A Primer on Optimal Mass Transportation and the Supremal Monge Problem

Orlando Ferrari

Supervisor: Prof. Elio Marconi

Department of Mathematics "Tullio Levi-Civita"
University of Padova

Padova, 6 May 2026

Università degli Studi di Padova

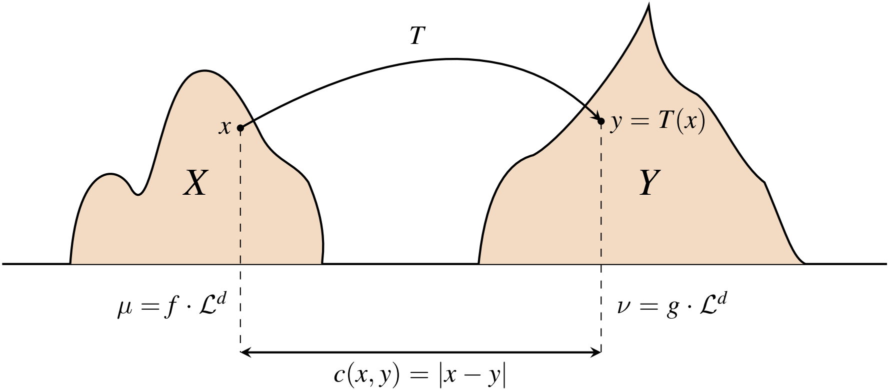
Seminar Notes

A full, rigorous write-up of these proofs and additional examples can be found in my seminar report.

refrainfr.github.io

The Monge Problem (1781)

How to transport a pile of dirt (*déblais*) to an excavation or embankment site (*remblais*) while minimizing the total physical effort, or transport cost?



The Monge Problem (1781)

Given metric spaces X, Y, measures $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$, and a cost $c : X \times Y \to [0, +\infty]$:

Definition (Transport Map)

*A measurable map $T : X \to Y$ is a **transport map** of mass μ to ν if for any Borel set $A \subset Y$:*

$$\int_A d\nu(y) = \int_{T^{-1}(A)} d\mu(x) \iff T_{\#}\mu(A) := \mu(T^{-1}(A)) = \nu(A).$$

Monge Problem (MP)

Find

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

a solution to (MP) is called an **optimal transport map**.

Issues with Monge Problem

Issue 1: Monge constraint is not closed
Define $f_n : [0,1] \to [0,1]$ by periodizing $x \mapsto nx - \lfloor nx \rfloor = \{nx\}$ in $[0,1]$. One can prove that $(f_n)_{\#}\mathcal{L} \llcorner [0,1] = \mathcal{L} \llcorner [0,1]$ for all $n \in \mathbb{N}$ using the Change of Variables formula and $f_n \rightharpoonup \frac{1}{2}$ where

$$\left(\frac{1}{2}\right)_{\#} \mathcal{L} \llcorner [0,1] = \delta_{\frac{1}{2}} \llcorner [0,1] \neq \mathcal{L} \llcorner [0,1].$$

Issue 2: Non-existence

- Let $\mu = \delta_{x_0}$ and ν be a continuous measure (no atoms).
- For any map T, $T_{\#}\delta_{x_0} = \delta_{T(x_0)}$.
- A map T **cannot split mass**, thus the admissible set is empty!

Issues with Monge Problem

Issue 1: Monge constraint is not closed

Define $f_n : [0,1] \to [0,1]$ by periodizing $x \mapsto nx - \lfloor nx \rfloor = \{nx\}$ in $[0,1]$. One can prove that $(f_n)_\# \mathcal{L} \llcorner [0,1] = \mathcal{L} \llcorner [0,1]$ for all $n \in \mathbb{N}$ using the Change of Variables formula and $f_n \rightharpoonup \frac{1}{2}$ where

$$\left(\frac{1}{2}\right)_\# \mathcal{L} \llcorner [0,1] = \delta_{\frac{1}{2}} \llcorner [0,1] \neq \mathcal{L} \llcorner [0,1].$$

Issue 2: Non-existence

- Let $\mu = \delta_{x_0}$ and ν be a continuous measure (no atoms).
- For any map T, $T_\# \delta_{x_0} = \delta_{T(x_0)}$.
- A map T **cannot split mass**, thus the admissible set is empty!

The Kantorovich Relaxation (1942)

To allow mass splitting, we relax the problem from deterministic maps T to probabilistic plans $\gamma \in \mathcal{P}(X \times Y)$.

