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

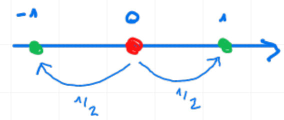
$$(MP) = \inf \left\{ \int_X c(x, T(x))\, d\mu(x) \;:\; T_\# \mu = \nu \right\}$$

$$(KP) = \inf \left\{ \int_{X\times Y} c(x,y)\, d\gamma(x,y) \;:\; \gamma \in \underbrace{\Pi(\mu,\nu)}_{\text{set of plans between } \mu \text{ and } \nu} \right\}$$

Some questions on existence and uniqueness

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

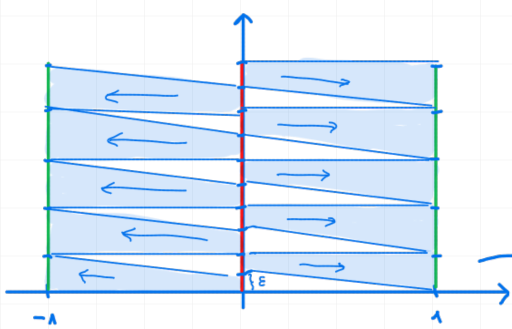
There are no maps, and there is a unique plan:



$$\gamma = \frac{1}{2}\delta_{(0,-1)} + \frac{1}{2}\delta_{(0,1)}.$$

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

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



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

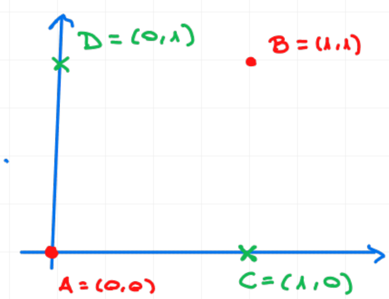
(MP) is not a min! for any map s.t. $T_\#\mu = \nu$,

$|Tx - x| \geq 1$ a.e. $\Rightarrow$ 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.

$\boxed{\text{Ex. 2}}$ (non-uniqueness of optimal plans)



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

- There are two maps $\begin{Bmatrix} A \to C \\ B \to D \end{Bmatrix}$ & $\begin{Bmatrix} A \to D \\ B \to C \end{Bmatrix}$. Both are optimal

- A plan is given by: $\gamma_\alpha := \alpha\,\delta_{(A,C)} + \left(\frac{1}{2} - \alpha\right)\delta_{(A,D)} + \left(\frac{1}{2}-\alpha\right)\delta_{(B,C)} + \alpha\,\delta_{(B,D)} \in \mathcal{P}(\mathbb{R}^4)$.

  for any $\alpha \in [0, 1/2]$.

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.

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

$$\mu = \frac{1}{M} \mathbb{1}_{[0,M]} \, dx \qquad \nu = \frac{1}{M} \mathbb{1}_{[1,M+1]} \, dx$$

What is optimal?

Ⓐ $T_1(x) = x+1$

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

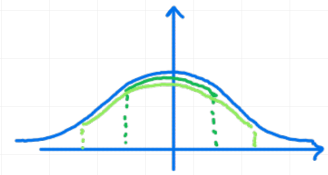
Ⓒ $T_3(x) = M+1-x$

☐ Preliminaries in measure theory

X locally compact, separable, metric space.

- $\mathcal{C}_c(X)$: continuous, compactly supported functions on X.

  ↳ this is not closed under uniform convergence!



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

- $\mathcal{C}_0(X)$: continuous, vanishing at infinity (closed under uniform convergence)
- $\mathcal{C}_b(X)$: continuous, bounded functions.

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

$$\int \varphi \, d\mu_k \longrightarrow \int \varphi \, d\mu \quad \forall \varphi \in \mathcal{C}_c(X).$$

<u>Theorem</u> (Banach-Alaoglu) (Blackbox) $\|\mu_k\| \leq 1$, then up to a subsequence, we have $\mu_k \xrightarrow{*} \mu$. (i.e., $\exists$ subseq. k_j s.t. $\mu_{k_j} \xrightarrow{*} \mu$).

→ TV (Total variation) of the measure.

[<u>Remark</u>: $(\mathcal{M}(X), \|\cdot\|_{TV})$ is the dual of $(\mathcal{C}_0(X), \|\cdot\|_{L^\infty})$ (or $(\mathcal{C}_c(X), \|\cdot\|_{L^\infty})$)
Related to the Riesz-Representation Theorem.]

