Introduction to Optimal Transport

François-Xavier Vialard

LIGM, UPEM

OT Workshop, CDC@Nice, Dec. 2019

Monge Problem: 1784 [1]

How to move a pile of sand with the least effort possible ?

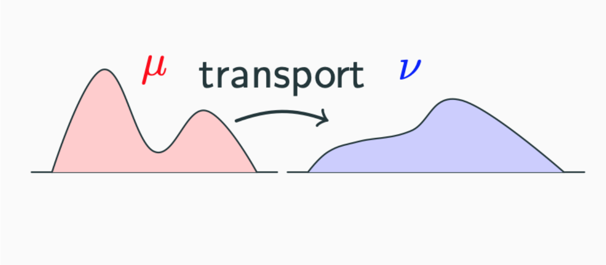


Figure – Pay a cost proportional to the distance per unit of mass

Discrete and continuous

1. Linear optimization problem.
2. Interplay between discrete and continuous settings.

1. Publication date for his "Mémoire sur la théorie des déblais et des remblais"

Contents

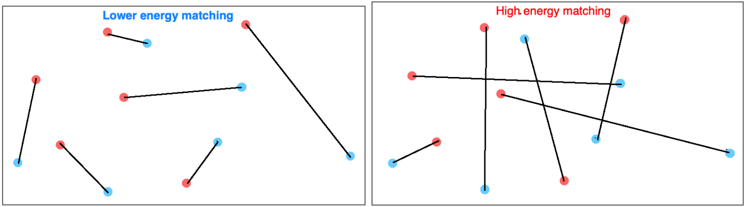
The matching problem

Data:

1. N red points x_i and N blue points y_j.
2. A cost: $c(i,j)$, e.g. $c(i,j) = d(x_i, y_j)$.

$$\textbf{Minimize} \sum_{i=1}^{N} c(i, \sigma(i))$$

over all possible bijections σ of $\{1, 2, \ldots, N\}$.

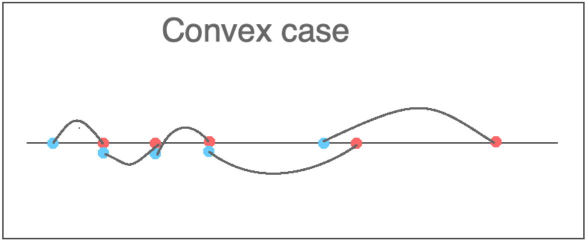


Naive brute force : $N!$ possibilities (e.g. $70! \approx 10^{100}$).

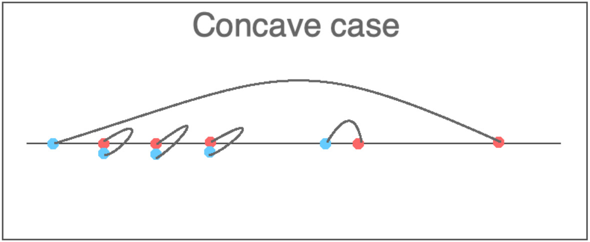
The 1D case:

Importance of the cost: $c(x, y) = f(y - x)$

1. f convex $\Longrightarrow$ monotone sorting, cost $N \log(N)$.



2. f concave $\Longrightarrow$ large displacements.

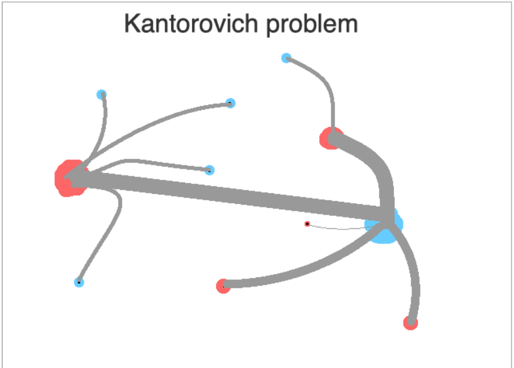


Matching with weights

1. N red sources x_i with μ_i capacity supply.
2. M blue clients y_j with ν_j demand.
3. Cost per unit $c(i,j) = c(x_i, y_j)$.

$$\textbf{Minimize} \sum_{i,j} c(i,j)\pi(i,j)$$

over $N \times M$ matrices π s.t. $\sum_i \pi(i,j) = \nu_j$ and $\sum_j \pi(i,j) = \mu_i$.



When Kantorovich = Monge

Setting:

1. When $m_i = n_j = 1/N$ for $i, j = 1, \ldots, N$.
2. Solve Kantorovich (instead of Monge): $\min \sum_{ij} \pi(i,j) c(i,j)$ under constraints.

Extremal points of bistochastic matrices $\implies \exists$ a permutation solution.

