L2: Existence & some properties of optimal transport plans.

Weierstrass criterion for existence of minimizers (direct method in calc. of variations)

Prokhorov's thm & similar preliminaries

<u>Def</u> A function $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is said to be lower semi-continuous (lsc) if $\forall\, x_n \to x$ $\quad f(x) \le \liminf_{n\to\infty} f(x_n)$.

<u>Thm (1.)</u> If $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is lsc & $\mathcal{X}$ is compact then $\inf_{x\in\mathcal{X}} f(x)$ is attained by some $\bar{x} \in \mathcal{X}$.

<u>Proof</u> If $f \equiv +\infty$ we are done. Otherwise let $\ell = \inf_{x\in\mathcal{X}} f(x)$ $\left(\in \mathbb{R} \cup \{-\infty\}\right)$. Let x_n be a minimizing sequence. Pick a converging subsequence $x_{n_k} \to \bar{x}$. Then $\ell \le f(\bar{x}) \le \liminf f(x_{n_k}) = \ell$ so we have $=$, (& $\ell \in \mathbb{R}$).

<u>Def</u> We say that a sequence of measures $\mu_n \in \mathcal{P}(\mathcal{X})$ converges weakly (or narrowly) to $\mu \in \mathcal{P}(\mathcal{X})$ if
$$\int \varphi(x)\, \mu_n(dx) \to \int \varphi(x)\, \mu(dx) \quad \forall \varphi \in C_b(\mathcal{X}).$$
We write $\mu_n \Rightarrow \mu$ or $\mu_n \to \mu$.

<u>Rk</u> $\mathcal{X} = \mathbb{R}$ this is equivalent to $F_{\mu_n}(x) \to F_\mu(x)$ $\forall x$ point of continuity of F_μ.

<u>Def</u> A family of measures $\{\mu_i : i \in \mathcal{I}\}$ in $\mathcal{P}(\mathcal{X})$ is said to be tight if $\forall \varepsilon > 0$ $\exists K_\varepsilon \subseteq \mathcal{X}$ compact s.t. $\mu_i(\mathcal{X} \setminus K_\varepsilon) < \varepsilon$ $\forall i \in \mathcal{I}$.

<u>Thm (Prokhorov)</u> Let $\mathcal{X}$ be Polish & $(\mu_n)_{n\ge1} \subseteq \mathcal{P}(\mathcal{X})$. (μ_n) is relatively compact (i.e., $\forall (\mu_{n_k}) \subseteq (\mu_n)$, $\exists (\mu_{n_{k_l}})$, $\mu_{n_{k_l}} \to \mu$ for some $\mu \in \mathcal{P}(\mathcal{X})$) iff $(\mu_n)_{n\ge1}$ is tight.

<u>Rk</u> $\mu(\mathcal{X} \setminus K_\varepsilon) \le \liminf \mu_n(\mathcal{X} \setminus K_\varepsilon) < \varepsilon$ by portmanteau thm so $\overline{(\mu_n)}$ is also tight.
($\mathcal{X} \setminus K_\varepsilon$ open)

<u>Ideas of Proof</u> "$\Leftarrow$"
For a compact $K \subseteq \mathcal{X}$ we have $C_0(K) = C_b(K) = C(K)$ so the dual is the space of measures & (μ_n) is a bounded sequence so (by Banach–Alaoglu since $C_0(K)$ is separable) has a weakly converging subsequence: $\mu_{n_k}|_K \to \mu_K$. Take $K_{1/i} \subseteq K_{1/(i+1)} \subseteq \ldots$ & through a diagonal argument build one subsequence μ_{m_n} which converges weakly to some ν_i on $K_{1/i}$.
Let $\mu(A) := \sup_i \nu_i(A \cap K_{1/i})$. For $\varphi \in C_b(\mathcal{X})$, $\int_{\mathcal{X}} \varphi\, d(\mu_{m_n} - \mu) \le 2\|\varphi\|_\infty \cdot \frac{1}{i} + \int_{K_i} \varphi\, d(\ldots)$
& $\mu(\mathcal{X}) = 1$. $\qquad \to 2\|\varphi\|_\infty / i$ $\quad K_i$
so $\to 0$.

"$\Rightarrow$" $\forall r > 0$, we can cover $\mathcal{X}$ by open balls $B_1, B_2, \ldots$ of radius r.
Let $G_k = B_1 \cup \ldots \cup B_k$. Then $\liminf \mu_n(G_k) = 1$. (*)

Indeed, otherwise $\sup_k \mu_{n_k}(G_{n_k}) = c < 1$. Taking subsequence, $\mu_{n_{k_l}} \to \mu$ &

$\mu(G_m) \leq \liminf \mu_{n_{k_l}}(G_m) \leq \liminf \mu_{n_{k_l}}(G_{n_{k_l}}) = c < 1$. & $m \to \infty$ gives $1 = \mu(X) < \ldots$ contr.

