$\begin{cases} \text{Liu Jiakun}, 2022 \\ \text{Schwanen Konig}. \end{cases}$

Introduction on Optimal transport

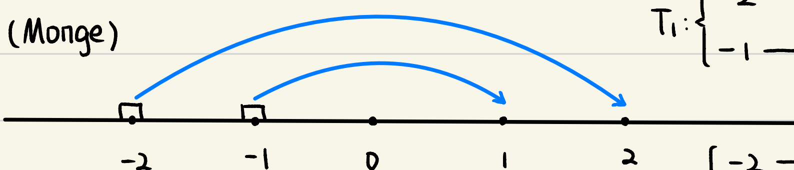
Some Big names $\begin{cases} \text{Monge. 1781 Monge problem} \\ \text{Kantorovich. (key figure in this seminar) 1942 Nobel prize in Economics} \\ \text{Yau. complex Monge-Ampére} \Rightarrow \text{Calabi conjecture 1976 Fields medal} \\ \text{Villani. 2010 Fields medal (boltzmann \& OT)} \\ \text{Figalli. 2018 Fields medal (Geometric measure theory \& OT)} \end{cases}$

[Ref]. 1) C. Villani, Optimal Transport: Old and New. Springer Berlin, Heidelberg, 2008.

2) F. Santambrogio. Optimal Transport for applied Mathematicians, Vol. 87 of Progress in Nonlinear Differential Equations and their Applications Birkhäuser, Cham, 2015

Discrete Eg 1. (Monge)

$T_1: \begin{cases} -2 \longrightarrow 1 \\ -1 \longrightarrow 2 \end{cases}$



$T_2: \begin{cases} -2 \longrightarrow 2 \\ -1 \longrightarrow 1 \end{cases}$

$X=\{-2,-1\}$ $Y=\{1,2\}$ $T: X \to Y$ transport map

cost function $c(x,y)=|x-y|^p$ $0<p<1$ concave $p=1$ linear $p>1$ convex

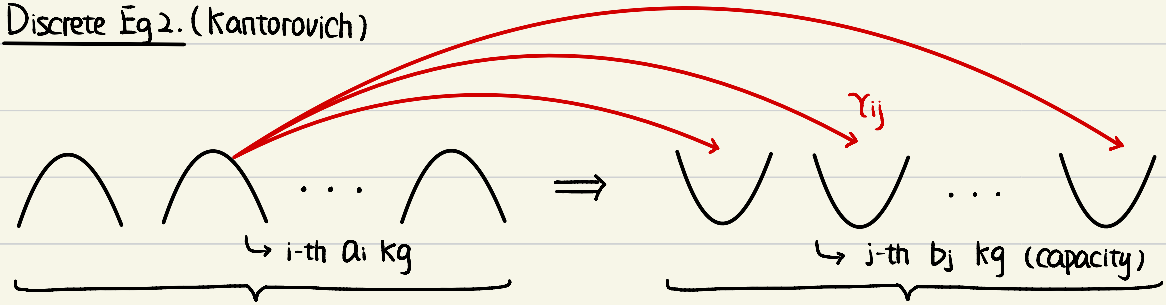
$p=1$ $C(T_1)=3+3=6=2+4=C(T_2)$ $\Rightarrow$ no strictly minimum

$p=\frac{1}{2}$ $C(T_1)=\sqrt{3}+\sqrt{3}>\sqrt{2}+\sqrt{4}=C(T_2)$ $\Rightarrow$ T_2 a "local" minimizer

$p=2$ $C(T_1)=3^2+3^2<2^2+4^2=C(T_2)$ $\Rightarrow$ T_1 a "global" minimizer

- (strictly) convex cost is more proper for optimal transport

Discrete Eg 2. (Kantorovich)



$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$ (total sand) $\quad \gamma_{ij} \Rightarrow$ sand transport from i-th piles to j-th pits

$a_i = \sum_{j=1}^{n} \gamma_{ij} \qquad b_j = \sum_{i=1}^{m} \gamma_{ij} \qquad c_{ij} \Rightarrow$ cost of transport unit mass from i to j

minimizing the total cost of the transport $\quad W = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\gamma_{ij}$

$$\min \quad W = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\gamma_{ij}$$