Kantorovich Problem (KP)

Find

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

where $\Pi(\mu,\nu) = \{\gamma \in \mathcal{P}(X \times Y) : (\pi_X)_{\#}\gamma = \mu \text{ and } (\pi_Y)_{\#}\gamma = \nu\}$ given that π_X and π_Y are the **projections** of elements in $X \times Y$ onto X and Y, respectively.
We refer a measure $\gamma \in \Pi(\mu,\nu)$ as a **transport plan** and if γ is a solution (KP) we refer it as an **optimal transport plan**.

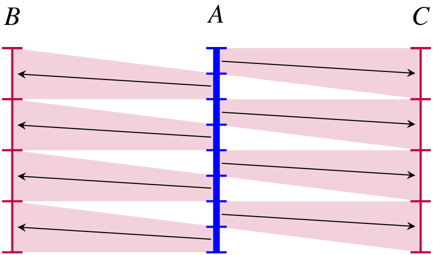
In this sense, we “allow" the transport map T to be multi-valued!

Another Issue of Monge Problem: The Ray Transport

Example (Solution of (KP) exists, but not (MP))

Let $c(x, y) = |x - y|^2$ *(or also* $c(x, y) = |x - y|$*) and consider* $\mu = \mathcal{H}^1 \llcorner A$ *and* $\nu = \frac{1}{2}\mathcal{H}^1 \llcorner B + \frac{1}{2}\mathcal{H}^1 \llcorner C$*, with*

$$A = \{[0, y] : y \in [0, 1]\}, \quad B = \{[-1, y] : y \in [0, 1]\}, \quad C = \{[1, y] : y \in [0, 1]\}.$$



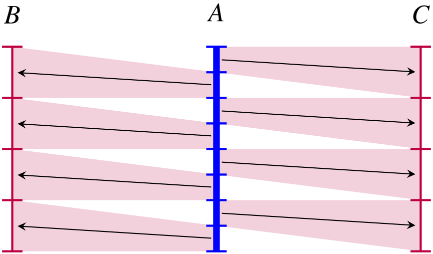
One can verify $\min(KP) = 1$ and attained by γ splitting the masses of A equally to B and C. However, we cannot have a solution for (MP) even though $\inf(MP) = 1$.

Another Issue of Monge Problem: The Ray Transport

Example (Solution of (KP) exists, but not (MP))

Let $c(x,y) = |x-y|^2$ *(or also* $c(x,y) = |x-y|$*) and consider* $\mu = \mathcal{H}^1 \llcorner A$ *and* $\nu = \frac{1}{2}\mathcal{H}^1 \llcorner B + \frac{1}{2}\mathcal{H}^1 \llcorner C$*, with*

$$A = \{[0,y] : y \in [0,1]\}, \quad B = \{[-1,y] : y \in [0,1]\}, \quad C = \{[1,y] : y \in [0,1]\}.$$



One can verify $\min(KP) = 1$ and attained by γ splitting the masses of A equally to B and C. However, we cannot have a solution for (MP) even though $\inf(MP) = 1$.

Direct Methods

Theorem (Weierstrass Critetion)

Let X be a ***sequentially compact*** *topological space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be* ***lower semicontinuous*** *(l.s.c.), then there exists $\bar{x} \in X$ such that $f(\bar{x}) = \min_{x \in X} f(x)$.*

Definition (Convergences on Space of (Signed) Measure $\mathcal{M}(X)$)

Let X be a metric space. We define:

- ***the weak topology*** *of $\mathcal{M}(X)$ with respect to the duality of $C_b(X)$ with the convergence denoted by $\mu_n \rightharpoonup \mu$.*
- ***the weak* topology*** *of $\mathcal{M}(X)$ with respect to the duality of $C_0(X)$ with the convergence denoted by $\mu_n \rightharpoonup^* \mu$ provided X is also locally compact.*

Since $C_0(X) \subset C_b(X)$ in general, the weak convergence is stronger. However, in the case of X being a compact space, both convergences coincide and instead we consider $\mathcal{M}(X)$ as the dual of $C(X)$.

