Cours M2Ri :

CV II, thm lim,

& transport optimal

Séance 12

27 Octobre 2021

$\boxed{\text{II}}$ Dualité de Kantorovich.

On s'interesse au pb de Kantorovich

$$\inf_{\pi \in \Pi(\mu,\nu)} \int c(x,y)\, \pi(dx,dy) = \inf_{\substack{\mathcal{L}(X)=\mu \\ \mathcal{L}(Y)=\nu}} \mathbb{E}\, c(X,Y)$$

où $\underset{\text{couplages}}{\Pi(\mu,\nu)} := \{\pi \in \mathcal{M}_1(\mathcal{X}\times\mathcal{Y}) \mid (P_1)_* \pi = \mu,\ (P_2)_* \pi = \nu\}$

$$= \{\pi \in \mathcal{M}_1(\mathcal{X}\times\mathcal{Y}) \mid \pi(A\times\mathcal{Y}) = \mu(A),\ \forall A \in \mathrm{Bor}(\mathcal{X}),\ \pi(\mathcal{X}\times B) = \nu(B),\ \forall B \in \mathrm{Bor}(\mathcal{Y})\}$$

Comme remarqué précédemment, c'est une optimisat° linéaire sous contraintes linéaires (Repenser au discret, très explicite). Plus généralement, nous avons un problème d'optimisat° convexe.

Thm [Dualité (convexe) de Kantorovich]

Soient $\begin{cases} \mathcal{X} \text{ et } \mathcal{Y} \text{ espaces polonais} \\ c : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}_+ \cup \{+\infty\} \text{ semi-continue inférieurement} \end{cases}$

Posons $I(\pi) := \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, \pi(dx,dy)$ pour $\pi \in \Pi(\mu,\nu)$

$J(\varphi,\psi) = \mu(\varphi) + \nu(\psi)$ pour $\varphi \in L^1(\mu)$, $\psi \in L^1(\nu)$

Alors $$\inf_{\pi \in \Pi(\mu,\nu)} I(\pi) = \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi,\psi)$$

où $\Phi_c = \{ (\varphi, \psi) \in L^1(\mu) \times L^1(\nu) \mid$
$\varphi(x) + \psi(y) \leq c(x,y), \ \forall x \in X \ \mu\text{-pp}$
$\forall y \in Y \ \nu\text{-pp} \}$

De plus l'infimum est atteint en un π^*

$$\inf_{\pi \in \Pi(\mu,\nu)} I(\pi) = I(\pi^*)$$

Raccourci pour $\mathcal{C}_b(X) \times \mathcal{C}_b(Y)$

et $$\sup_{(\varphi,\psi) \in \Phi_c} J(\varphi,\psi) = \sup_{(\varphi,\psi) \in \Phi_c \cap \mathcal{C}_b} J(\varphi,\psi)$$

↳ restreint aux continus bornés

Quelques dénominat°:

- En analyse convexe, les fonct° φ, ψ sont appelées variables duales, ou multiplicateurs de Lagrange.

- Dans la littérature du transport optimal, ils sont appelés "Potentiels de Kantorovich".

① Existence de minimiseurs π^*

Propositi°:

$$\exists \ \pi^* \in \Pi(\mu,\nu), \quad I(\pi^*) = \inf_{\pi \in \Pi(\mu,\nu)} I(\pi)$$

Preuve : Il s'agit de "la méthode directe" en calcul des variations :

① un argument de compacité,

② Semi-continuité inférieure.

① $\Pi(\mu,\nu)$ est compact pour la CV faible.

Trivialement fermé. Il suffit donc de pouvoir appliquer Prokhorov pour avoir la relative compacité. Mq $\Pi(\mu,\nu)$ tendu.

Comme les singletons dans $\mathcal{M}_1(E)$ sont tendus pour E polonais :

$$\forall \varepsilon>0, \quad \exists\, K_{\mathcal{X}} \subset \mathcal{X}, \quad \mu(K_{\mathcal{X}}^c) \leq \varepsilon$$
$$\exists\, K_{\mathcal{Y}} \subset \mathcal{Y}, \quad \nu(K_{\mathcal{Y}}^c) \leq \varepsilon.$$