Using convexity

Notations:

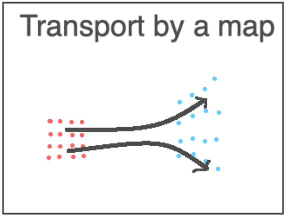
1. $\langle x, y\rangle = \sum_i x(i)y(i)$,
2. $\pi_1 = \sum_j \pi(i,j)$ and $\pi_2 = \sum_j \pi(i,j)$,
3. Let C be a convex set, $\iota_C(x) = \begin{cases} 0 \text{ if } x \in C \\ +\infty \text{ otherwise.} \end{cases}$.

$$\min\{\langle \pi, c\rangle \mid \pi \geq 0; \pi_1 = \mu; \pi_2 = \nu\} = \sup_{p_1, p_2} \{\langle p_1, \mu\rangle + \langle p_2, \nu\rangle \mid p_1 + p_2 \leq c\}.$$

Contents

From discrete to continuous

1. $T : \mathbb{R}^d \mapsto \mathbb{R}^d$.
2. Large number of points: $\mu = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$ represented by $\rho(x)\,\mathrm{d}x$.



$$T_*(\mu) = \frac{1}{N} \sum_{i=1}^{N} \delta_{T(x_i)} \quad (1)$$

$$T_*(\rho(x)) = \mathrm{Jac}(T^{-1}(x))\rho(T^{-1}(x))\,. \quad (2)$$

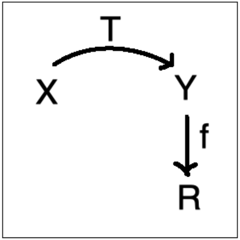
More formally:

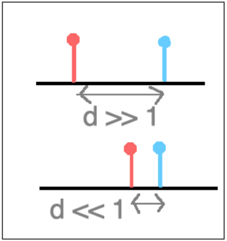
Definition

1. Let X, Y be (compact) spaces; $\mathcal{P}(X), \mathcal{P}(Y)$: probability measures on X, Y and $c : X \times Y \mapsto c(x, y)$ be a cost function.
2. A map $T : X \mapsto Y$ induces $T_* : \mathcal{P}(X) \mapsto \mathcal{P}(Y)$, i.e. $T_*(\mu)(B) = \mu(T^{-1}(B))$.
3. Also, $T^* : \mathcal{F}(Y, \mathbb{R}) \mapsto \mathcal{F}(X, \mathbb{R})$ i.e. $f \mapsto f \circ T$.

- Duality between measures and functions: $\langle T^* f, \mu\rangle = \langle f, T_* \mu\rangle$.
- Need of a topology to compare measures: weak* convergence:

$\mu_n \to \mu \Leftrightarrow \lim\langle f, \mu_n\rangle = \langle f, \mu\rangle$.





Monge formulation

Monge Problem

Fix $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$. Find a map $T : X \mapsto Y$ s.t.

$$\inf_T \int_X c(x, T(x)) \, \mathrm{d}\mu(x) \text{ s.t. } T_*\mu = \nu \,. \tag{3}$$

Problem: not feasible for all couples μ, ν.

Kantorovich formulation

Kantorovich Problem

Fix $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$. A coupling plan for μ, ν is:

$$\gamma \in \mathcal{P}(X \times Y) \text{ s.t. } [p_1]_* \gamma = \nu \text{ and } [p_2]_* \gamma = \mu \,. \tag{4}$$

Kantorovich problem is

$$MK_c(\mu, \nu) = \min_{\gamma \in \mathcal{C}(\mu,\nu)} \langle \gamma(x, y), c(x, y) \rangle \tag{5}$$

Proposition

1. *The set of coupling plans $\mathcal{C}(\mu, \nu)$ is compact for weak-* convergence of measures.*
2. *If $c(x, y)$ is continuous[a], then existence of a minimizing coupling plan.*

---

a. Actually, lower semicontinuous and bounded below is sufficient.

Duality

Fenchel-Rockafellar theorem implies:

Proposition

$$\min_{\gamma \in \mathcal{C}(\mu,\nu)} \langle \gamma(x,y), c(x,y) \rangle = \sup_{f,g \in C(X)\times C(Y)} \{ \langle f, \mu \rangle + \langle g, \nu \rangle \,|\, f(x) + g(y) \leq c(x,y) \}\,. \quad (6)$$

Existence of dual problem

1. Invariance: $(f,g) \mapsto (f+c, g-c)$, $\forall c \in \mathbb{R}$.
2. Given f, best possible $f^c(y) = \inf_x c(x,y) - f(x) \implies$ control on f^c.
3. Then, $(f,g) \mapsto (f, f^c) \mapsto (f^{cc}, f^c) \mapsto \ldots$ [a]
4. Existence by Arzela-Ascoli theorem.

a. In fact, $f^{ccc} = f^c$.