Direct Methods

Theorem (Weierstrass Critetion)

*Let X be a **sequentially compact** topological space and $f : X \to \mathbb{R} \cup \{+\infty\}$ be **lower semicontinuous** (l.s.c.), then there exists $\bar{x} \in X$ such that $f(\bar{x}) = \min_{x \in X} f(x)$.*

Definition (Convergences on Space of (Signed) Measure $\mathcal{M}(X)$)

Let X be a metric space. We define:

- ***the weak topology** of $\mathcal{M}(X)$ with respect to the duality of $C_b(X)$ with the convergence denoted by $\mu_n \rightharpoonup \mu$.*
- ***the weak* topology** of $\mathcal{M}(X)$ with respect to the duality of $C_0(X)$ with the convergence denoted by $\mu_n \rightharpoonup^* \mu$ provided X is also locally compact.*

Since $C_0(X) \subset C_b(X)$ in general, the weak convergence is stronger. However, in the case of X being a compact space, both convergences coincide and instead we consider $\mathcal{M}(X)$ as the dual of $C(X)$.

Solutions of (KP)

Theorem (Existence of Solution, c Continuous)

Let X and Y be compact metric spaces, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, and $c : X \times Y \to \mathbb{R}$ be a continuous function. Then (KP) admits a solution.

Proof. K is continuous and $\Pi(\mu, \nu)$ is compact by Banach-Alaoglu followed by Riesz-Representation Theorem of $\mathcal{M}(X) = (C(X))'$.

Remarks.

Same result still holds on these cases

- If $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and bounded below, we can approximate c as an increasing limit of n-Lipschitz and bounded functions $\{c_n\}$ to achieve same result.
- Thanks to Prokhorov Theorem, if X and Y are instead Polish, i.e., separable & complete metric spaces, $\Pi(\mu, \nu)$ is still compact.

Solutions of (KP)

Theorem (Existence of Solution, c Continuous)

Let X and Y be compact metric spaces, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, and $c : X \times Y \to \mathbb{R}$ be a continuous function. Then (KP) admits a solution.

Proof. K is continuous and $\Pi(\mu, \nu)$ is compact by Banach-Alaoglu followed by Riesz-Representation Theorem of $\mathcal{M}(X) = (C(X))'$.

Remarks.

Same result still holds on these cases

- If $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and bounded below, we can approximate c as an increasing limit of n-Lipschitz and bounded functions $\{c_n\}$ to achieve same result.
- Thanks to Prokhorov Theorem, if X and Y are instead Polish, i.e., separable & complete metric spaces, $\Pi(\mu, \nu)$ is still compact.

Kantorovich Duality

Notice that the space of solutions for (KP) is convex. We can view (KP) as a linear optimization under convex constraints, we can form a dual problem.

The Dual Problem (DP)

Let X and Y be metric spaces. Given $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, and the cost function $c : X \times Y \to [0, +\infty)$ we consider the problem

$$\sup \left\{ \int_X \phi \, d\mu + \int_Y \psi \, d\nu : \varphi \in C_b(X), \psi \in C_b(Y) \text{ s.t. } \varphi \oplus \psi \le c \right\}. \quad \text{(DP)}$$

Key Tool for Strong Duality:

- **c-transform:** For $c : X \times Y \to [-\infty, +\infty]$, we define $\varphi^c(y) = \inf_{x \in X} \{c(x, y) - \varphi(x)\}$. It Allows us to restrict to c-concave functions, i.e. φ such that ψ^c for some ψ.

Kantorovich Duality

Notice that the space of solutions for (KP) is convex. We can view (KP) as a linear optimization under convex constraints, we can form a dual problem.

The Dual Problem (DP)

Let X and Y be metric spaces. Given $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, and the cost function $c : X \times Y \to [0, +\infty)$ we consider the problem

$$\sup \left\{ \int_X \phi \, d\mu + \int_Y \psi \, d\nu : \varphi \in C_b(X), \psi \in C_b(Y) \text{ s.t. } \varphi \oplus \psi \leq c \right\}. \quad \text{(DP)}$$

Key Tool for Strong Duality:

- ***c*-transform:** For $c : X \times Y \to [-\infty, +\infty]$, we define $\varphi^c(y) = \inf_{x \in X} \{c(x, y) - \varphi(x)\}$. It Allows us to restrict to c-concave functions, i.e. φ such that ψ^c for some ψ.

More on Duality

One suggestive interpretation of the dual problem as follows.