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

<u>Ex</u>: $\delta_m \in \mathcal{P}(\mathbb{R})$, $\delta_m \xrightarrow{*} 0$ as $m \to \infty$ (since $\int \varphi \, d\delta_m = \varphi(m) \xrightarrow{\text{as } m\to\infty} 0$).

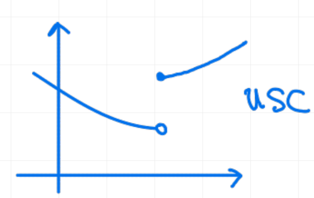
<u>Def</u>: $(\mu_k)_{k\in\mathbb{N}} \in \mathcal{P}(X)$, μ_k <u>narrowly converges</u> to μ ($\mu_k \to \mu$) if one of the following equivalent definitions hold:

① $\int \varphi \, d\mu_k \longrightarrow \int \varphi \, d\mu \quad \forall \varphi \in \mathcal{C}_b(X)$.

② $\int \varphi \, d\mu_k \longrightarrow \int \varphi \, d\mu \quad \forall \varphi \in Lip_b(X)$.

③ $\limsup_k \int \varphi \, d\mu_k \leq \int \varphi \, d\mu \quad \forall \varphi$ USC, $\varphi \leq C$.

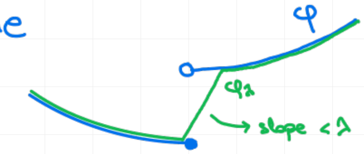
④ $\liminf_k \int \varphi \, d\mu_k \geq \int \varphi \, d\mu \quad \forall \varphi$ l.s.c., $\varphi \geq -C$.



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:

+ $\varphi_\lambda \overset{y=x}{\leq} \varphi(x)$, $\quad \varphi_{\lambda'} \leq \varphi_\lambda \leq \varphi$ if $\lambda' < \lambda$
+ $\varphi_\lambda \uparrow \varphi$ as $\lambda \to \infty$ pointwise (idea: suppose x_λ minimum in the def. of $\varphi_\lambda(x)$, then $x_\lambda \to x$ as $\lambda \to +\infty$; exercise)
+ φ_λ is Lipschitz (λ-Lipschitz):

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

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

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

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

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

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

Ex in $\mathbb{R}$, $\mu_k = \left(1 - \frac{1}{k}\right)\delta_0 + \frac{1}{k}\delta_k \longrightarrow \delta_0$ (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

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

Remark (How to verify tightness)

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

Then A is tight.

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

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

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

Proof in $\mathbb{R}^n$: $\longrightarrow$ K_ε large enough ball

$\longrightarrow$ η_ε 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 $\Longrightarrow$.

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

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

$$\mu_1|_{K_1}, \mu_2|_{K_1}, \dots \xrightarrow{*} \mu^{(1)} \text{ in } K_1$$

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

$$\vdots$$

repeat and take the diagonal $\downarrow$

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

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

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

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

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

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

$$\leq \limsup_n \limsup_i \left| \int_{X \setminus \mathring{K}_n} \varphi \, d\mu_{j_i} \right| + \left| \int_{X \setminus \mathring{K}_n} \varphi \, d\hat{\mu} \right| + \left| \int_{\mathring{K}_n} \varphi \, d\mu_{j_i} - \int_{\mathring{K}_n} \varphi \, d\hat{\mu} \right| \leq$$

$$\leq \underbrace{\limsup_{n,i} \|\varphi\|_{L^\infty} \cdot \overbrace{\left(\mu_{j_i}(X \setminus \mathring{K}_n) + \hat{\mu}(X \setminus \mathring{K}_n) \right)}^{\frac{2}{n-1} \text{ (since } \mu_{j_i}(X \setminus \mathring{K}_n) \leq \mu_{j_i}(X \setminus K_{n-1}))}}_{0} + \underbrace{\left| \int_{\mathring{K}_n} \varphi \, d\mu_{j_i} - \int_{\mathring{K}_n} \varphi \, d\hat{\mu} \right|}_{\hookrightarrow 0 \text{ as } i \to \infty \ \forall n, \text{ by } (+)} \leq$$

$$\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \leq \varepsilon \quad (\forall \varepsilon > 0, \text{ take } n \text{ and } i \text{ large enough}) \qquad \square$$