Economic interpretation of duality

Idea: externalize the cost of transport via buy/sell. Sending μ to ν is equivalent to

1. Sell $\mu(x)$ at $s(x)$ and buy $\nu(y)$ at $b(y)$.
2. You are willing to buy at y and sell at x, if this cost is less or equal to the "do it yourself" $c(x, y)$: $b(y) - s(x) \leq c(x, y)$.
3. Thus, the external company will try to maximize:

$$\sup_{s,b} \langle -s, \mu \rangle + \langle b, \nu \rangle \text{ s.t. } b(y) - s(x) \leq c(x, y)\,. \tag{7}$$

c-transforms

Definition (c-transform)

Let $f : X \mapsto \mathbb{R}$ and $c(x,y)$ continuous, define

$$f^c(y) = \inf_x c(x,y) - f(x)\,. \tag{8}$$

Example

$$f^c(y) = \inf_x \frac{1}{2}|x-y|^2 - f(x) = \frac{1}{2}|y|^2 + \inf_x -\langle x,y\rangle - (f - \frac{1}{2}|x|^2)\,.$$

$$\frac{1}{2}|y|^2 - f^c(y) = \sup_x \langle x,y\rangle - (\frac{1}{2}|x|^2 - f)\,.$$

$\frac{1}{2}|y|^2 - f^c(y)$ is **convex**.

Kantorovich boils down to Monge !

1. Consider $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ such that μ has density w.r.t. Lebesgue.
2. Consider $c(x, y) = \frac{1}{2}|x - y|^2$.
3. Consider optimal dual pair: (f, g) s.t. $f(x) + g(y) = c(x, y)$ then $\frac{1}{2}|x|^2 - f(x)$ is convex, $\implies$ [2] $\nabla f(x) = x - y$.

Brenier's map: optimal Monge map is the gradient of a convex function. $y = \nabla(\frac{1}{2}|x|^2 - f(x))$

2. Rademacher theorem

More rigourously

Definition

A set S in $X \times Y$ is said "c-cyclically monotone" if for all $n \geq 1$ and $(x_i, y_i)_{i=1,\dots,n}$

$$\sum_i c(x_i, y_i) \leq \sum_i c(x_i, y_{\sigma(i)}) \quad \forall \sigma \, \mathrm{Bij}(\{1, \dots, n\}) \, .$$

Proposition

1. *If* γ *is optimal then,* $\mathrm{supp}(\gamma)$ *is c-cyclically monotone.*
2. *If* Γ *is cyclically monotone, then there exists p c-convex,* $\Gamma \subset \{(x, y) \,|\, p(x) + p^c(y) = c(x, y)\}$.
3. *If* Γ *is c-cyclically monotone then it is optimal between its marginal.*

How rare is Monge?

How to get Monge ?

Recall optimality $p(x) + p^c(y) = c(x, y)$ and take differentiation if possible:

$$\nabla_x p(x) = \nabla_x c(x, y) \,. \tag{9}$$

It defines a unique y, if $y \mapsto \nabla_x c(x, y)$ is injective (Twist condition).

Example

1. $c(x, y) = d^2(x, y)$ on a compact Riemannian manifold.
2. $h(x - y)$ on $\mathbb{R}^d$ for h strictly convex.

Generically on costs: [3]

1. Fix μ, ν, uniqueness of optimal transport plan for generic costs c.
2. For a generic set of c, there exist μ, ν absolutely continuous, for which unique optimal plan not a map.

3. McCann - Rifford

Smoothness of Monge map?

1. Consider a cost, which satisfies twist condition.
2. Need μ, ν to be at least with density w.r.t. Lebesgue.

Monge-Ampère equation:

$$[\nabla p]_*(\mu) = \nu \Leftrightarrow \rho_\mu(x) = \det(\nabla^2 p)(x)\rho_\nu(\nabla p(x))\,. \tag{10}$$

Theorem (MTW - Ma-Trudinger-Wang)

Important quantity for smoothness:

$$MTW_{x,y}(\xi,\eta) = \sum_{ijklrs} c_{ij,r}c^{r,s}c_{s,kl} - c_{ij,kl})\xi^i\xi^j\eta^k\eta^l \geq 0 \tag{11}$$

+ other geometric properties.

Example

1. Euclidean spaces,
2. Spheres, but not ellipsoids.
3. C^4 perturbations of costs.

Wasserstein distances.

Definition

Let X, d be a complete metric space. The Wasserstein L^p distance is defined as

$$W_p(\mu, \nu) = MK_{d^2}(\mu, \nu)^{1/2}.$$

Proposition

It metrizes the weak convergence. If X is a length space, so is $(\mathcal{P}(X), W_2)$.