1. Suppose that X is a set of bakeries and Y is a set of cafes, then the problem in the (KP) corresponds to minimizing costs of a consortium between bakeries and cafes.
2. Now suppose that there is a transport company which buys at price $\phi(x)$ from X and sells at price $\psi(y)$ to Y. To be competitive with the direct agreement between bakeries and cafes, it must hold $\psi(y) - \phi(x) \leq c(x, y)$. Then the profit is $\int \psi(y)\, d\nu - \int \phi(x)\, d\mu$.
3. By changing sign, the problem for the transport company is the same as the formulation of (DP).

Notice that from the condition $\varphi \oplus \psi \leq c$ we immediately have $\sup(\mathrm{DP}) \leq \inf(\mathrm{KP})$. The converse inequality can be proved under mild assumptions.

More on Duality

One suggestive interpretation of the dual problem as follows.

1. Suppose that X is a set of bakeries and Y is a set of cafes, then the problem in the (KP) corresponds to minimizing costs of a consortium between bakeries and cafes.
2. Now suppose that there is a transport company which buys at price $\phi(x)$ from X and sells at price $\psi(y)$ to Y. To be competitive with the direct agreement between bakeries and cafes, it must hold $\psi(y) - \phi(x) \leq c(x, y)$. Then the profit is $\int \psi(y)\, d\nu - \int \phi(x)\, d\mu$.
3. By changing sign, the problem for the transport company is the same as the formulation of (DP).

Notice that from the condition $\varphi \oplus \psi \leq c$ we immediately have $\sup(\mathrm{DP}) \leq \inf(\mathrm{KP})$. The converse inequality can be proved under mild assumptions.

Definition

*c-Cyclically Monotone Set A set $\Gamma \subset X \times Y$ is **c-cyclically monotone (c-CM)** if for all or any $k \in \mathbb{N}$, any permutation $\sigma \in S_k$, and any finite elements $(x_1, p_1), \ldots, (x_k, p_k) \in \Gamma$, we have $\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_i, y_{\sigma(i)})$.*

Theorem (Rockafellar in c-CM)

If $\Gamma \neq \varnothing$ is c-CM, $c : X \times Y \to \mathbb{R}$. Then there exists a c-concave function $\varphi \not\equiv -\infty$ such that

$$\Gamma \subset \{(x, y) \in X \times Y : \varphi(x) + \varphi^c(y) = c(x, y)\}.$$

Proof. The desired function φ is

$$\varphi(x) = \inf \left\{ c(x, y_k) - c(x_k, y_k) + \sum_{i=0}^{k-1} c(x_{i+1}, y_i) - c(x_i, y_i) \right. \\ \left. : k \in \mathbb{N}, (x_i, y_i) \in \Gamma, \forall i = 1, \ldots, k \right\}.$$

We need $c(x, y) \neq \pm\infty$ to make sure case like $\infty - \infty$ does not happen.

Definition

*c-Cyclically Monotone Set A set $\Gamma \subset X \times Y$ is c-**cyclically monotone** (c-**CM**) if for all or any $k \in \mathbb{N}$, any permutation $\sigma \in S_k$, and any finite elements $(x_1, p_1), \ldots, (x_k, p_k) \in \Gamma$, we have $\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_i, y_{\sigma(i)})$.*

Theorem (Rockafellar in c-CM)

If $\Gamma \neq \varnothing$ is c-CM, $c : X \times Y \to \mathbb{R}$. Then there exists a c-concave function $\varphi \not\equiv -\infty$ such that

$$\Gamma \subset \{(x, y) \in X \times Y : \varphi(x) + \varphi^c(y) = c(x, y)\}.$$

Proof. The desired function φ is

$$\varphi(x) = \inf \Bigg\{ c(x, y_k) - c(x_k, y_k) + \sum_{i=0}^{k-1} c(x_{i+1}, y_i) - c(x_i, y_i) \\ : k \in \mathbb{N}, (x_i, y_i) \in \Gamma, \forall i = 1, \ldots, k \Bigg\}.$$

We need $c(x, y) \neq \pm\infty$ to make sure case like $\infty - \infty$ does not happen.

We assume X and Y are Polish.

Theorem

If γ is an optimal transport plan for the cost c in (KP) and c is continuous, then $\operatorname{supp}(\gamma)$ is a c-CM set. Moreover, if c is uniformly continuous and bounded, then (DP) admits a solution (φ, φ^c) and we have $\max(DP) = \min(KP)$.