Donc pour $K = K_{\mathcal{X}} \times K_{\mathcal{Y}}$ compact, on a :

$$\begin{aligned}
\forall \pi \in \Pi(\mu,\nu), \quad \pi(\mathcal{X}\times\mathcal{Y} \setminus K) &= \mathbb{E}_\pi\left(\mathbb{1}_{\{(X,Y)\notin K\}}\right) \\
&= \mathbb{E}_\pi\left(\mathbb{1}_{\{X\notin K \text{ ou } Y \notin K\}}\right) \\
&\leq \mathbb{E}_\pi\left(\mathbb{1}_{\{X\notin K_{\mathcal{X}}\}} + \mathbb{1}_{\{Y\notin K_{\mathcal{Y}}\}}\right) \\
&= \mu(K_{\mathcal{X}}) + \nu(K_{\mathcal{Y}}) \leq 2\varepsilon
\end{aligned}$$

Donc $\forall \varepsilon>0, \exists K, \quad \sup_{\pi\in\Pi(\mu,\nu)} \pi(K^c) \leq 2\varepsilon$

$\leadsto \Pi(\mu,\nu)$ tendu !

(2) I semi-continuité inférieurement ie

$$\left(\pi_n \to \pi \text{ faiblement} \Rightarrow I(\pi) \leq \liminf_{n\to+\infty} I(\pi_n)\right)$$

Soit $c_k \to c$ où c_k suite croissante continue et bornée

(Par ex, $c_k = k \wedge c$)

Alors
$$\int c_k \pi \overset{CV}{=} \lim_{n\to+\infty} \int c_k \pi_n \overset{\text{Monotonie}}{\leq} \liminf_{n\to+\infty} \int c\, \pi_n = \liminf_{n\to+\infty} I(\pi_n)$$

$\int c_k \pi \xrightarrow{\text{CV monotone}} \int c\pi$

Donc $I(\pi) \leq \liminf_{n\to+\infty} I(\pi_n)$

(3) Mise en musique :

Soit π_n une suite minimisante

pour $\inf_{\pi\in\Pi(\mu,\nu)} I(\pi) = \lim_{n\to+\infty} I(\pi_n)$

Par compacité, $\exists\, \varphi$ extract°, $\pi_{\varphi(n)} \to \pi^* \in \Pi(\mu,\nu)$

puis $I(\pi^*)$

$$\leq \liminf_{n\to+\infty} I(\pi_{\varphi(n)}) \quad \text{(s.c inférieure)}$$

$$= \inf_{\pi\in\Pi(\mu,\nu)} I(\pi) \leq I(\pi^*)$$

d'où le résultat $\boxtimes$

② Partie facile de la dualité de Kantorovich.

Propositi° :

$$\sup_{\Phi_c \cap \mathcal{C}_b} \mu(\varphi) + \nu(\psi) \leq \sup_{\Phi_c} \mu(\varphi) + \nu(\psi)$$

$$\leq \inf_{\pi \in \Pi(\mu,\nu)} I(\pi)$$

Preuve: $\mathcal{C}_b \subset L^1$ donc (première inégalité) triviale. Soit $\pi \in \Pi(\mu,\nu)$. Il suffit de mq

$$\sup_{\Phi_c} \mu(\varphi) + \nu(\psi) \leq I(\pi)$$

Par def de π : $\begin{cases} \int \varphi(x)\, \pi(dx,dy) = \mu(\varphi) \\ \int \psi(y)\, \pi(dx,dy) = \nu(\psi) \end{cases}$

Donc $\forall \varphi, \psi \in L^1(\mu) \times L^1(\nu)$

$$\mu(\varphi) + \nu(\psi) = \int \underbrace{(\varphi(x) + \psi(y))} \pi(dx, dy)$$

Lorsque $(\varphi,\psi) \in \Phi_c$, on a $\leq c(x,y)$! (*)

Donc $\sup_{(\varphi,\psi)\in\Phi_c} \mu(\varphi) + \nu(\psi) \leq \int c(x,y)\, \pi(dx, dy)$

$$= I(\pi).$$

⚠ Il manque que (*) soit vrai π-pp.

Or on a juste

$$\varphi(x) + \psi(y) \leq c(x,y) \quad \forall x \ \mu\text{-pp}, \ \forall y \ \nu\text{-pp}$$

$\overset{\text{def}}{\iff}$ (pas clair avant) $\exists N_{\mathcal{X}} \subset \mathcal{X}, \exists N_{\mathcal{Y}} \subset \mathcal{Y}$,

$\forall (x,y) \in N_{\mathcal{X}}^c \times N_{\mathcal{Y}}^c, \quad \varphi(x) + \psi(y) \leq c(x,y)$

Donc il suffit de mq $N = (N_{\mathcal{X}}^c \times N_{\mathcal{Y}}^c)^c$ est π-négligeable.

