{R}^n$, $|\lambda x + (1-\lambda) y|^2 \leq \lambda |x|^2 + (1-\lambda) |y|^2$ IV

Solut°: $|\lambda x + (1-\lambda) y|^2 = \lambda^2 |x|^2 + (1-\lambda)^2 |y|^2 + 2\lambda(1-\lambda) \langle x, y\rangle$

$$= \lambda |x|^2 + (1-\lambda)|y|^2 \underbrace{- \lambda(1-\lambda)|x|^2 - \lambda(1-\lambda)|y|^2 + 2\lambda(1-\lambda)\langle x,y\rangle}_{-\lambda(1-\lambda)|x-y|^2}.$$

$\square$

In the end, $\forall (a,b) \in \operatorname{Conv} \mathcal{S}(A(\omega) - z') \times \operatorname{Conv} \mathcal{S}(B - z')$,

$$d_c(z, A)^2 \leq (1-\lambda)^2 + \lambda |a|^2 + (1-\lambda)|b|^2$$

Taking the minimum:

$$d_c(z, A)^2 \leq (1-\lambda)^2 + \lambda\, d_c(z', A(\omega))^2 + (1-\lambda)\, d_c(z', B)^2$$

Idea 2: Just like in Chernoff / Cramer bounds, it is more efficient to minimize in λ AFTER averaging over $\mathbb{E}$.

Almost sure optimizat° is often of poor quality as it accounts for worst case scenarios.

Now if $X = (X', \omega)$

$$\mathbb{E}\left(e^{\frac{1}{4} d_c(X, A)} \,\middle|\, \omega \right)$$

$$\leq \mathbb{E}\left[e^{\frac{1}{4}(1-\lambda)^2} e^{\frac{\lambda}{4} d_c(X', A(\omega))^2} e^{\frac{1-\lambda}{4} d_c(X', B)^2} \right]$$

$$\leq e^{\frac{1}{4}(1-\lambda)^2} \mathbb{E}\left[e^{\frac{1}{4} d_c(X', A(\omega))^2} \right]^{\lambda} \mathbb{E}\left[e^{\frac{1}{4} d_c(X', B)^2} \right]^{1-\lambda}$$

Hölder $\frac{1}{p} = \lambda \quad q = 1-\lambda$

At this point, either $A(\omega)=\emptyset$ & take $\lambda=0$

$$\mathbb{E}\left(e^{\frac14 d_c(X,A)}\,\middle|\,\omega\right) \leqslant \mathbb{E}\left(e^{\frac14 + \frac14 d_c(X',B)^2}\right) \underset{\text{hyp}}{\overset{\text{induct°}}{\leqslant}} e^{\frac14}/\mathbb{P}(B) \leqslant 2/\mathbb{P}(B) \quad (\star)$$

or $A(\omega)\neq\emptyset$. Applying the induct° hypothesis:

$$\mathbb{E}\left(e^{\frac14 d_c(X,A)}\,\middle|\,\omega\right) \leqslant e^{\frac14(1-\lambda)^2}\,\mathbb{P}(A(\omega))^{-\lambda}\,\mathbb{P}(B)^{-(1-\lambda)}$$

$$= \frac{1}{\mathbb{P}(B)}\, e^{\frac14(1-\lambda)^2}\left[\underbrace{\frac{\mathbb{P}(A(\omega))}{\mathbb{P}(B)}}_{u\in[0,1]}\right]^{-\lambda}$$

$u\in[0,1]$ as $A(\omega)\subset B$

It is only now that we optimize in λ with $u\in[0,1]$ fixed.

$$\inf_{\lambda\in[0,1]} e^{\frac14(1-\lambda)^2}\,u^{-\lambda} \leqslant 2-u \quad \text{(Exercise)}$$

Hence
$$\mathbb{E}\left(e^{\frac14 d_c(X,A)^2}\,\middle|\,\omega\right) \leqslant \frac{1}{\mathbb{P}_{\Omega_n}(B)}\left(2-\frac{\mathbb{P}_{\Omega_n}(A(\omega))}{\mathbb{P}_{\Omega_n}(B)}\right)$$