Contents

Fréchet means - Wasserstein barycenters

Define averages on a metric space:

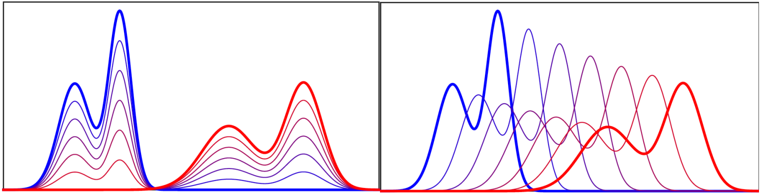
$$\arg\min_x \sum_{i=1}^N d(x, x_i)^p \tag{12}$$

On Wasserstein space, called Wasserstein barycenters:

$$\mathrm{Bar}((t_i, \rho_i)) = \arg\min_\rho \sum_{i=1}^N t_i W_2^2(\rho, \rho_i) \tag{13}$$

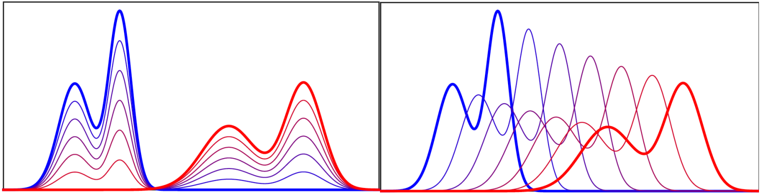
1. Convex problem.
2. Multimarginal optimal transport.

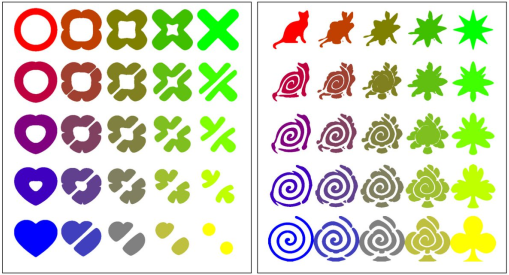
An example [4]



4. Gabriel Peyré's illustrations

An example [4]





4. Gabriel Peyré's illustrations

The 1D case

Proposition

If $c(x,y) = h(y-x)$ with strictly convex h, then, uniqueness of optimal transport plan and

$$\mathrm{MK}(\mu,\nu) = \int_0^1 h(\mathrm{icdf}(\mu)(s) - \mathrm{icdf}(\nu)(s))\,\mathrm{d}s \tag{14}$$

where $\mathrm{icdf}(\mu)$ *is the inverse*[a] *of* $\mu(]-\infty,x])$.

a. generalized!

Example

The L^2 Wasserstein distance on $\mathbb{R}$ is of "Hilbertian type".[a]
Sliced Wasserstein distances:

$$\mathrm{SW}(\mu,\nu)^p = \int_{S_{d-1}} W([p_\theta]_*\mu, [p_\theta]_*\nu)^p\,\mathrm{d}\theta\,. \tag{15}$$

a. Not true in higher dimension

Miscellanous

Proposition (Gaussian case)

Let A, B be two SDP matrices. Define corresponding Gaussians g_A, g_B, then

$$W_2(g_A, g_B)^2 = \operatorname{tr}(A) + \operatorname{tr}(B) - 2\operatorname{tr}(A^{1/2}BA^{1/2})^{1/2} \tag{16}$$

1. Affine map defined by SDP matrices are gradient of $\frac{1}{2}\langle x, Ax + \langle b, x\rangle$.
2. Known as the Bures metric (quantum mechanic) when restricted on unit trace matrices.

Miscellanous

Proposition (Gaussian case)

Let A, B be two SDP matrices. Define corresponding Gaussians g_A, g_B, then

$$W_2(g_A, g_B)^2 = \mathrm{tr}(A) + \mathrm{tr}(B) - 2\,\mathrm{tr}(A^{1/2}BA^{1/2})^{1/2} \tag{16}$$

1. Affine map defined by SDP matrices are gradient of $\frac{1}{2}\langle x, Ax + \langle b, x\rangle$.
2. Known as the Bures metric (quantum mechanic) when restricted on unit trace matrices.

Proposition (Probabilistic interpretation:)

If X, Y are of respective laws μ, ν,

$$MK_c(\mu, \nu) = \min_{X\sim\mu, Y\sim\nu} \mathbb{E}[c(X, Y)] \tag{17}$$

W_1 case

The c-transform for $c(x, y) = d(x, y)$ is $f \mapsto -f$

W_1 case

The c-transform for $c(x,y) = d(x,y)$ is $f \mapsto -f$

Proposition

$$W_1(\mu, \nu) = \sup_{f \in \mathrm{Lip}_1} \langle f, \mu - \nu \rangle . \tag{18}$$

W_1 case

The c-transform for $c(x,y) = d(x,y)$ is $f \mapsto -f$

Proposition

$$W_1(\mu,\nu) = \sup_{f \in \mathrm{Lip}_1} \langle f, \mu - \nu \rangle \,. \tag{18}$$