Proof. Suppose $\operatorname{supp}(\gamma)$ is not c-CM, we can construct a strictly more optimal transport plan $\tilde{\gamma}$. utilizing the non-c-CM elements. Thus, contradicting optimality of γ.
For the later part, we want uniform continuity and boundedness to ensure φ (and thus φ^c) is continuous and bounded, hence admissible to (DP).

Remarks.

- If $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and bounded below, we can do approximation argument to have $\sup(DP) = \min(KP)$.
- As a consequence to approximation argument, in the case c same as above, γ is concentrated on a c-CM set Γ which will not be closed in general.

We assume X and Y are Polish.

Theorem

If γ is an optimal transport plan for the cost c in (KP) and c is continuous, then $\mathrm{supp}(\gamma)$ *is a c-CM set. Moreover, if c is uniformly continuous and bounded, then (DP) admits a solution* (φ, φ^c) *and we have* $\max(DP) = \min(KP)$.

Proof. Suppose $\mathrm{supp}(\gamma)$ is not c-CM, we can construct a strictly more optimal transport plan $\tilde{\gamma}$. utilizing the non-c-CM elements. Thus, contradicting optimality of γ.
For the later part, we want uniform continuity and boundedness to ensure φ (and thus φ^c) is continuous and bounded, hence admissible to (DP).

Remarks.

- If $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and bounded below, we can do approximation argument to have $\sup(DP) = \min(KP)$.
- As a consequence to approximation argument, in the case c same as above, γ is concentrated on a c-CM set Γ which will not be closed in general.

General $c(x,y) = h(x-y)$ where h is strictly convex

Proposition

If c is C^1, φ is a Kantorovich potential for the cost c, and γ is an optimal transport plan in the transport from μ to ν, and $(x_0, y_0) \in \operatorname{supp}(\gamma)$, then $\nabla\varphi(x_0) = \nabla_x c(x_0, y_0)$ provided that φ is differentiable at x_0.

Proof. For a fixed $(x_0, y_0) \in \operatorname{supp}(\gamma)$ we have $c(x_0, y_0) = \varphi(x_0) + \varphi^c(y_0)$ and thus the function $\mathbb{R}^d \ni x \mapsto c(x, y_0) - \varphi(x)$ achieve its minimum at $x = x_0$. Therefore, $\nabla_x c(x_0, y_0) = \nabla\varphi(x_0)$.

If h is strictly convex, it suggests that in differentiable points of φ at x_0, the equality from above Proposition (extended to subdifferential) gives us

$$\nabla\varphi(x_0) \in \partial h(x_0 - y_0) \implies y_0 = x_0 - (\partial h)^{-1}(\nabla\varphi(x_0))$$

where the relation $(\partial h)^{-1}$ is justified since h is strictly convex.

General $c(x, y) = h(x - y)$ where h is strictly convex

Proposition

If c is C^1, φ is a Kantorovich potential for the cost c, and γ is an optimal transport plan in the transport from μ to ν, and $(x_0, y_0) \in \operatorname{supp}(\gamma)$, then $\nabla\varphi(x_0) = \nabla_x c(x_0, y_0)$ provided that φ is differentiable at x_0.

Proof. For a fixed $(x_0, y_0) \in \operatorname{supp}(\gamma)$ we have $c(x_0, y_0) = \varphi(x_0) + \varphi^c(y_0)$ and thus the function $\mathbb{R}^d \ni x \mapsto c(x, y_0) - \varphi(x)$ achieve its minimum at $x = x_0$. Therefore, $\nabla_x c(x_0, y_0) = \nabla\varphi(x_0)$.

If h is strictly convex, it suggests that in differentiable points of φ at x_0, the equality from above Proposition (extended to subdifferential) gives us

$$\nabla\varphi(x_0) \in \partial h(x_0 - y_0) \implies y_0 = x_0 - (\partial h)^{-1}(\nabla\varphi(x_0))$$

where the relation $(\partial h)^{-1}$ is justified since h is strictly convex.

Existence of solution for (MP)

Our previous considerations is culminated formally in the following theorem

Theorem (Brenier for General h strictly convex)

Given μ and $\nu \in \mathcal{P}(\Omega)$ for a compact domain $\Omega \subset \mathbb{R}^d$. There exists an optimal transport plan γ for the cost $c(x,y) = h(x-y)$ with h is strictly convex on $\mathbb{R}^d$ and real valued. It is unique and of the form $(id, T)_{\#}\mu$, provided μ is absolutely continuous to the Lebesgue measure and $\partial\Omega$ is μ-negligible. Moreover, there exists a Kantorovich potential μ, and T and the potentials μ are linked by

$$T(x) = x - (\nabla h)^{-1}(\nabla \varphi(x)).$$

Remarks.