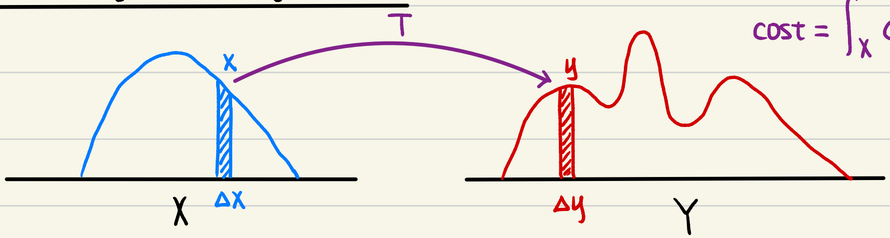
$$\text{s.t.} \quad \sum_{j=1}^{n} \gamma_{ij} = a_i \qquad i = 1, 2, \cdots, m$$

$$\sum_{i=1}^{m} \gamma_{ij} = b_j \qquad j = 1, 2, \cdots, n$$

Generalized Eg 1. $\Rightarrow$ Monge Problem

$c(x, y): X \times Y \to [0, +\infty]$

$\text{cost} = \int_X c(x, T(x))\, d\mu$



- transportation : from (X, μ) to (Y, ν) measure space $\quad T: (X, \mu) \to (Y, \nu)$

  measurable

  one-to-one.

lower semi continuous. l.s.c.

$f \qquad x_n \to x \qquad f(x) \le \liminf_{n\to\infty} f(x_n)$

u.s.c. $\quad f(x) \ge \limsup_{n\to\infty} f(x_n)$.

- cost function : $c: X \times Y \longrightarrow [0,+\infty]$ continuous / semi-continuous

$\Rightarrow X=Y=\Omega \subseteq \mathbb{R}^d \quad c(x,y)=h(x-y) \quad h \Rightarrow$ strictly convex

- measure preserving : pushfoward $T^{\#}\mu(A)=\mu(T^{-1}(A)) \quad \forall A \subseteq Y$ measurable

↪ if $T^{\#}\mu=\nu$, we call $T:(X,\mu)\to(Y,\nu)$ a measure preserving map

Problem (Monge Problem)

Given (X,μ) (Y,ν) and cost function $c: X\times Y \longrightarrow [0,+\infty]$, solve the problem

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

- Original Monge problem : $c(x,y)=|x-y|$ hard ! Evans. Gangbo 1999

but when $c(x,y)=h(|x-y|)$ strictly convex (i.e. $c(x,y)=|x-y|^p$, $p>1$) relatively easy

(Calculus of variations method)

To minimize functional $I[f]$, $f\in H$

(Calculus of variation) choose some ϕ, s.t. $f+t\phi \in H$ $(0<|t|<\varepsilon,\ t\in\mathbb{R}$ small$)$

If f is minimizer $\Rightarrow \frac{d}{dt} I[f+t\phi]\big|_{t=0}$ should be 0 for $\forall\phi$ admissable

Example. (Dirichlet energy and Laplace equation)

$\{u \mid \int_\Omega |\nabla u|^2 dx < \infty\}$

$u: \Omega(\subseteq\mathbb{R}^n) \longrightarrow \mathbb{R} \qquad \min\left\{ I[u]=\int_\Omega |\nabla u|^2 dx \;\middle|\; u\in \dot{H}^1(\Omega),\ u=g\in L^2(\partial\Omega) \text{ on } \partial\Omega \right\}$

$H=\{u \mid u\in \dot{H}^1(\Omega), u=g \text{ on } \partial\Omega\} \qquad u\in H \quad \forall \varphi\in C_c^\infty(\Omega) \quad u+t\varphi\in H \quad \forall t\in\mathbb{R}$

if u is minimizer of $I[u]$ $\Rightarrow 0=\frac{d}{dt}\big|_{t=0} I[u+t\varphi] = \frac{d}{dt}\big|_{t=0} \int_\Omega |\nabla u|^2 + 2t\nabla u\cdot\nabla\varphi + t^2|\nabla\varphi|^2$

$= -\int_\Omega \Delta u\,\varphi \quad \forall\varphi\in C_c^\infty(\Omega) \qquad \overset{\shortparallel}{\int_\Omega \nabla u\cdot\nabla\varphi}$.

Hence the minimizer u satisfies $\begin{cases} \Delta u=0 & \text{in } \Omega \\ u=g & \text{on } \partial\Omega \end{cases}$

$I(T).$ $\quad$ T minimizer.

$\quad\quad\quad\quad\quad\quad \frac{d}{dt} I(T_t)|_{t=0} = 0$

$\shortparallel$

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

- How to disturb T such that $T\circ\phi^{-1}$ is measure preserving

$\Rightarrow$ we need a measure preserving (diffeomorphism) family $\phi_t : X \to X$

The idea is obtained from the incompressible condition in fluid mechanics.

