

Recall: Given $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, $c: X \times Y \rightarrow [0, \infty]$,

$$(MP) = \inf \left\{ \int_X c(x, T(x)) d\mu(x) : T_{\#}\mu = \nu \right\}$$

set of plans between μ and ν

$$(KP) = \inf \left\{ \int_{X \times Y} c(x, y) d\gamma(x, y) : \gamma \in \Pi(\mu, \nu) \right\}$$

Some questions on existence and uniqueness

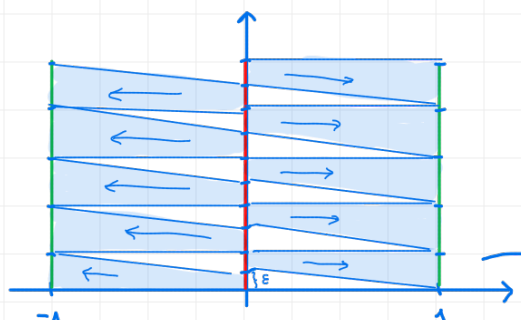
Ex. 1 (Non-existence of maps) In $\mathbb{R}$, $\mu = \delta_0$, $\nu = \frac{\delta_{-1} + \delta_1}{2}$.

There are no maps, and there is a unique plan:



Thus γ is the minimizer of (KP) for every cost.

Ex. 1.5 (Non-existence of maps) In $\mathbb{R}^2$, $\mu = \mathcal{L}_{1 \otimes \chi_{[0,1]}}$, $\nu = \frac{\mathcal{L}_{1 \otimes \chi_{[0,1]}} + \mathcal{L}_{1 \otimes \chi_{[0,1]}}}{2}$.



$$c(x, y) = |x - y|.$$

(MP) is not a min! for any map s.t. $T_{\#}\mu = \nu$,

$$|Tx - x| \geq 1 \text{ a.e.} \Rightarrow \text{cost}(T) = \int |Tx - x| d\mu \geq 1.$$

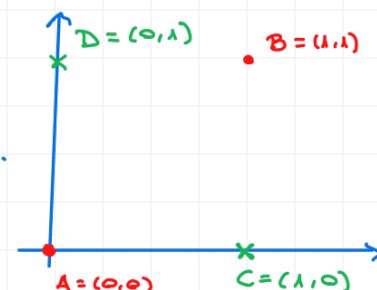
If this is T_ε , $\inf_\varepsilon \text{cost } T_\varepsilon = 1 = (MP)$.

There is no map realizing the inf, but there is a plan: the one splitting each point into two.

Ex. 2 (non-uniqueness of optimal plans)

$$X = Y = \mathbb{R}^2, \quad c(x, y) = |x - y|^2, \quad \mu = \frac{\delta_A + \delta_B}{2}, \quad \nu = \frac{\delta_C + \delta_D}{2}.$$

• There are two maps $\begin{cases} A \rightarrow C \\ B \rightarrow D \end{cases}$ & $\begin{cases} A \rightarrow D \\ B \rightarrow C \end{cases}$. Both are optimal



• A plan is given by: $\gamma_\alpha := \alpha \delta_{(A,C)} + (\frac{1}{2} - \alpha) \delta_{(A,D)} + (\frac{1}{2} - \alpha) \delta_{(B,C)} + \alpha \delta_{(B,D)} \in \mathcal{P}(\mathbb{R}^4)$.
for any $\alpha \in [0, 1/2]$.

$$\text{Cost } \gamma_\alpha = \int |x - y|^2 d\gamma_\alpha = 1 \quad \forall \alpha \in [0, 1/2].$$

All the 1-parameter family of plans is optimal.

Ex. 3 (book shuffling) in $\mathbb{R}$, $c(x, y) = |x - y|$

$$\mu = \frac{1}{M} \mathbb{1}_{[0, M]} \delta x \quad \nu = \frac{1}{M} \mathbb{1}_{[1, M+1]} \delta x$$



What is optimal? $M > 1$.

$$\textcircled{A} \quad T_1(x) = x + 1 \quad \begin{cases} x + M & x \in [0, 1] \\ x & x \in [1, M] \end{cases}$$

$$\textcircled{B} \quad T_2(x) = \begin{cases} x + M & x \in [0, 1] \\ x & x \in [1, M] \end{cases}$$

$$\textcircled{C} \quad T_3(x) = M + 1 - x$$

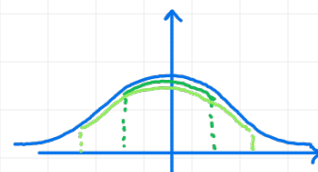
□ Preliminaries in measure theory

X locally compact, separable, metric space.

- $C_c(X)$: continuous, compactly supported functions on X .

↳ this is not closed under uniform convergence!

e.g. in $\mathbb{R}$, $f_m(x) = \frac{1}{1+x^2} \mathbb{1}_{[-m,m]} \xrightarrow{\text{unif.}} \frac{1}{1+x^2}$.



- $C_0(X)$: continuous, vanishing at infinity (closed under uniform convergence)
- $C_b(X)$: continuous, bounded functions.