Take $r = \frac{1}{m}$ & write G^m_k. $\forall \varepsilon > 0$, by (*), $\exists k_1, k_2, \ldots$ $\inf_n \mu_n(G^m_{k_m}) \geq 1 - \varepsilon 2^{-m}$

Let $A := \bigcap_m G^m_{k_m}$ then $\inf_n \mu_n(A) \geq 1-\varepsilon$ & $\bar{A}$ is complete & totally bounded $\Rightarrow$ compact

$\left(\mu_n(\bar{A}^c) \leq \mu_n(A^c) = \mu_n(\bigcup (G.)^c) \leq \sum_i \mu_n(G^m_{k_m}) \leq \varepsilon \sum_{i=1}^{\infty} 2^{-m} = \varepsilon\right)$

uses Ax. of C

(Can be covered by a finite union of balls of any given radius).

Existence of solutions to OT pb

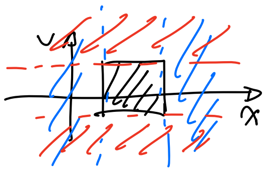
Lemma 21 Let $P \subseteq \mathcal{P}(\mathcal{X})$ & $Q \subseteq \mathcal{P}(\mathcal{Y})$ be tight. Then

$$\Pi(P,Q) = \{\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) : \mathrm{proj}_{\mathcal{X}} \# \pi \in P \ \& \ \mathrm{proj}_{\mathcal{Y}} \# \pi \in Q\} \text{ is tight.}$$

Proof. Fix $\varepsilon > 0$ & $K_{\mathcal{X}}, K_{\mathcal{Y}}$ compacts with ...

$\forall \pi \in \Pi(P,Q)$, $\pi(\mathcal{X} \times \mathcal{Y} \setminus K_{\mathcal{X}} \times K_{\mathcal{Y}}) \leq \pi(\mathcal{X} \times \mathcal{Y} \setminus K_{\mathcal{X}} \times \mathcal{Y}) + \pi(\mathcal{X} \times \mathcal{Y} \setminus \mathcal{X} \times K_{\mathcal{Y}})$

$= \mu(\mathcal{X} \setminus K_{\mathcal{X}}) + \nu(\mathcal{Y} \setminus K_{\mathcal{Y}}) < 2\varepsilon$.



Lemma 22 $\Pi(\mu,\nu)$ is compact.

Proof It is relatively compact by Prokhorov & the above lemma so we just have to establish closedness.

Let π be a limit of π_n. Then $\forall \varphi, \psi \in C_b$ $\int \varphi(x) + \psi(y) \, d\pi = \lim_n \int \varphi(x) + \psi(y) \, d\pi_n = \int \varphi(x) d\mu + \int \psi(y) d\nu$

so $\pi \in \Pi(\mu,\nu)$.

Rk We used that $C_b(\mathcal{X})$ determine elements in $\mathcal{P}(\mathcal{X})$. In fact one can construct a countable family of functions that does that.

Lemma 2.3 Suppose $c: \mathcal{X}\times\mathcal{Y} \mapsto \mathbb{R}\cup\{+\infty\}$ is lsc and bounded from below. Then

$$\pi \longmapsto \int c\, d\pi \quad \text{is lsc} \quad \text{on } \mathcal{P}(\mathcal{X}\times\mathcal{Y}) \text{ with topology of weak cv.}$$

Lemma 2.4. For $c: \mathcal{Z} \mapsto \mathbb{R}\cup\{+\infty\}$ bounded from below

$$c \text{ is lsc} \iff c(z) = \sup_k f_k(z) \text{ for a family } (f_k)_{k\geq 1} \text{ of Lipschitz functions on } \mathcal{Z}.$$

Proof (2.4).

$\Leftarrow$ ① $f_k(x) \leq \liminf_n f_k(x_n) \leq \liminf_n c(x_n)$ since $c \geq f_k$.

& taking $\sup_k$ $\quad c(x) \leq \liminf_n c(x_n)$.

(Rk) More generally a sup of lsc functions is lsc.

② (2^{nd} proof) c lsc $\iff$ the epigraph $\{(z,u): u \geq c(z)\}$ is closed in $\mathcal{Z}\times\mathbb{R}$

but —"— of sup $= \bigcap$ epigraphs.

$\Rightarrow$ Wlog $c \geq 0$.

