06/01/22

→ Nicolas J.

Distance de Wasserstein, Entropie et inégalité de Talagrand

I. Espace $P_p(X)$ et distance W_p de Wasserstein d'ordre $p \geq 1$

I.1. Inégalité de Hölder

Si $a, b > 0$ et $p \geq 1$

$(a+b)^p \leq 2^{p-1}(a^p + b^p)$ Cas spéciaux : $p=1$, $p=2$

Démo : $(a\times 1 + b\times 1) \leq (a^p + b^p)^{1/p} \underbrace{\left(1^{p/p-1} + 1^{p/p-1}\right)}_{=2}{}^{\frac{p-1}{p}}$ ■

I.2. Définition

Définition : Sur (X, d), on pose $P(X)$ l'espace des mesures de proba boréliennes.

Pour $p \geq 1$, on définit $P_p(X) = \left\{ \mu \in P(X), \int d(y, x_0)^p d\mu(y) < \infty \right\}$

Remarques : a) $\int d(y, x_0')^p d\mu \leq 2^{p-1} \underbrace{\int d(y, x_0)^p d\mu}_{<\infty} + 2^{p-1} \int d(x_0, x_0') d\mu$ si $\mu \in P_p(X)$.

b) $\left(\int d(y, x_0)^p d\mu \right)^{1/p} = \| d(\cdot, x_0) \|_{L^p(\mu)}$. (p-ième moment fini) $\Updownarrow$ $\mu \in P_p(X)$.

On a $P_1 \supset P_2 \supset \dots$

I.3. Fait.

Définition : $\mathcal{T}_p(\mu_0, \mu_1) = \min\limits_{\pi \in Marg(\mu_0, \mu_1)} \displaystyle\int_{X \times X} \underbrace{d(x,y)^p}_{= c(x,y)} \, d\pi(x,y)$

Si $\mu_0, \mu_1 \in \mathcal{P}_p(X)$, on a

$$\mathcal{T}_p(\mu_0, \mu_1) \leq 2^{p-1} \mathcal{T}_p(\mu_0, \delta_{x_0}) + 2^{p-1} \mathcal{T}_p(\delta_{x_0}, \mu_1).$$

I.4. Distance de Wasserstein

6 6 4 6 4 6 6 6

$$W_p(\mu, \mu') := \left(\min_{\pi \in Marg(\mu, \mu')} \int d(x,y)^p \, d\pi(x,y) \right)^{1/p}$$

Rq : $\mathcal{P}_p(X) = \left\{ \mu,\ W_p(\mu, \delta_{x_0}) < \infty \right\}$

II. Inégalité de Talagrand et entropie

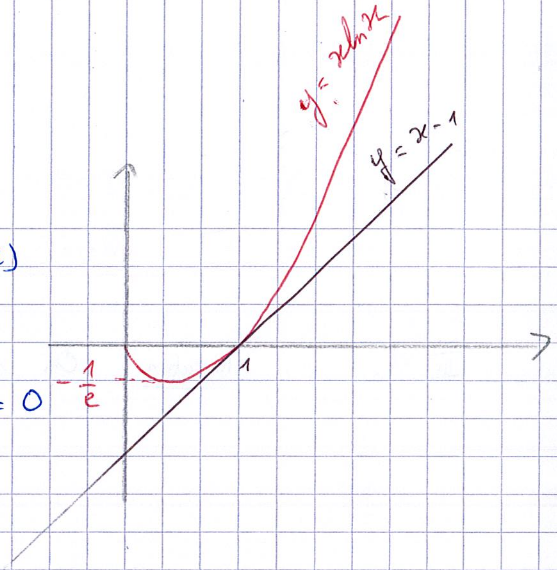
II.1 Espaces métriques mesurés

On dit qu'un espace métrique probabilisé $(X, \underset{\text{distance}}{d}, \underset{\text{mesure proba}}{\nu})$ ~~est~~ satisfait l'inégalité de Talagrand $(T_p(C))$ $p \geq 1$ si

$$\forall \mu \in \mathcal{P}(X),\quad W_p(\mu, \nu) \leq \sqrt{C \,.\, \underbrace{Ent(\mu | \nu)}_{\text{divergence de Kullback-Leibler}}},$$

avec $Ent(\mu | \nu) = \begin{cases} \displaystyle\int \rho \ln(\rho) \, d\nu = \int \ln(\rho) \, d\nu \in [0, +\infty] \text{ si } \mu \ll \nu,\ \rho = \dfrac{d\mu}{d\nu} \\ +\infty \text{ sinon.} \end{cases}$

