Recall: Formulation of optimal transport problem. X, Y complete, separable, locally cp. space

Monge's formulation. $\mu \in P(X)$. $\nu \in P(Y)$. Borel measu

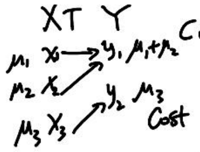
$T: X \to Y$ Borel map. $T\#\mu = \nu$. $C: X \times Y \to \mathbb{R} \cup \{+\infty$ bdd below.

minimize $\int_X c(x, Tx)\, d\mu(x)$

Kontorovitch's formulation: Look for $\gamma \in P(X \times Y)$. $(\mathrm{proj}_1)_\# \gamma = \mu$. $(\mathrm{proj}_2)_\# \gamma = \nu$.

wish to minimize $\int_{X \times Y} c(x, y)\, d\gamma(x, y)$.

Relation: If Monge's formulation has a solution $T: X \to Y$, then $\gamma = (id \times T)$ is a "candidate" for minimizer for Kontorovitch formulation



$C_{ij} = c(x_i, y_j)$. $1 \le i \le 3$, $1 \le j \le 2$.

Cost of $T = \mu_1 \cdot C_{11} + \mu_2 C_{21} + \mu_3 \cdot C_{32}$.

Kontorovitch: $\gamma \in P(X \times Y)$.

$\gamma_{ij} = \gamma(\{x_i\} \times \{y_j\})$.

Marginal condition:

$\gamma_{11} \quad \gamma_{21} \quad \gamma_{31} \longrightarrow \mu_1 + \mu_2$

γ.

$\gamma_{12} \quad \gamma_{22} \quad \gamma_{32} \longrightarrow \mu_3$

$\downarrow \mu_1 \quad \downarrow \mu_2 \quad \downarrow \mu_3$

Minimizing set: $\gamma_{ij} \ge 0$.

$\gamma_{11} + \gamma_{21} = \mu_1$. $\gamma_{21} + \gamma_{22} = \mu_2$

$\gamma_{31} + \gamma_{32} = \mu_3$

$\gamma_{11} + \gamma_{21} + \gamma_{31} = \mu_1 + \mu_2$

$\gamma_{12} + \gamma_{22} + \gamma_{32} = \mu_3$ $\times$

Total cost: $\sum_{i,j} \gamma_{ij} C_{ij}$.

Kontorovitch cost $= \min_\gamma \left(\sum_{i,j} \gamma_{ij} C_{ij} \right)$.

If you choose $C_{11}, C_{21}, C_{32} >> 1$.

then Monge cost $>$ Kontorovitch cost.

Thm: There exists $\gamma_0 \in \Pi(\mu, \nu)$. s.t

$I[\gamma_0] = \inf_{\gamma \in \Pi(\mu,\nu)} I[\gamma]$.

Skech of proof: Let γ_n be a minimizing sequence $\exists$ subseq $\gamma_{n_j} \longrightarrow \gamma$.

since $\gamma_{n_j} \in \Pi(\mu, \nu)$. $\gamma \in \Pi(\mu, \nu)$ as well.

Moreover $\int c(x,y)\, d\gamma \le \liminf_{j \to \infty} \int c(x,y)\, d\gamma_{n_j}$

So γ is a minimizer.

Some examples:

(1) $\mu \in P(X)$. $\nu = \delta_a$ $a \in Y$.

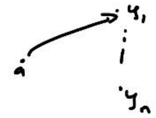
Then $\Pi(\mu, \nu) = \mu \otimes \delta_a = \gamma = (id \times T)\# \mu$.

$T: X \to Y$ the const map.

(2) $\mu = \delta_a$. $Y = \{y_1, \cdots y_n\}$ $\nu = \frac{1}{n} \sum^n \delta$

then $\Pi(\mu, \nu) = \{\frac{1}{n} \sum_{i=1}^n \delta_{(a,}$

There doesn't exist $T: X \to Y$. s.t T



$\nu(A)$

(3) $\mu = \mu_1 \delta_{x_1} + \mu_2 \delta_{x_2} +$

$\nu = (\mu_1 + \mu_2) \delta_{y_1} + $

Monge's formulation:

Need to find $T: \{x_1, x_2$ s.t $T\#\mu = \nu$.

T uniq...

$\underline{Q}$: When is Monge's formulation equivalent to Kantorovich?
How to characterize the minimizer?

A Duality Theorem: X, Y complete, separable, locally cpt metric space
$c: X \times Y \to \mathbb{R} \cup \{+\infty\}$ l.s.c & bdd below

(1). $\pi \in \Pi(\mu, \nu)$ define $I[\pi] = \int_{X \times Y} c(x,y) d\pi(x,y)$.

(2) $(\varphi, \psi) \in C_b(X) \times C_b(Y)$. with $\varphi(x) + \psi(y) \leq c(x,y) \rightsquigarrow \Phi_c$

Define $J(\varphi, \psi) = \int_X \varphi \, d\mu + \int_Y \psi \, d\nu$.