Definition (MMD distances, K reproducing kernel)

$$\|\mu - \nu\|_K \overset{\text{def.}}{=} \sup_{f \in B(1)} \langle f, \mu - \nu \rangle$$

$$\|\mu - \nu\|_K = \|K^{1/2}(\mu - \nu)\|_{L^2} \,.$$

Example ($g(x,y) = e^{-\|x_i - x_j\|^2/\sigma^2}$)

If $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$, then

$$\|\mu\|^2 = \sum_{i,j} g(x_i, x_j) \,.$$

Contents

The Riemannian(-like) metric

Consider $c(x,y) = d^2(x,y)$ and $\mu = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}$ and $\nu_\varepsilon = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i+\varepsilon v(x_i)}$, Then, $\varepsilon << 1$,

$$W_2(\mu,\nu)^2 = \frac{1}{N}\sum_{i=1}^{N}\varepsilon^2|v(x_i)|^2\,. \tag{19}$$

The Riemannian(-like) metric

Consider $c(x,y) = d^2(x,y)$ and $\mu = \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ and $\nu_\varepsilon = \frac{1}{N}\sum_{i=1}^N \delta_{x_i+\varepsilon v(x_i)}$. Then, $\varepsilon << 1$,

$$W_2(\mu,\nu)^2 = \frac{1}{N}\sum_{i=1}^N \varepsilon^2 |v(x_i)|^2 \,. \tag{19}$$

In general, for a density

$$g(\rho)(\delta\rho,\delta\rho) = \int_\Omega |\nabla(\Delta_\rho)^{-1}(\delta\rho)|^2 \rho(x)\,\mathrm{d}x \tag{20}$$

with $\Delta_\rho(f) = -\operatorname{div}(\rho\nabla f)$. Variational point of view:

$$g(\rho)(\delta\rho,\delta\rho) = \inf_v \left\{ \int_\Omega |v(x)|^2 \rho(x)\,\mathrm{d}x\,;\, -\operatorname{div}(\rho v) = \delta\rho \right\} \tag{21}$$

Benamou-Brenier formulation

For geodesic costs, for instance $c(x,y) = \frac{1}{2}|x-y|^2$

$$W_2(\mu,\nu) = \inf \mathcal{E}(v) = \frac{1}{2}\int_0^1 \int_M |v(x)|^2 \rho(x)\, \mathrm{d}x\, \mathrm{d}t \ , \tag{22}$$

s.t.

$$\begin{cases} \dot{\rho} + \nabla \cdot (v\rho) = 0 \\ \rho(0) = \mu \text{ and } \rho(1) = \nu \, . \end{cases} \tag{23}$$

Benamou-Brenier formulation

For geodesic costs, for instance $c(x,y) = \frac{1}{2}|x-y|^2$

$$W_2(\mu,\nu) = \inf \mathcal{E}(v) = \frac{1}{2}\int_0^1 \int_M |v(x)|^2 \rho(x)\, \mathrm{d}x\, \mathrm{d}t \ , \tag{22}$$

s.t.

$$\begin{cases} \dot{\rho} + \nabla \cdot (v\rho) = 0 \\ \rho(0) = \mu \text{ and } \rho(1) = \nu \,. \end{cases} \tag{23}$$

Convex reformulation: Change of variable: momentum $m = \rho v$,

$$\inf \mathcal{E}(m) = \frac{1}{2}\int_0^1 \int_M \frac{|m(x)|^2}{\rho(x)}\, \mathrm{d}x\, \mathrm{d}t \,, \tag{24}$$

s.t.

$$\begin{cases} \dot{\rho} + \nabla \cdot m = 0 \\ \rho(0) = \mu \text{ and } \rho(1) = \nu \,. \end{cases} \tag{25}$$

where $(\rho, m) \in \mathcal{M}([0,1] \times M, \mathbb{R} \times \mathbb{R}^d)$.

Benamou-Brenier formulation

For geodesic costs, for instance $c(x,y) = \frac{1}{2}|x-y|^2$

$$W_2(\mu,\nu) = \inf \mathcal{E}(v) = \frac{1}{2}\int_0^1 \int_M |v(x)|^2 \rho(x) \, \mathrm{d}x \, \mathrm{d}t \ , \tag{22}$$

s.t.

$$\begin{cases} \dot\rho + \nabla\cdot(v\rho) = 0 \\ \rho(0) = \mu \text{ and } \rho(1) = \nu \, . \end{cases} \tag{23}$$

Convex reformulation: Change of variable: momentum $m = \rho v$,

$$\inf \mathcal{E}(m) = \frac{1}{2}\int_0^1 \int_M \frac{|m(x)|^2}{\rho(x)} \, \mathrm{d}x \, \mathrm{d}t \, , \tag{24}$$

s.t.

$$\begin{cases} \dot\rho + \nabla\cdot m = 0 \\ \rho(0) = \mu \text{ and } \rho(1) = \nu \, . \end{cases} \tag{25}$$

where $(\rho, m) \in \mathcal{M}([0,1]\times M, \mathbb{R}\times\mathbb{R}^d)$.