- $c(x,y) = |x-y|^p$ for $1 < p < +\infty$ admits a solution for (MP).
- An interesting case is $p = 2$ which was initiated by a result of Brenier that was done independently from OT approach. Moreover, on $p = 2$, the minimizer is of the form $T = \nabla u$ for a convex function u.

The Supremal Case (L^∞)

What if we want to minimize the *maximum* transport cost instead of the integral?

$$\min_{\gamma \in \Pi} \|x - y\|_{L^\infty(\gamma)}$$

The Challenge: The L^∞ problem is highly degenerate and not a type we considered beforehand!

Let us define

$$\begin{aligned} K_\infty(\gamma) := \|x - y\|_{L^\infty(\gamma)} &= \inf\{m \in \mathbb{R} : |x - y| \leq m \text{ for } \gamma\text{-a.e. } (x, y)\} \\ &= \max\{|x - y| : (x, y) \in \operatorname{supp}(\gamma)\}. \end{aligned}$$

Within this discussion, we limit our consideration on a compact domain $\Omega \subset \mathbb{R}^d$.

Secondary Variational Problem

Let $\Omega \subset \mathbb{R}^d$ be a compact domain.

Lemma (Solvability of Supremal Kantarovich)

For every $\gamma \in \mathcal{P}(\Omega \times \Omega)$, we have $K_p(\gamma)^{\frac{1}{p}} \uparrow K_\infty(\gamma)$.
In particular, $K_\infty(\gamma) = \sup_{p \geq 1} K_p(\gamma)^{\frac{1}{p}}$ and K_∞ is l.s.c. for the weak convergence in $\mathcal{P}(\Omega \times \Omega)$.
Thus, it admits a minimizer over $\Pi(\mu, \nu)$, which is compact by direct method.

The Secondary Variational Problem

We isolate the "best" L^∞ plan by minimizing a strictly convex secondary cost $K_2(\gamma)$ over the set of L^∞ minimizers $O_\infty(\mu, \nu)$:

$$\min\{K_2(\gamma) : \gamma \in O_\infty(\mu, \nu)\}$$

which admits a solution $\bar{\gamma}$.

Secondary Variational Problem

Let $\Omega \subset \mathbb{R}^d$ be a compact domain.

Lemma (Solvability of Supremal Kantarovich)

For every $\gamma \in \mathcal{P}(\Omega \times \Omega)$, we have $K_p(\gamma)^{\frac{1}{p}} \uparrow K_\infty(\gamma)$.
In particular, $K_\infty(\gamma) = \sup_{p \geq 1} K_p(\gamma)^{\frac{1}{p}}$ and K_∞ is l.s.c. for the weak convergence in $\mathcal{P}(\Omega \times \Omega)$.
Thus, it admits a minimizer over $\Pi(\mu, \nu)$, which is compact by direct method.

The Secondary Variational Problem

We isolate the "best" L^∞ plan by minimizing a strictly convex secondary cost $K_2(\gamma)$ over the set of L^∞ minimizers $O_\infty(\mu, \nu)$:

$$\min\{K_2(\gamma) : \gamma \in O_\infty(\mu, \nu)\}$$

which admits a solution $\bar{\gamma}$.

Minimizer of Secondary Variational

Let $L = \inf_{\gamma \in \Pi(\mu,\nu)} K_\infty(\gamma)$. Solving the secondary problem is equivalent to solve (KP) for the cost function

$$c(x,y) = \begin{cases} |x-y|^2 & \text{if } |x-y| \leq L \\ +\infty & \text{if } |x-y| \leq L \end{cases}$$

that is bounded below and l.s.c. Hence, $\bar{\gamma}$ is concentrated in a c-CM set Γ.

We can intersect Γ with $\text{supp}(\gamma)$ to have a c-CM set $\Gamma \subset \{(x,y) : |x-y| \leq L\}$. In particular, by c-CM, Γ has the property

$$(\forall (x_1,y_1),(x_2,y_2) \in \Gamma) |x_1 - y_2|, |x_2 - y_1| \leq L \implies \langle x_1 - x_2, y_1 - y_2 \rangle \geq 0. \tag{1}$$

This property suggests that Γ is necessarily concentrated on a map. It does! However, we need to work on this set up to removal of negligible points.