Def: $\{\mu_k\} \subseteq \mathcal{M}(X)$. We say $\mu_k \xrightarrow{*} \mu$ (weakly-*) if

$$\int \varphi d\mu_k \rightarrow \int \varphi d\mu \quad \forall \varphi \in C_c(X).$$

↳ TV (Total Variation) of the measure.

Theorem (Banach-Alaoglu) $\|\mu_k\| \leq 1$, then up to a subsequence,

(blackbox)

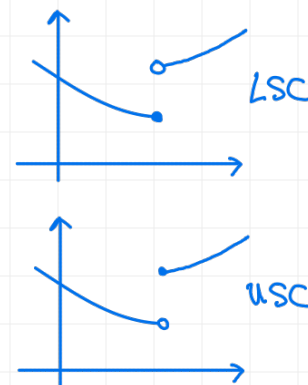
we have $\mu_k \xrightarrow{*} \mu$. (i.e., $\exists$ subseq. k_j s.t. $\mu_{k_j} \xrightarrow{*} \mu$).

[Remark: $(\mathcal{M}(X), \|\cdot\|_{TV})$ is the dual of $(C_0(X), \|\cdot\|_{L^\infty})$ (or $(C_c(X), \|\cdot\|_{L^\infty})$)
Related to the Riesz-Representation Theorem.]

Remark: $(\mu_k)_{k \in \mathbb{N}} \in \mathcal{P}(X)$, $\mu_k \xrightarrow{*} \mu$. Then μ may not be a probability.

Def: $(\mu_k)_{k \in \mathbb{N}} \in \mathcal{P}(X)$, μ_k narrowly converges to μ ($\mu_k \rightarrow \mu$) if one of the following equivalent definitions hold:

- ① $\int \varphi d\mu_k \rightarrow \int \varphi d\mu \quad \forall \varphi \in C_b(X)$.
- ② $\int \varphi d\mu_k \rightarrow \int \varphi d\mu \quad \forall \varphi \in \text{Lip}_b(X)$.
- ③ $\limsup_k \int \varphi d\mu_k \leq \int \varphi d\mu \quad \forall \varphi \text{ USC}, \varphi \leq c$.
- ④ $\liminf_k \int \varphi d\mu_k \geq \int \varphi d\mu \quad \forall \varphi \text{ l.s.c.}, \varphi \geq -c$.
- ⑤ $\liminf_k \mu_k(A) \geq \mu(A) \quad \forall A \text{ open}$.



Proof:

① $\Rightarrow$ ④ wlog. $C = 0 \leq \varphi$ (otherwise translate by a constant, $C \in \mathcal{C}_b(X)$).

inf-convolution strategy. For every $\lambda > 0$, define

$$\varphi_\lambda(x) := \inf_y \{ \varphi(y) + \lambda \overbrace{|x-y|}^{d(x,y) \text{ in metric space}} \}$$



with properties:

+ $\varphi_\lambda \leq \varphi(x)$, $\varphi_{\lambda'} \leq \varphi_\lambda \leq \varphi$ if $\lambda' < \lambda$

+ $\varphi_\lambda \uparrow \varphi$ as $\lambda \rightarrow \infty$ pointwise (idea: suppose x_λ minimum in the def. of $\varphi_\lambda(x)$, then $x_\lambda \rightarrow x$ as $\lambda \rightarrow +\infty$; exercise)

+ φ_λ is Lipschitz (λ -Lipschitz):

$$\varphi_\lambda(x') \leq \varphi(y) + \lambda|x'-y| \leq \varphi(y) + \lambda|x'-x| + \lambda|x-y|. \text{ Take inf in } y:$$

$$\varphi_\lambda(x') \leq \varphi_\lambda(x) + \lambda|x-x'|: \text{ by reversing roles of } x, x', \text{ we get Lipschitz.}$$

Define $\varphi_{\lambda, M}(x) := \min \{ \varphi_\lambda, M \} \leq \varphi$.

$$\liminf_k \int \varphi d\mu_k \geq \liminf_k \int \varphi_{\lambda, M} d\mu_k \stackrel{①}{=} \int \varphi_{\lambda, M} d\mu.$$

Take $M \rightarrow \infty$, $\lambda \rightarrow +\infty$, by the monotone convergence theorem we get ④.

Drawback: δ_m , $m \in \mathbb{N}$, does not converge narrowly as $m \rightarrow +\infty$.

Ex in $\mathbb{R}$, $\mu_k = (1 - \frac{1}{k})\delta_0 + \frac{1}{k}\delta_k \rightarrow \delta_0$ (there can be some "escape of mass" but it vanishes at infinity.)

Def: $A \subseteq \mathcal{P}(X)$ is tight if $\forall \varepsilon > 0$, $\exists K_\varepsilon \subset X$ compact s.t.

$$\mu(X \setminus K_\varepsilon) \leq \varepsilon, \forall \mu \in A.$$

Theorem (Prokhorov) $A \subseteq \mathcal{P}(X)$ is tight $\Leftrightarrow$ A is relatively compact for narrow convergence

($\rightarrow$ for every seq. $\{\mu_k\}_k \subset A$, there exists a converging subsequence).