$(x_1, \cdots, x_n)$ $\quad\quad\quad\quad \sum_{i=1}^{n} \frac{\partial X^i}{\partial x^i} = 0$

<u>Lemma</u>: Let $\vec{X} = (X^1, \cdots, X^n) \in C_c^\infty(U;\mathbb{R}^n)$ $(U \subseteq \mathbb{R}^n)$ be a vector field with $\operatorname{div} X = 0$

Then the flow $\vec{\Phi}_t = (\phi^1, \cdots, \phi^n)$ generated by $\vec{X}$

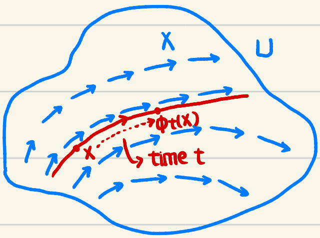
$$\begin{cases} \vec{\Phi}(x,0) = x & \text{for } x \in U \quad (\vec{\Phi}|_{t=0} = Id) \\ \frac{d\vec{\Phi}}{dt}(x,t) = \vec{X}(\vec{\Phi}(x,t)) & \text{in } U \times (-\infty, \infty) \end{cases}$$

is a Lebesgue measure preserving diffeomorphism $\vec{\Phi}(\cdot,t) : U \to U$ $\forall t$

<u>proof.</u> Change of variable formula $\int_U f(\phi(y))\,dy = \int_U f(y)\,dy = \int_U f(\phi(x))\,d\phi(x)$

$\Rightarrow$ hence $\vec{\Phi}$ is measure preserving if $\det D\vec{\Phi} = 1$

$\vec{\Phi}(\cdot,0) = Id$ $\quad$ when $t=0$ $\quad \det D\vec{\Phi} = 1$



Aim is $\det D\vec{\Phi} = 1$ $\forall t$.

$$\frac{d}{dt}(\det D\vec{\Phi}) = A^{ij}(D\vec{\Phi}) \frac{d}{dt}[D_i \phi_j] \quad = \det(D\vec{\Phi}) \operatorname{div}(X) = 0$$

$$= A^{ij}(D\vec{\Phi}) D_i(X_j) = A^{ij}(D\vec{\Phi}) \frac{\partial X_j}{\partial \phi_k} D_i \phi_k$$

$$= \det(D\vec{\Phi})\, \delta_k^j \frac{\partial X_j}{\partial \phi_k} = 0 \quad (\text{by div-free})$$

▨

- T measure preserving. $T_t = T\circ\vec{\Phi}_t^{-1}$ measure preserving

$$\frac{d}{dt}\vec{\Phi}_t = X(\phi_t)$$

$$\begin{aligned}0 &= \frac{d}{dt}\Big|_{t=0} \int_U c(x, \underbrace{T\circ\vec{\Phi}_t^{-1}}_{T_t}(x))\, d\mu(x) \qquad z = \vec{\Phi}_t^{-1}(x) \\ &= \frac{d}{dt}\Big|_{t=0} \int_U c(\vec{\Phi}_t(z), T(z))\, d\mu(z) \\ &= \int_U D_x c(z, T(z)) \frac{d}{dt}\vec{\Phi}_t \, d\mu(z) = \int_U \underbrace{D_x c(z, T(z))}_{D\psi(z)} \cdot \vec{X}(z)\, dz \qquad \forall \operatorname{div}\vec{X} = 0\end{aligned}$$

$\int_U D\psi(z)\cdot\vec{X}(z)\,dz$ $\operatorname{div}\vec{X} = 0$.

(integration by parts)

$D_x c(z, T(z)) + D_z\psi(z) = 0 \Rightarrow$ solve T. we need $(D_x c)(x,\cdot)$ injective c strictly convex

$T \sim D\psi$ ↳ where strictly convex is needed

Appell's work 1887 Bordin Prize French Academy of Science

Generalized Eq 2. $\Longrightarrow$ Kantorovich problem

Def. (Coupling)

(X,μ) (Y,ν) two probability space $\Omega = X\times Y$, $\pi_X : X\times Y \to X$ $(x,y)\longmapsto x$

$\pi_Y : X\times Y\to Y$ $(x,y)\longmapsto y$ coupling of (μ,ν) is definded by   (marginal)

$$\Pi[\mu,\nu] = \{\gamma \text{ is probability measure} \mid (\pi_X)^{\#}\gamma = \mu,\ (\pi_Y)^{\#}\gamma = \nu\}$$