Existence of minimizers: Fenchel-Rockafellar.
Numerics: First-order splitting algorithm: Douglas-Rachford.

Question: Extend Wasserstein to positive Radon measures.

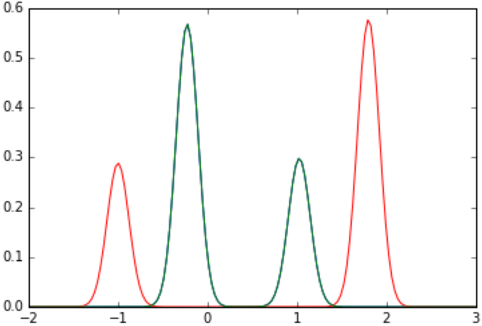


Figure – Optimal transport between bimodal densities

Question: Extend Wasserstein to positive Radon measures.

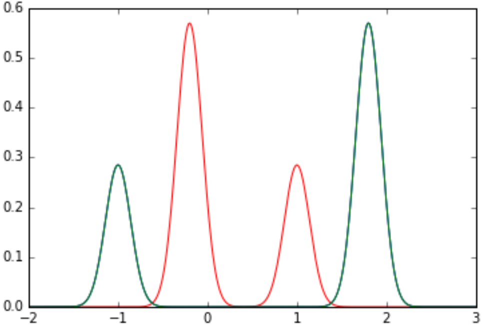


Figure – Another transformation

Two different approaches

- Extend static formulation:

$$\min \lambda KL(\mathrm{Proj}^1_* \gamma, \rho_1) + \lambda KL(\mathrm{Proj}^2_* \gamma, \rho_2) + \int_{M^2} \gamma(x,y) d(x,y)^2 \, \mathrm{d}x \, \mathrm{d}y \quad (26)$$

  Good for numerics, but is it a distance ?

- Extend dynamic formulation: on the tangent space of a density, choose a metric on the transverse direction.
  Built-in metric property but does there exist a static formulation ?

An extension of Benamou-Brenier formulation

Add a source term in the constraint: (weak sense)

$$\dot{\rho} = -\nabla \cdot (\rho v) + \alpha\rho \,,$$

where α can be understood as the growth rate.

$$\mathrm{WF}(m,\alpha)^2 = \frac{1}{2}\int_0^1 \int_M |v(x,t)|^2 \rho(x,t) \,\mathrm{d}x\,\mathrm{d}t + \frac{\delta^2}{2}\int_0^1 \int_M \alpha(x,t)^2 \rho(x,t)\,\mathrm{d}x\,\mathrm{d}t \,.$$

where δ is a length parameter.

A relaxed static OT formulation

Define

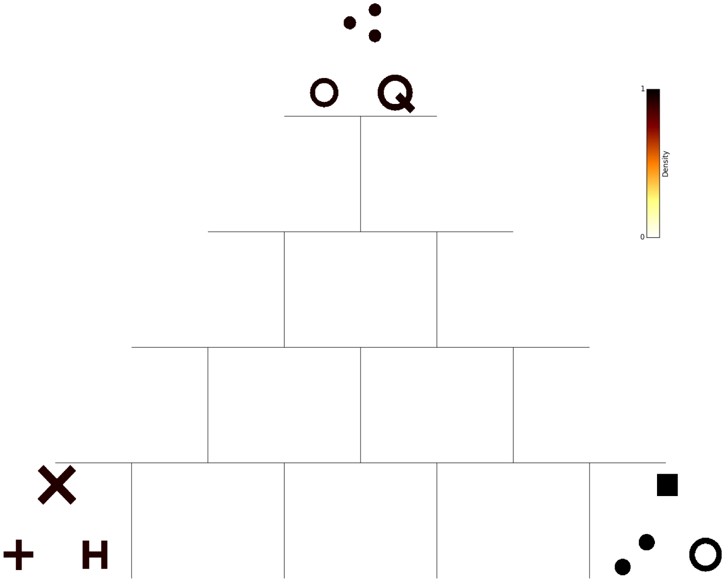
$$KL(\gamma,\nu) = \int \frac{\mathrm{d}\gamma}{\mathrm{d}\nu} \log\left(\frac{\mathrm{d}\gamma}{\mathrm{d}\nu}\right) \mathrm{d}\nu + |\nu| - |\gamma|$$

$$WF^2(\rho_1,\rho_2) = \inf_{\gamma} KL(\mathrm{Proj}^1_* \gamma, \rho_1) + KL(\mathrm{Proj}^2_* \gamma, \rho_2) - \int_{M^2} \gamma(x,y) \log(\cos^2(d(x,y)/2 \wedge \pi/2)) \,\mathrm{d}x\,\mathrm{d}y$$