Let $f_k(z) = \inf_{u\in\mathcal{Z}} \left(c(u) + k\, d(z,u)\right)$

- $|f_k(z_1) - f_k(z_2)| = \left| \inf_{u_1\in Z}\left(c(u_1) + k\, d(z_1,u_1)\right) - \inf_{u_2\in Z}\left(c(u_2) + k\, d(z_2,u_2)\right)\right|$ (wlog sup. ≥ 0)

$$\leq \inf_{u_1\in Z} c(u_1) + k\, d(z_1,u_1) - c(u_1) - k\, d(z_2,u_1)$$

$$= k \inf_{u\in Z} d(z_1,u_1) - d(z_2,u_1) \leq k\, d(z_1,z_2) \quad \text{is } k\text{-Lip.}$$

- $f_k \leq f_{k+1} \leq c$
- $\lim_k f_k(z) = \sup_k f_k(z) \leq c(z)$ Suppose the $=$ does not hold $\ell := \lim_k f_k(z) < c(z) \leq \infty$ for some $z\in\mathcal{Z}$

$\forall_k$ pick $u_k \in \mathcal{Z}$ s.t. $\quad c(u_k) + k\, d(u_k,z) < f_k(z) + 1/k \leq \ell + 1/k$

$$d(u_k,z) \leq \frac{\ell + 1/k - c(u_k)}{k} \leq \frac{\ell + 1/k}{k} \longrightarrow 0$$

taking limits in we get $c(z) \leq \liminf c(u_k) \leq \ell$ a contrad. ☒

Rk Taking $g_k = f_k \wedge k$ we may assume the sequence is of bdd functions

Proof (2.3)

We know we can take a sequence $c_n \nearrow c$ of Lip & bounded functions.

Then $\pi \mapsto J_n(\pi) = \int c_n\, d\pi$ is cont $\Rightarrow \int c\, d\pi \overset{MCT}{=} \lim J_n(\pi) = \sup J_n(\pi)$ is lsc

(by the Rk above)

☒

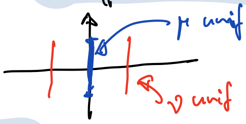
Thm 25 Let $\mathcal{X},\mathcal{Y}$ be Polish & $\mu\in\mathcal{P}(\mathcal{X})$, $\nu\in\mathcal{P}(\mathcal{Y})$ & $c:\mathcal{X}\times\mathcal{Y}\to\mathbb{R}$ lsc & bounded below. Then the Kantorovich OT problem is attained

$$\mathcal{P}(\mu,\nu)=\inf_{\pi\in\Pi(\mu,\nu)}\int c\,d\pi=\int c\,d\pi^*\quad\text{for some } \pi^*\in\Pi(\mu,\nu)$$

Proof We know that $\Pi(\mu,\nu)$ is compact, $\pi\mapsto\int c\,d\pi$ is lsc so we conclude by Weierstrass.

Rk We already noted that while $\Pi(\mu,\nu)$ is non-empty, the set of transports $\mathcal{T}(\mu,\nu)=\{\pi\in\Pi(\mu,\nu):\ \exists T:\mathcal{X}\to\mathcal{Y}\ \ \pi=(Id,T)_\#\mu\}$ may well be empty (e.g. $\mu=\delta_{\{0\}}$, $\nu=N(0,1)$).

Lemma If $\mu,\nu\in\mathcal{P}(\mathbb{R}^d)$ & μ is atomless then $\mathcal{T}(\mu,\nu)\neq\emptyset$. If μ,ν are supported on a compact then $\mathcal{T}(\mu,\nu)$ is dense in $\Pi(\mu,\nu)$ & for c continuous: $\inf_{\pi\in\mathcal{T}(\mu,\nu)}\int c\,d\pi=\min_{\pi\in\Pi(\mu,\nu)}\int c\,d\pi$.

Example: 

$c(x,y)=|x-y|^2 \Rightarrow \pi^*$ splits mass, but can be approximated with 

Extensions to MOT We consider here $\mathcal{Y}=\mathcal{X}$ & ONE step martingales.

Recall the $\mathcal{M}(\mu,\nu)=\{\pi\in\Pi(\mu,\nu):\ \mathbb{E}_\pi[Y|\sigma(X)]=X\ \ \mu\text{-a.s.}\}=\{\pi\in\Pi(\mu,\nu):\ \pi=\mu\otimes\theta\ \&\ \int y\,\theta(x,dy)=x\ \ \mu(dx)\text{-a.e.}\}$

This set may be empty. In fact:

Thm (Strassen) Let $\mu,\nu\in\mathcal{P}_1(\mathbb{R}^d)=\{\gamma\in\mathcal{P}(\mathcal{X}):\ \int|x|\,\gamma(dx)<\infty\}$

$$\mathcal{M}(\mu,\nu)\neq\emptyset\ \text{ iff }\ \mu\leq_{cx}\nu,\ \text{ i.e., }\ \int f\,d\mu\leq\int f\,d\nu\quad\forall f:\mathbb{R}^d\to\mathbb{R}\ \text{convex}$$

Lemma If $\mathcal{M}(\mu,\nu)\neq\emptyset$ then $\mathcal{M}(\mu,\nu)$ is compact.