Then $\inf_{\gamma \in \Pi(\mu,\nu)} I[\pi] = \sup_{(\varphi,\psi) \in \Phi_c} J(\varphi, \psi)$.

Moreover, the minimizer to (1) always exists. If X, Y cpt, the (2) has a maximizer.

For π_0- a.e (x_0, y_0).
$\nabla \varphi_0(x_0) = \nabla_x c(x_0, y_0)$.

ASSUME: $\forall x \in \Omega_1$. $y \mapsto \nabla_x c(x,y)$ invertible.
Then we can find $T(x_0)$. s.t
$\nabla \varphi_0(x_0) = \nabla_x c(x_0, T(x_0))$.
i.e for π_0- a.e (x_0, y_0). $y_0 = T(x_0)$
$\pi_0 = (id \times T)_\# \mu$.

ASSUME FURTHER: $\forall y \in \Omega_2$. $x \mapsto \nabla_y c(x,y)$ invertible.
$x_0 = S(y_0)$ a.e. $S(y_0)$ solves $\nabla \psi_0(y_0) = \nabla_y c(x_0, y_0)$.

We can also write down a MA type equation as follows.
optimal map $T: \Omega_1 \to \Omega_2$. which needs to satisfy

$$\nabla \varphi_0(x) = \nabla_x c(x, T(x))$$

Using the push-forward condition of $T_\#(f dx) = g dy$

$$\det DT = \frac{f(x)}{g(T(x))}.$$

$\int \varphi(y) g(y) dy \underset{y=Tx}{=} \int \varphi(Tx) g(Tx) \det DT dx$

T can be represented in terms of $\nabla \varphi_0$ $= \int \varphi(T(x)) \cdot f(x) dx$

A formal proof (using a min-max principle)

$\inf \geq \sup$ is obvious.

Need to show $\inf \leq \sup$.

$$LHS = \inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = \inf_{\pi \in M_+(\mu,\nu)} \left(I(\pi) + \begin{cases} 0 & \text{if } \pi \in \Pi(\mu,\nu) \\ +\infty & \text{if } \pi \notin \Pi(\mu,\nu) \end{cases} \right)$$

$$\begin{cases} 0 \\ +\infty \end{cases} = \sup_{(\varphi,\psi) \in C_b(X)\times C_b(Y)} \left(\int \varphi \, d\mu + \int \psi \, d\nu - \int (\varphi(x)+\psi(y)) \, d\pi(x,y) \right)$$

$$= \inf_{\pi \in M_+(\mu,\nu)} \sup_{\substack{(\varphi,\psi) \\ \in C_b(X)\times C_b(Y)}} \left(\int c(x,y) \, d\pi + \int \varphi \, d\mu + \int \psi \, d\nu - \int (\varphi(x)+\psi(y)) \, d\pi(x,y) \right)$$

$$\overset{?}{=} \sup \inf_{\pi \in M_+(X\times Y)} \left(\qquad \right)$$

$$\overset{?}{=} \sup \inf_{\pi \in M_+} \left(\int \varphi \, d\mu + \int \psi \, d\nu + \int (c(x,y) - (\varphi(x)+\psi(y))) \, d\pi \right).$$

$$\inf_{\pi \in M_+} \int (c(x,y) - (\varphi(x)+\psi(y))) \, d\pi = \begin{cases} 0 & \text{if } \underbrace{\varphi(x)+\psi(y) \leq c(x,y)} \\ -\infty & \text{if not true.} \end{cases}$$

say at (x_0,y_0).
$\pi = \lambda \delta_{(x_0,y_0)}$
$\lambda \to +\infty$

Villani: Topics in optimal transport.

<u>Q</u>: How does the duality theorem help with characterizing the optimal transport plan?

Assume $X = \Omega_1$, $Y = \Omega_2 \subseteq \mathbb{R}^n$.

Let π_0 be minimizer to (1).

$(\varphi_0(x), \psi_0(y))$ be maximizers.

$c(x,y) \geq \varphi_0(x) + \psi_0(y) \rightsquigarrow (\varphi_0,\psi_0) \in \Phi_c$

Moreover $\int c(x,y) \, d\pi_0(x,y) = \int \varphi_0(x) \, d\mu + \int \psi_0(y) \, d\nu(y)$.

$$= \int (\varphi_0(x) + \psi_0(y)) \, d\pi_0(x,y)$$

So $c(x,y) = \varphi_0(x) + \psi_0(y)$ π_0-a.e.

WLOG, can assume φ_0, ψ_0 Lipschitz (will see later).

Let $(x_0,y_0) \in \operatorname{supp} \pi_0$ then $c(x_0,y_0) = \varphi_0(x_0) + \psi_0(y_0)$.

$c(x,y_0) \geq \varphi_0(x) + \psi_0(y_0)$

So $\underline{\nabla_x c(x_0,y_0) = \nabla \varphi_0(x_0)}$

eg: $c(x,y) = |x-y|^2$

$2(x_0 - y_0) = \nabla \varphi_0(x_0)$