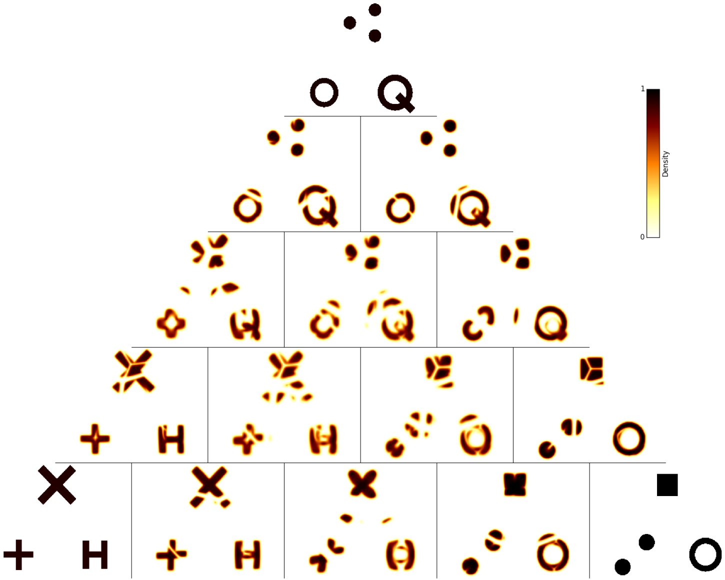
Theorem

On a Riemannian manifold (compact without boundary), the static and dynamic formulations are equal.

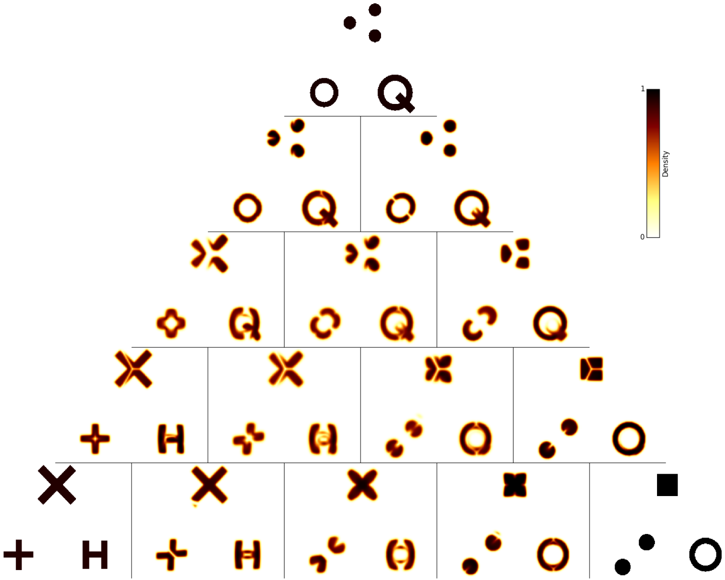
Barycenters : Wasserstein



Barycenters : Wasserstein



Barycenters : Unbalanced (GHK)



Contents

Gradient flows in metric spaces

Let $F : \mathbb{R}^d \mapsto \mathbb{R}$ be a function. The gradient flow of F w.r.t. the Euclidean metric is given by

$$\begin{cases} \dot{x} = -\nabla F(x) \\ x(0) = x_0 \in \mathbb{R}^d \, . \end{cases} \tag{27}$$

Proposition

If F is convex then $|x_1(t) - x_2(t)| \leq |x_1(0) - x_2(0)|$, therefore uniqueness of solutions.

Functionals on probability measures

1. A potential energy $\int_\Omega V(x)\,\mathrm{d}\mu(x)$.
2. An integral of a function of a density: $\int_\Omega f(\rho(x))\,\mathrm{d}x$.
3. Interaction energy: $\iint_\Omega W(x,y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$.
4. An optimal transport term: $\mu \mapsto \mathrm{MK}_c(\mu,\nu)$.

Geodesic convexity

On a length space X, the function V is geodesically convex if for every x_0, x_1 $\exists$ a geodesic $c(t)$ joining them s.t. $t \mapsto V(x(t))$ is convex.

1. $\int_\Omega V(x)\,\mathrm{d}\mu(x)$ displacement convex iff V is convex.
2. Integral functional on $\mathbb{R}^d$: f convex, superlinear, $s \mapsto s^{-d} f(s^d)$ is convex decreasing: entropy.

Heat flow

Proposition

$$\partial_t \rho = \Delta \rho \,, \tag{28}$$

is the gradient flow of the Dirichlet energy $\int_\Omega |\nabla \rho|^2 \,\mathrm{d}x$*.*

Proposition

The heat equation is a gradient flow of $E(\rho) = \int_\Omega \rho \log(\rho) \,\mathrm{d}x$ *for the Wasserstein metric.*

Interest ?