Rmk. ① $\gamma(A\times B) \Rightarrow$ probability of transport A to B (in discrete case $\Longrightarrow \gamma_{ij}$)

② $\forall A\subseteq X$, $B\subseteq Y$ measurable, then $\gamma(A\times Y) = \mu(A)$, $\gamma(X\times B) = \nu(B)$

$A = \{i\}$ $B = \{j\}$.   total mass of A is transported to Y ↳ …

$X = \{1,\cdots,m\}$ $Y = \{1,\cdots,n\}$. $\sum_{j=1}^{n}\gamma_{ij} = a_i$ $\sum_{i=1}^{m}\gamma_{ij} = b_j$

③ all measurable func. φ, ψ on X, Y $\Rightarrow \int_{X\times Y}(\varphi(x)+\psi(y))\,d\gamma(x,y) = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu$

$\gamma\in\Pi[\mu,\nu]$ $\varphi = \chi_A(x)$ $\psi = 0$.

$\varphi = 0$ $\psi = \chi_B(y)$

$\dot{x} \quad \dot{y}.\ \int_{X\times Y} c(x,y)\, d\gamma(x,y).$

Problem (Kantorovich problem)

Given (X,μ) (Y,ν) and cost function $c: X\times Y \longrightarrow [0,+\infty]$. solve the problem

$$\inf_{\gamma\in\Pi[\mu,\nu]} G_K(\gamma) = \inf\left\{\int_{X\times Y} c(x,y)\, d\gamma(x,y) \,\middle|\, \gamma\in\Pi[\mu,\nu]\right\} \longrightarrow \gamma = ?$$

- MP $\Rightarrow$ KP

$T \longmapsto \gamma = \mu\otimes T_\#\mu = \mu\otimes\nu \quad \int_{X\times Y} c(x,y)\, d((I\otimes T)^{\#}\mu) = \int_X c(x,T(x))\, d\mu(x)$

Hence $\inf_{\gamma\in\Pi[\mu,\nu]} G_K(\gamma) \leq \inf_{T_\#\mu=\nu} G_M(T)$ (How about the converse?) $\Rightarrow$ by Duality KP

- proof of existence for KP

X, Y compact. c – lower semicontinuous

(Idea) $K_c: \Pi[\mu,\nu] \longrightarrow \mathbb{R} \quad \gamma \longmapsto \int_{X\times Y} c\, d\gamma$

① K_c lower semicontinuous

② $\Pi[\mu,\nu]$ compact

$\hookrightarrow$ weak topology

proof: Claim 1. if $f: X\to\mathbb{R}\cup\{+\infty\}$ is l.s.c. $\exists$ family $f_k \uparrow$ bdd continuous $f = \sup_k f_k$

WLOG $\exists x_0\in X$ $f(x_0) < +\infty$. we define $g_k(x) = \inf_{y\in X} f(y) + k\,d(x,y)$

$\leq f(x_0) + k\,d(x,x_0)$

g_k is k-Lip func. and $g_k(x) \leq g_l(x) \leq f(x)$ when $k\leq l$

Next we check $f(x) = \sup_k g_k(x)$

Given x $\forall k$, $\exists x_k\in X$ s.t. $f(x_k) + k\,d(x,x_k) \leq g_k(x) + \frac{1}{k} \leq f(x) + \frac{1}{k}$ ⊛

$d(x,x_k) \leq \frac{1}{k}\left(f(x) + \frac{1}{k} - f(x_k)\right) \leq \frac{1}{k}\left(f(x) + \frac{1}{k} - M\right) \to 0 \quad (k\to\infty) \quad x_k \to x$

f l.s.c. $f(x) \leq \liminf_{k\to\infty} f(x_k) \leq \sup_k g_k(x)$ (by ⊛)

$\Rightarrow$ finally let $f_k = \min(g_k, k)$ $\Rightarrow f_k$ continuous bdd by k $\quad f = \sup_k f_k$

<u>Claim 2.</u> K_c l.s.c. if c l.s.c.

by Claim 1. $\exists\, c_k \uparrow c$ bdd continuous

by Def of weak convergence, if $\gamma_m \to \gamma$ $\quad \int c_k \, d\gamma_m \to \int c_k \, d\gamma$

K_{c_k} is continuous $\quad K_c = \sup_k K_{c_k} \Rightarrow$ l.s.c.