Minimizer of Secondary Variational

Let $L = \inf_{\gamma \in \Pi(\mu,\nu)} K_\infty(\gamma)$. Solving the secondary problem is equivalent to solve (KP) for the cost function

$$c(x,y) = \begin{cases} |x-y|^2 & \text{if } |x-y| \leq L \\ +\infty & \text{if } |x-y| \leq L \end{cases}$$

that is bounded below and l.s.c. Hence, $\bar{\gamma}$ is concentrated in a c-CM set Γ.

We can intersect Γ with $\mathrm{supp}(\gamma)$ to have a c-CM set $\Gamma \subset \{(x,y) : |x-y| \leq L\}$. In particular, by c-CM, Γ has the property

$$(\forall (x_1,y_1),(x_2,y_2) \in \Gamma)|x_1 - y_2|, |x_2 - y_1| \leq L \implies \langle x_1 - x_2, y_1 - y_2 \rangle \geq 0. \tag{1}$$

This property suggests that Γ is necessarily concentrated on a map. It does! However, we need to work on this set up to removal of negligible points.

Thank You!
Terima Kasih!
Grazie Mille!

Any questions?

Appendix: Refining $\bar{\Gamma}$

For practicality, $\mu \ll \mathcal{L}^d$. We shall define a "refinement" $\bar{\Gamma}$ as follows

1. Let $\{B_i\}$ be a countable topological basis of $\mathbb{R}^d$ and define $A_i := \pi_X(\Gamma \cap (\Omega \times B_i)) \subset \Omega$.
2. Set $N_i = A_i \backslash (\mathrm{Leb}(A_i))$ and we have $|N_i| = 0 \implies \mu(N_i) = 0$. Moreover, $\mu\left(\bigcup N_i\right) = 0$. Hence, we can now consider

$$\bar{\Gamma} = \Gamma \backslash \left(\left(\bigcup N_i\right) \times \Omega\right)$$

Indeed γ is still concentrated in $\bar{\Gamma}$ that is also c-CM.

Furthermore, $\bar{\Gamma}$ satisfies the property: for any pair $(x_0, y_0) \in \bar{\Gamma}$, any errors $\epsilon, \delta > 0$, any unit vector ξ (as a direction), and every sufficiently small $r > 0$, there are points $x \in (B(x_0, r) \backslash B(x_0, r/2)) \cap C(x_0, \xi, \delta)$, $y \in B(y_0, \epsilon)$ such that $(x, y) \in \bar{\Gamma}$, where $C(x_0, \xi, \delta)$ is the convex cone

$$C(x_0, \xi, \delta) := \{x : \langle x - x_0, \xi \rangle > (1 - \delta)|x - x_0|\} .$$

Appendix: Refining $\bar{\Gamma}$

For practicality, $\mu \ll \mathcal{L}^d$. We shall define a "refinement" $\bar{\Gamma}$ as follows

1. Let $\{B_i\}$ be a countable topological basis of $\mathbb{R}^d$ and define $A_i := \pi_X(\Gamma \cap (\Omega \times B_i)) \subset \Omega$.
2. Set $N_i = A_i \backslash (\text{Leb}(A_i))$ and we have $|N_i| = 0 \implies \mu(N_i) = 0$. Moreover, $\mu\left(\bigcup N_i\right) = 0$. Hence, we can now consider

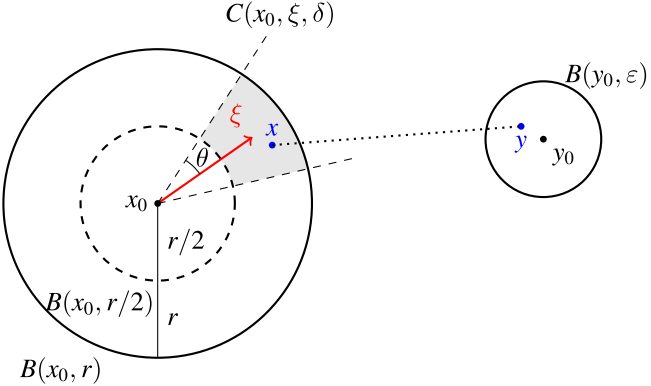
$$\bar{\Gamma} = \Gamma \backslash \left(\left(\bigcup N_i\right) \times \Omega\right)$$

Indeed γ is still concentrated in $\bar{\Gamma}$ that is also c-CM.