En effet

$$\begin{aligned} \pi((N_{\mathcal{X}}^c \times N_{\mathcal{Y}}^c)^c) &= \pi(N_{\mathcal{X}} \times \mathcal{Y} \cup \mathcal{X} \times N_{\mathcal{Y}}) \\ &\leq \pi(N_{\mathcal{X}} \times \mathcal{Y}) + \pi(\mathcal{X} \times N_{\mathcal{Y}}) \\ &= \mu(N_{\mathcal{X}}) + \nu(N_{\mathcal{Y}}) \\ &= 0. \end{aligned}$$

$\square$

③ Preuve formelle : Moralement juste, rigoureusement fausse.

Convent° : $\infty \times 0 = 0$

Mesures positives de masse finie.

$$\inf_{\pi \in \Pi(\mu,\nu)} I(\pi) \overset{\text{Astuce}}{=} \inf_{\pi \in \mathcal{M}_+(\mathcal{X} \times \mathcal{Y})} I(\pi) + \infty \mathbb{1}_{\{\pi \notin \Pi(\mu,\nu)\}}$$

Or $$\sup_{\varphi,\psi} \mu(\varphi) + \nu(\psi) - \int (\varphi(x) + \psi(y)) \, \pi(dx,dy) = \infty \mathbb{1}_{\{\pi \notin \Pi(\mu,\nu)\}}$$

Donc $\displaystyle\inf_{\pi\in\Pi(\mu,\nu)} I(\pi)$

$$= \inf_{\pi\in\mathcal{M}_+(X\times Y)} I(\pi) + \sup_{\varphi,\psi} \mu(\varphi)+\nu(\psi) - \int(\varphi(x)+\psi(y))\,\pi(dx,dy)$$

$$= \inf_{\pi\in\mathcal{M}_+(X\times Y)} \sup_{\varphi,\psi} \mu(\varphi)+\nu(\psi) - \int \pi(dx,dy)\,(\varphi(x)+\psi(y)-c(x,y))$$

⚠ Principe du minimax non-rigoureux.

$$= \sup_{\varphi,\psi}\ \inf_{\pi\in\mathcal{M}_+(X\times Y)} \overline{\qquad\qquad}$$

$$= \sup_{\varphi,\psi} \mu(\varphi)+\nu(\psi) - \underbrace{\sup_{\pi\in\mathcal{M}_+(X\times Y)} \int \overbrace{(\varphi(x)+\psi(y)-c(x,y))}^{\xi(x,y)}\,\pi(dx,dy)}_{(*)}$$

Le supremum vaut $+\infty$ dès que π charge

$\{(x,y) \mid \xi(x,y)>0\}$ ⇝ Contrainte.

Sous cette contrainte, π-pp, $\xi(x,y)\leq 0$

$\iff \varphi(x)+\psi(y)\leq c(x,y)$

et dans ce cas, le sup vaut 0 pour $\pi=0$.

Donc $(*) = \infty\,\mathbb{1}_{\{(\varphi,\psi)\notin\Phi_c\}}$

d'où $\displaystyle\inf_{\pi\in\Pi(\mu,\nu)} I(\pi) = \sup_{(\varphi,\psi)\in\Phi_c} \mu(\varphi)+\nu(\psi) - \cancel{\infty\,\mathbb{1}_{\{(\varphi,\psi)\notin\Phi_c\}}}$

4) Un minimax rigoureux.

Thm (Dualité de Legendre-Fenchel-Rockafellar)

Soit $\begin{cases} E \text{ un evn}, & E^* \text{ son dual topologique.} \\ F, G : E \to \mathbb{R} \cup \{+\infty\} & \text{fonct° convexes} \end{cases}$

Posons F^*, G^* = transformées de Legendre-Fenchel

ie $$F^* : \begin{array}{rcl} E^* & \longrightarrow & \mathbb{R} \cup \{+\infty\} \\ z^* & \longmapsto & \sup_{z \in E} \langle z^*, z \rangle - F(z) \end{array}$$

et supposons $\begin{cases} \exists z_0 \in E, & F(z_0) < +\infty, \ G(z_0) < +\infty \\ F \text{ continue en } z_0 \end{cases}$

Alors $$\inf_{z \in E} F(z) + G(z) = \max_{z^* \in E^*} \left[-F^*(-z^*) - G^*(z^*) \right]$$

<ins>Remarque</ins> : Le second membre vaut :

$$\max_{z^* \in E^*} -F^*(-z^*) - G^*(z^*)$$

$$= \max_{z^* \in E^*} \inf_{(x,y) \in E \times E} F(x) + G(y) - \langle z^*, y - x \rangle$$

C'est bien un minimax !

Preuve : Admise (cf Villani).

Utilise Hahn - Banach.