Proof (for $\mathbb{R}$ i.e. $d=1$) Indeed, it is a subset of a compact set so we just need to prove it is closed. Recall that $\pi\in\Pi(\mu,\nu)$ belong to $\mathcal{M}(\mu,\nu)$ iff

$$\int_{\mathcal{X}\times\mathcal{Y}}\varphi(x)(y-x)\,d\pi=0\quad\forall\varphi\in C_b(\mathcal{X})$$

$$\parallel$$

$$\int_{\mathcal{X}}\varphi(x)\Big(x-\int_{\mathcal{Y}}y\,\theta(x,dy)\Big)\mu(dx)$$

Let $\pi_n\in\mathcal{M}(\mu,\nu)$ conv. weakly to $\pi\in\Pi(\mu,\nu)$. Fix $K>0$ & $f_K=\begin{cases}1 & \text{on } [-K,K]^2\\ 0 & \text{on } \mathbb{R}^2\setminus[-K-1,K+1]\end{cases}$ cont. bdd by 1.

$g_K=\varphi(x)(y-x)f_K(x,y)$ is C_b so $\int g_K\,d\pi_n\to\int g_K\,d\pi$

$$\forall_{\varepsilon>0}\ \exists K \text{ s.t. } \int |g - g_K| \, d\pi \le \int_{\mathbb{R}^2 \setminus [-K,K]^c} b\,(y+x)\, d\pi = b\left(\int_{[-K,K]^c} x \, d\mu + \int_{[-K,K]^c} y \, d\nu\right) \le \varepsilon$$

$\|\varphi\| \le b$

$$\Rightarrow \left|\int g \, d\pi\right| \le 3\varepsilon + \left|\int g \, d\pi_n\right| = 3\varepsilon \searrow 0 \quad \Rightarrow \pi \in \mathcal{M}(\mu,\nu)$$ ☒.

Rk The above extends to mg with finite discrete time $\mathcal{M}(\mu_1,\dots,\mu_n)$ but fails, e.g., in continuous time with $\mathcal{M}_{[0,1]}(\mu_0,\mu_1)$.

Rk Going back to OT, a restriction of an optimal tr. plan is still optimal:

Prop 2.6. In the setting of Thm 2.5, if π is a minimiser & $\pi' \le \pi$ is a non-negative measure with $\pi'(\mathcal{X}\times\mathcal{Y})>0$ then $\hat{\pi} := \dfrac{\pi'}{\pi'(\mathcal{X}\times\mathcal{Y})}$ is an optimal tr. plan for marginals $\hat{\mu} = \mathrm{proj}_{\mathcal{X}\#}\hat{\pi}$, $\hat{\nu} = \mathrm{proj}_{\mathcal{Y}\#}\hat{\pi}$.

Proof (Ex?)

If $\hat{\pi}$ not optimal then take minimiser $\bar{\pi}$, $\int c \, d\bar{\pi} < \int c \, d\hat{\pi}$

$\bar{\pi}, \hat{\pi} \in \Pi(\hat{\mu},\hat{\nu})$

Let $\tilde{\pi} := (\pi - \pi') + \pi'(\mathcal{X}\times\mathcal{Y}) \cdot \bar{\pi} = \pi + \pi'(\mathcal{X}\times\mathcal{Y}) \cdot (\bar{\pi} - \hat{\pi}) \in \Pi(\mu,\nu)$

& $\int c \, d\tilde{\pi} < \int c \, d\pi \Rightarrow$ contradiction. ☒

Some properties of the optimal solutions

A natural way to try to improve a given transport plan π is to consider if we can lower the cost via a cyclical relabelling of points.

Def (c-cyclical monotonicity) For $c: \mathcal{X} \times \mathcal{Y} \longmapsto \mathbb{R} \cup \{+\infty\}$ a subset $\Gamma \subseteq \mathcal{X} \times \mathcal{Y}$ is said to be c-cyclically monotone if $\forall_{N \in \mathbb{N}}$ $\forall_{(x_1, y_1), \dots, (x_N, y_N) \in \Gamma}$

$$\sum_{i=1}^{N} c(x_i, y_i) \le \sum_{i=1}^{N} c(x_i, y_{i+1}) , \quad \text{where } y_{N+1} = y_1 .$$

current cost in Γ

cost after cyclical re-routing

A transport plan $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ is said to be c-cyclically monotone if it is concentrated on a —, —, — set.

Intuition: $\pi^* \in \Pi(\mu, \nu)$ optimal $\Longrightarrow$ π^* is c-cyclically monotone

Insight from duality: $\Longleftarrow$ also holds.