Furthermore, $\bar{\Gamma}$ satisfies the property: for any pair $(x_0, y_0) \in \bar{\Gamma}$, any errors $\epsilon, \delta > 0$, any unit vector ξ (as a direction), and every sufficiently small $r > 0$, there are points $x \in (B(x_0, r) \backslash B(x_0, r/2)) \cap C(x_0, \xi, \delta),\ y \in B(y_0, \epsilon)$ such that $(x, y) \in \bar{\Gamma}$, where $C(x_0, \xi, \delta)$ is the convex cone

$$C(x_0, \xi, \delta) := \{x : \langle x - x_0, \xi \rangle > (1 - \delta)|x - x_0|\}.$$

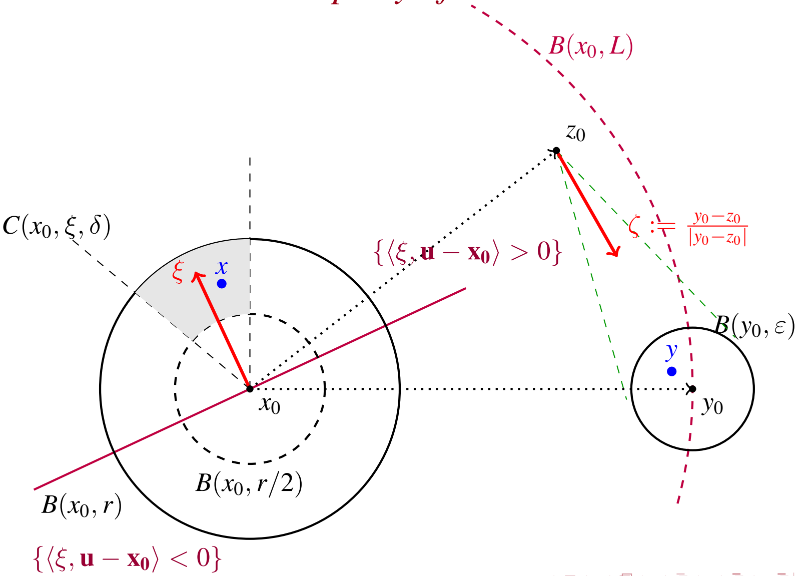
Appendix: Geometric Property of $\bar{\Gamma}$



Theorem

For any $(x_0, y_0), (x_0, z_0) \in \bar{\Gamma}$ we have $y_0 = z_0$. Thus $\bar{\Gamma}$ is graph of a (measurable) transport map T. Moreover, any solution to the Supremal Monge Problem is unique.

Appendix: Geometric Property of $\bar{\Gamma}$ in Final Theorem



Appendix: Sketch of Proof for Final Theorem

Suppose $y_0 \neq z_0$, the geometric constraints inside the cone would violate the c-CM property as follows:

1. We shall take a pair $(x, y) \in \bar{\Gamma}$ from an appropriate direction of unit vector ξ with respect to the point (x_0, y_0) such that $|x - z_0|$, $|x_0 - y| \leq L$. Then, c-CM of $\bar{\Gamma}$ implies $\langle x - x_0, y - z_0 \rangle \geq 0$.
2. However, the direction of $x - x_0$ is almost that of ξ (up to an error of $O(\sqrt{\delta})$) and the direction of $y - z_0$ is almost that of $\zeta := \frac{y_0 - z_0}{|y_0 - z_0|}$ (up to an error $O(\epsilon)$). If we choose ξ such that $\langle \xi, y_0 - z_0 \rangle < 0$, this means that for small enough δ and ϵ, we would get a contradiction
3. The computation for such (x, y) is done by considering the expansions

$$|x - z_0|^2 = |x - x_0|^2 + |x_0 - z_0|^2 + 2\langle x - x_0, x_0 - z_0 \rangle$$
$$|x_0 - y|^2 = |x_0 - x|^2 + |x - y|^2 + 2\langle x_0 - x, x - y \rangle$$

putting r, ϵ, δ small enough gives or goal for $|x - z_0|$, $|x_0 - y| \leq L$.

Therefore: The secondary optimal plan is induced by a (unique) map T.