Remark (How to verify tightness)

Assume $\exists \varphi: X \rightarrow [0, +\infty)$ s.t. $\{ \varphi \leq \lambda \}$ is compact $\forall \lambda \in \mathbb{R}$ and $\sup_{\mu \in A} \int \varphi d\mu < +\infty$.

Then A is tight.

Indeed, $\mu(\underbrace{X \setminus \{ \varphi \leq \lambda \}}_{\text{compactness used here}}) = \mu(\{ \varphi > \lambda \}) \leq \frac{1}{\lambda} \int \varphi d\mu \leq \frac{C}{\lambda}$

Markov / Chebyshev

Lemma $\odot$ Given $\mu \in \mathcal{P}(X)$:

- $\forall \varepsilon > 0, \exists K_\varepsilon \subset X$ compact s.t. $\mu(K_\varepsilon) \geq 1 - \varepsilon$ (μ is tight).
- $\forall \varepsilon > 0, \exists \eta_\varepsilon \in C_c(X), |\eta_\varepsilon| \leq 1$, s.t. $\int \eta_\varepsilon d\mu \geq 1 - \varepsilon$.

Proof in $\mathbb{R}^n$: $\rightarrow K_\varepsilon$ large enough ball

$\rightarrow \eta_\varepsilon$ cuts off K_ε , $\eta_\varepsilon = 1$ in $K_\varepsilon = B_R$, $\eta_\varepsilon = 0$ in B_{R+1}^c .

Proof (Prokhorov) : we prove only $\Rightarrow$.

- idea :
- Banach-Alaoglu applies if we restrict $\mu_j|_K$ for some sequence μ_j , K compact.
 - Do a diagonal argument filling X with increasing compacts K_n .

Since A tight, let $K_n \subset K_{n+1}$ compact s.t. $\mu(X \setminus K_n) \leq \frac{1}{n} \forall \mu \in A$ (*).

Given $(\mu_j)_j \subseteq A$, by Banach-Alaoglu $(\mu_j|_{K_n})_j$ has a converging

subsequence:

$$\mu_{j_1}|_{K_1}, \mu_{j_2}|_{K_1}, \dots \xrightarrow{*} \mu^{(1)} \text{ in } K_1$$

$$\mu_{j_1}|_{K_2}, \mu_{j_2}|_{K_2}, \dots \xrightarrow{*} \mu^{(2)} \text{ in } K_2$$

$\vdots$

repeat and
take the
diagonal

There exists a sequence j_i s.t. $\mu_{j_i}|_{K_n} \xrightarrow{*} \mu^{(n)}$ as $i \rightarrow \infty$ $\forall n \in \mathbb{N}$ (**)

- $\mu^{(n)}$ vanishes outside of K_n .
- $\mu^{(n)}(X \setminus K_m) \leq \frac{1}{m}$ by (*), $\forall n, m \in \mathbb{N}$, (***)
- $\mu^{(n+1)}|_{K_n^\circ} = \mu^{(n)} \quad (0)$
- $\mu^{(n)} \leq \mu^{(n+1)}$

Candidate for the limit μ_j is $\hat{\mu} := \sup_n \mu^{(n)}$

- From (0), $\hat{\mu}|_{K_n^\circ} = \mu^{(n)}|_{K_n^\circ} \quad \forall n$
- From (**), $\mu_{j_i}|_{K_n^\circ} \xrightarrow{*} \hat{\mu}|_{K_n^\circ}$ as $j \rightarrow +\infty$ (+)
- From (***), $\hat{\mu}(X \setminus K_m) \leq \frac{1}{m}$.

Take $\varphi \in \mathcal{C}_b(X)$:

$$\limsup_{i \rightarrow \infty} \left| \int_X \varphi d\mu_{j_i} - \int_X \varphi d\hat{\mu} \right| \leq$$

$$\begin{aligned} &\leq \limsup_n \limsup_i \left| \int_{X \setminus K_n^\circ} \varphi d\mu_{j_i} \right| + \left| \int_{X \setminus K_n^\circ} \varphi d\hat{\mu} \right| + \left| \int_{K_n^\circ} \varphi d\mu_{j_i} - \int_{K_n^\circ} \varphi d\hat{\mu} \right| \leq \\ &\leq \underbrace{\limsup_{n,i} \|\varphi\|_{L^\infty} \cdot \left(\mu_{j_i}(X \setminus K_n^\circ) + \hat{\mu}(X \setminus K_n^\circ) \right)}_{\frac{2}{n-1} \text{ (since } \mu_{j_i}(X \setminus K_n^\circ) \leq \mu_{j_i}(X \setminus K_{n-1}^\circ)} + \underbrace{\left| \int_{K_n^\circ} \varphi d\mu_{j_i} - \int_{K_n^\circ} \varphi d\hat{\mu} \right|}_{\rightarrow 0 \text{ as } i \rightarrow \infty, \forall n, \text{ by (+)}} \leq \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leq \varepsilon \quad (\forall \varepsilon > 0, \text{ take } n \text{ and } i \text{ large enough}) \quad \square \end{aligned}$$