Remarques : a) $x \in [0, +\infty] \xrightarrow{\phi} x\ln(x)$



Jensen $\Rightarrow$ $\displaystyle\int \phi(\rho)\, d\nu \geq \phi\Big(\underbrace{\int \rho\, d\nu}_{=1}\Big) = 0$

b) Si $p \leq q$, si (X, d, ν) satisfait à $(T_q(c))$, alors (X, d, ν) ——— $(T_p(c))$ aussi

II. 2. Propriétés

a) $\displaystyle Ent(\mu \times \tilde{\mu} \mid \nu \times \tilde{\nu}) = \int \rho(x)\tilde{\rho}(y) \underbrace{\ln[\rho(x)\tilde{\rho}(y)]}_{\ln \rho(x) + \ln \tilde{\rho}(y)} d\mu(x)\, d\tilde{\mu}(x).$

b) $dm = \overset{\text{densité de proba}}{f(x,y)}\, d(\nu \times \tilde{\nu})(x,y)$

$m(dx, dy) = \mu(dx)\,(id, m_x)\,dy$, ce qui veut dire :

$$\forall \Psi \in \mathcal{C}_0(X \times X, \mathbb{R}), \quad \iint \Psi(x,y)\, m(dx, dy)$$

$$= \iint \Psi(x,y) \frac{f(x,y)\, d\nu(x)\, d\tilde{\nu}(y)}{m}$$

$$= \int \underbrace{\Psi(x,y) \frac{f(x,y)}{g(x)} d\tilde{\nu}(y)}_{m_x(dy)} \cdot \underbrace{\int f(x,t)\, d\tilde{\nu}(t)}_{=g(x)}\, d\nu(x)$$

$d\mu(x)$

Alors $\displaystyle Ent(\underset{f(x,y)\,d\nu\,d\tilde{\nu}}{m} \mid \nu \times \tilde{\nu}) = Ent(\mu \mid \nu) + \int Ent(m_x \mid \tilde{\nu})\, d\nu.$

III. Concentration de la mesure

III.1.

<u>Notation</u> : On note, $A \subset X$ et $r \geq 0$, A_r l'ensemble

$$A_r = \{ y \in X, \ d(y, A) \leq r \}.$$

$$A \subset A_0 = \bar{A} \subset \underset{r>0}{A_r}$$

<u>Théorème</u> : Soit (X, d, ν) satisfaisant $(T_1(c))$, alors il existe C $(= \sqrt{2} ?)$ t.q. pour tout $A \subset X$ de ν-mesure $\geq \frac{1}{2}$ et tout $r \geq 0$,

(i) $\nu(A_r) \geq 1 - C \exp\left(-\frac{r^2}{2c}\right)$

(ii) De plus, si $f : X \to \mathbb{R}$ est 1-lip et $a \in \mathbb{R}_+$ t.q.

$$\nu\left(\{x \in X, \ f(x) \geq a\}\right) \geq \frac{1}{2}$$

$$\nu\left(\{x \in X, \ f(x) \leq a\}\right) \geq \frac{1}{2},$$

alors $\nu\left(\{x \in X, \ -r \leq f(x) - a \leq r\}\right) \geq 1 - C \exp\left(-\frac{r^2}{2C}\right)$

(iii) La dernière relation (ii) reste vraie en remplaçant a par $b := \int f \, d\nu$, en remplaçant $(2c, 2C)$ par $\left(8c, e^{c^2 \pi}\right)$.

III.2. Tensorisation de $(T_2(c))$

$$\left.\begin{array}{l} (X, d, \nu) \text{ vérifie } (T_2(c)) \\ (\tilde{X}, \tilde{d}, \tilde{\nu}) \text{ ———} \end{array}\right\} \Rightarrow \left(X \times \tilde{X}, \sqrt{d^2 + \tilde{d}^2}, \nu \otimes \tilde{\nu}\right) \text{ vérifie } T_2(c).$$