<u>Claim 3.</u> $\Pi[\mu, \nu]$ compact

Take $\gamma_n \in \Pi[\mu, \nu]$ since μ, ν is probability measure

$\Rightarrow \gamma_n$ bounded in the dual of $C(X \times Y)$

$\exists$ weak convergence subsequence $\gamma_{n_k} \to \gamma$, only need to check $\gamma \in \Pi[\mu, \nu]$

since $\gamma_{n_k} \in \Pi[\mu, \nu]$ $\quad (\pi_x)^{\#} \gamma_{n_k} = \mu \Rightarrow \int \phi \circ \pi_x \, d\gamma_{n_k} = \int \phi \, d\mu \quad \forall \phi \in C(X)$

$$\downarrow$$

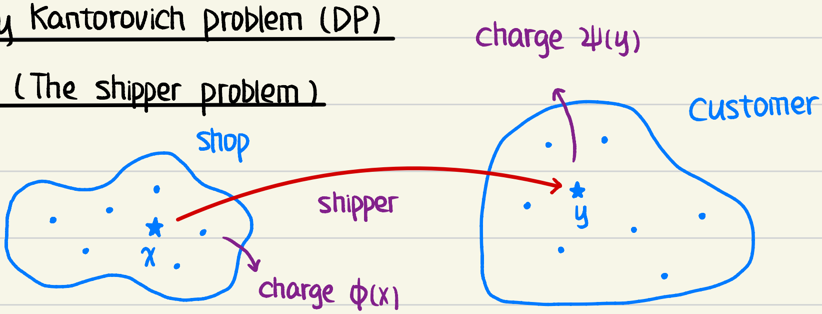
$$\int \phi \circ \pi_x \, d\gamma = \int \phi \, d(\pi_x)^{\#} \gamma$$

$\Rightarrow (\pi_x)^{\#} \gamma = \mu$ similarly $(\pi_y)^{\#} \gamma = \nu \Rightarrow \gamma \in \Pi[\mu, \nu]$

- Finally, l.s.c. on compact set must attains its minimum $\quad \blacksquare$

• <u>Duality Kantorovich problem (DP)</u>

<u>Example. (The shipper problem)</u>



• $c(x,y)$ is given $\phi(x)+\psi(y)\le c(x,y)$

total charge $\int_X \phi\, d\mu + \int_Y \psi\, d\nu = J(\phi,\psi)$

shopper try to minimize his cost, shipper try to maximize his charge

Problem. (Duality Problem)

$(X,\mu), (Y,\nu)$, $c: X\times Y \to [0,+\infty)$ continuous, solve the problem

$$\sup\{\underbrace{\int_X \phi\, d\mu + \int_Y \psi\, d\nu}_{= I(\phi,\psi)} \mid \phi\in C_b(X), \psi\in C_b(Y), \phi(x)+\psi(y)\le c(x,y)\}$$

To begin with, we want to give a sol. to DP (By Ascoli-Arzela) ↓ M

Def. (c- and $\bar{c}$-transform)

$\chi: X\to\bar{\mathbb{R}}$ c-transform $\Rightarrow \chi^c(y) = \inf_{x\in X}\{c(x,y)-\chi(x)\}$

$\xi: Y\to\bar{\mathbb{R}}$ $\bar{c}$-transform $\Rightarrow \xi^{\bar{c}}(x) = \inf_{y\in Y}\{c(x,y)-\xi(y)\}$

Rmk. x quantity, y selling price $c(x,y) \Rightarrow$ total price when buy x with price y

$\chi(x) \Rightarrow$ cost when buy x

$\Rightarrow \chi^c(y) = \inf_{x\in X}\{c(x,y)-\chi(x)\} \Rightarrow$ best amount of purchase when price is y

similarly $\xi^{\bar{c}}$

Def. (c-concave & $\bar{c}$-concave)

If f c-conjugate of some function $\Rightarrow f$ is c-concave

f $\bar{c}$-conjugate of some function $\Rightarrow f$ is $\bar{c}$-concave

Lemma. $\phi\in C_b(X), \psi\in C_b(Y)$ satisfying $\phi(x)+\psi(y)\le c(x,y)$, denote $I(\phi,\psi)=\int_X\phi\, d\mu+\int_Y\psi\, d\nu$, then $I(\phi,\psi)\le I(\phi,\phi^c)\le I(\phi^{c\bar{c}},\phi^c)$

proof. $\forall x \in X,\ y \in Y \quad \psi(y) \le c(x,y) - \phi(x) \Rightarrow \psi(y) \le \phi^c(y) \quad \phi(x) \le \psi^{\bar{c}}(x)$