⚠ works if $A(\omega)=\emptyset$

$$\Rightarrow \mathbb{E}\left(e^{\frac14 d_c(X,A)^2}\right)$$

$\longmapsto \int_{[0,1]}\mu(d\omega)\int_{\Omega_n}\mathbb{1}_{\{A(\omega)\}}\mu^{\otimes n}$

$$\leqslant \frac{1}{\mathbb{P}_{\Omega_n}(B)}\left(2-\frac{\mathbb{E}_\omega\,\mathbb{P}_{\Omega_n}(A(\omega))}{\mathbb{P}_{\Omega_n}(B)}\right)$$

$$\overset{\text{Fubini}}{=} \frac{1}{\mathbb{P}_{\Omega_n}(B)}\left(2-\frac{\mathbb{P}_{\Omega_n\otimes\mu}(A)}{\mathbb{P}_{\Omega_n}(B)}\right)$$

$$= \frac{1}{\mathbb{P}(A)}\underbrace{\frac{\mathbb{P}_{\Omega_{n+1}}(A)}{\mathbb{P}_{\Omega_n}(B)}\left(2-\frac{\mathbb{P}(A)}{\mathbb{P}(B)}\right)}_{\leqslant 1}$$

4 An equivalent definition of convex distance

If $\begin{cases} x \in \Omega_n = [0,1]^n \\ y \in \Omega_n \end{cases}$ then the Hamming distance

is $d_H(x,y) = \#\{ i \mid x_i \neq y_i \}$

$= \sum_{i=1}^{n} \mathbb{1}_{\{x_i \neq y_i\}}$

Weighted Hamming distance, with weight $\omega \in \mathbb{R}_+^n$:

$$d_\omega(x,y) := \sum_{i=1}^{n} \omega_i \mathbb{1}_{\{x_i \neq y_i\}}$$

Proposition: The convexe distance is the smallest distance dominating all d_ω ; $\omega \in \mathbb{R}_+^n$.

$$D(x,A) := \sup_{|\omega|_2 = 1} d_\omega(x,A) = d_c(x,A)$$

Proof: ($\leqslant$) For all ω, $|\omega|_2 = 1$,

$$d_\omega(x,A) = \inf_{y \in A} \sum_i \omega_i \mathbb{1}_{\{x_i \neq y_i\}}$$

$$= \inf_{s \in \mathcal{S}(A-x)} \langle \omega, s \rangle \quad (\text{def}^\circ \text{ of } \mathcal{S}(A-x))$$

$$= \inf_{s \in \operatorname{Conv} \mathcal{S}(A-x)} \langle \omega, s \rangle \quad (\text{inf of linear functo on convex set is reached on extreme points})$$

$$\leqslant \inf_{s \in \operatorname{Conv} \mathcal{S}(A-x)} |s|_2 |\omega|_2$$

$$= d_c(x,A)$$

Hence $\sup_{|\omega|_2 = 1} d_\omega(x,A) \leqslant d_c(x,A)$

$\geqslant$ Let $z \in \operatorname{Conv} \mathcal{S}(A-x)$ s.t $d_c(x,A) = |z|$

either $z = 0$ & then $D(x,A) = 0$

either $|z| > 0$: $\begin{cases} \omega := \frac{z}{|z|} \\ v \in \operatorname{Conv} \mathcal{S}(A-x) \end{cases}$

By convexity, $\forall \lambda \in [0,1]$, $\operatorname{conv} \mathcal{S}(A-x) \ni \lambda z + (1-\lambda) v$

& by z acheeving inf:

$$d_c(x,A)^2 = |z|^2 \leqslant |z + \lambda(v-z)|^2$$

$$= |z|^2 + 2\lambda\langle z, v-z\rangle + \lambda^2 |v-z|^2$$

$$\Rightarrow \quad \langle z, v-z\rangle \geqslant 0$$

First order deformato

$$\Leftrightarrow \quad |z|^2 \leqslant \langle z, v\rangle$$

$$\Leftrightarrow \quad |z| \leqslant \langle \omega, v\rangle$$

Hence $\forall v \in \operatorname{Conv} \mathcal{S}(A-x)$, $d_c(x,A) = |z| \leqslant \langle \omega, v\rangle$

Therefore $D(x,A) = \sup\limits_{|\omega|_2 = 1} d_\omega(x,A)$

$$\geqslant d_\omega(x,A) = \inf_{y \in A} \langle \omega, (\mathbb{1}_{\{x_i \neq y_i\}})_i\rangle$$

$$= \inf_{v \in \operatorname{Conv} \mathcal{S}(A-x)} \langle \omega, v\rangle \geqslant d_c(x,A)$$

$\blacksquare$