$I(\phi,\psi) \le I(\phi,\phi^c) \le I(\phi^{c\bar{c}},\phi^c)$ $\blacksquare$

Thm (Solution of DP)

X, Y compact, c continuous, then $\exists$ c-concave ϕ such that (ϕ, ϕ^c) is optimal pair of Dual Problem, $\max(DP) = \max\limits_{\phi \in c\text{-conc}(X)} \int_X \phi \, d\mu + \int_Y \phi^c d\nu$

↳ all c-concave function

proof. suppose (ϕ_n, ψ_n) is maximizing sequence

$c(x,y)$ uniform continuous on $X \times Y \Rightarrow \{c(x,y) - \phi_n(x)\}, \{c(x,y) - \psi_n(y)\}$ equicontinuous

take inf $\{\phi_n^c\}, \{\psi_n^c\}$ equicontinuous $\Rightarrow \{\phi_n^{c\bar{c}}\}$ equicontinuous

By Ascoli-Arzela $\Rightarrow \exists$ subseq $\phi_{n_k} \to \phi,\ \psi_{n_k} \to \psi$

$I(\phi_n, \psi_n) \le I(\phi_n^{c\bar{c}}, \phi_n^c) \longrightarrow \max(DP) = I(\phi,\psi) = I(\phi^{c\bar{c}}, \phi^c)$

$\phi^{c\bar{c}} \in c\text{-conc}(X),\ \phi^c \in \bar{c}\text{-conc}(Y)$

min(KP) = max(DP) $\begin{cases} ① \text{ cyclic monotonicity} \\ ② \text{ Fenchel-Moreau} \Rightarrow \end{cases}$

Legendre transform $f^*(y) = \sup\limits_{x \in X} \{x \cdot y - f(x)\}$

Thm (Fenchel-Moreau) f lower semi-continuous, convex $\Rightarrow f = f^{**}$

Thm (Dual formula) X, Y compact, $c: X \times Y \to \mathbb{R}$ continuous, then $\min(KP) = \max(DP)$

proof: $H: C(X,Y) \longrightarrow \mathbb{R} \Rightarrow$

$H(p) = -\sup\{\int_X \phi \, d\mu + \int_Y \psi \, d\nu : \phi \in C_b(X), \psi \in C_b(Y), \phi(x) - \psi(y) \le (c-p)(x,y)\}$

↳ l.s.c.

By F-M $\Rightarrow H(0) = H^{**}(0)$ and $-H(0) = \max(DP)$

suffice to prove $-H^{**}(0) = \min(KP)$

$H^*(\gamma) = \sup_{p \in C(X\times Y)} \{ \int_{X\times Y} p \, d\gamma - H(p) \}$ $\quad \gamma \in \mathcal{M}(X\times Y)$ signed measure

↳ if γ not non-negative, take $np \le 0$ $n \nearrow \infty$ ruled out

Hence $\gamma \ge 0$ take $p = c - \phi - \psi$ to maximize $\int p \, d\gamma$

$\Rightarrow H^*(\gamma) = \sup_{\phi \in C_b(X), \psi \in C_b(Y)} \{ \int_{X\times Y} c(x,y) - \phi(x) - \psi(y) \, d\gamma + \int_X \phi \, d\mu + \int_Y \psi \, d\nu \}$

then $-H^{**}(0) = -\sup_\gamma \{ \langle 0, \gamma \rangle - H^*(\gamma) \} = \inf_\gamma H^*(\gamma) = \min(KP)$ ▨

Wasserstein metric

(X, d) metric space cost func. $c(x,y) = d(x,y)^p$

p-th moment finite space $\mathcal{P}_p(X) := \{ \mu \in \mathcal{P}(X) : \int_X |x|^p d\mu < \infty \}$

Def. (p-Wasserstein metric)

(X, d) metric space $\mu, \nu \in \mathcal{P}_p(X)$. $W_p(\mu, \nu) := \min \{ \int_{X\times X} d(x,y)^p d\gamma : \gamma \in \Pi[\mu, \nu] \}^{\frac{1}{p}}$

Rmk. ① W-metric $\Rightarrow P_p(X)$ metric

② $P_p(X)$ weak-* topology induced by W-metric

i.e. when X compact $\mu_n \rightharpoonup \mu \iff W_p(\mu_n, \mu) \to 0$