1. Define similar equation without gradient structure.
2. Define new numerical schemes: JKO scheme

More generally

- Let $F : \mathcal{P}(\Omega) \mapsto \mathbb{R}$ be a functional.

$$\dot{\rho} = \operatorname{div}\left(\rho \nabla \frac{\delta F}{\delta \rho}(\rho)\right) . \tag{29}$$

comes from

$$\arg\min_{v} \frac{1}{2} \int_M \|v(t,x)\|^2 \,\mathrm{d}\rho(x) + \left\langle \frac{\delta F}{\delta \rho}(\rho), -\operatorname{div}(\rho v) \right\rangle , \tag{30}$$

Example

$F(\rho) = \int_X \rho(x)(\log(\rho(x)) - 1)\,\mathrm{d}x + \int_X V(x)\rho(x)\,\mathrm{d}x$ for which $\frac{\delta F}{\delta \rho}(\rho) = \log(\rho) + V(x)$

$$\dot{\rho} = \Delta\rho + \operatorname{div}(\rho \nabla V) . \tag{31}$$

Gradient flows and growth models

Let $G : \mathcal{M}_+(X) \mapsto \mathbb{R}$ a functional.

JKO scheme

Implicit gradient descent: $x_{k+1} = x_k - \tau \nabla V(x_{k+1})$.

$$x_{k+1} = \arg\min V(x) + \frac{1}{2\tau}\|x - x_k\|^2$$

$$\mu_{k+1} = \arg\min F(\mu_k) + \frac{1}{2\tau} W(\mu, \mu_k)^2$$

Gradient flows and growth models

Let $G : \mathcal{M}_+(X) \mapsto \mathbb{R}$ a functional.

JKO scheme

Implicit gradient descent: $x_{k+1} = x_k - \tau \nabla V(x_{k+1})$.

$$x_{k+1} = \arg\min V(x) + \frac{1}{2\tau}\|x - x_k\|^2$$

$$\mu_{k+1} = \arg\min F(\mu_k) + \frac{1}{2\tau} W(\mu, \mu_k)^2$$

Gradient flow with *WF*

$$\partial_t \mu = \operatorname{div}(\mu \nabla \mathcal{G}'(\mu)) - 4\mu\mathcal{G}'(\mu) \tag{32}$$

JKO scheme with *WF*

$$\mu_{k+1}^{\tau} \in \operatorname{argmin}_{\mu \in \mathcal{M}_+(X)} \mathcal{G}(\mu) + \frac{1}{2\tau} \mathrm{WF}^2(\mu, \mu_k^{\tau}) \,. \tag{33}$$

Hele-Shaw type model

Preprint: *A tumor growth Hele-Shaw problem as a gradient flow*, S. Di Marino, L. Chizat.
Applied to $G(\rho) = -\rho(X) + \iota(\rho \leq 1)$, it leads to:

$$\begin{cases} \partial_t \rho - \nabla \cdot (\rho \nabla p) = 4(1-p)_+ \rho \\ p(1-\rho) = 0\,. \end{cases} \tag{34}$$

Hele-Shaw type model

Preprint: *A tumor growth Hele-Shaw problem as a gradient flow*, S. Di Marino, L. Chizat.
Applied to $G(\rho) = -\rho(X) + \iota(\rho \leq 1)$, it leads to:

$$\begin{cases} \partial_t \rho - \nabla \cdot (\rho \nabla p) = 4(1-p)_+ \rho \\ p(1-\rho) = 0\,. \end{cases} \tag{34}$$

Results

- Proof of existence and uniqueness of weak solutions (improving Perthame et al.)
- Implicit numerical scheme handling L^∞ constraints.
- Simple physical interpretation.

Simulations with JKO

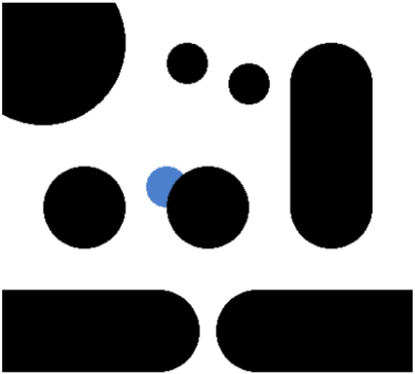


Figure – Flow of the density

Simulations with JKO

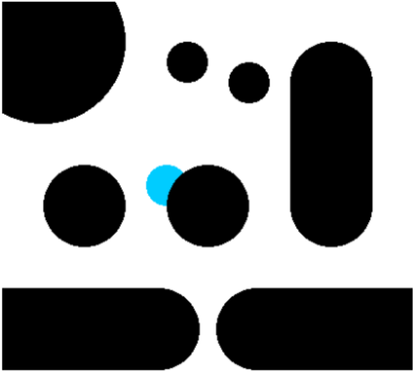


Figure – Level set of the pressure

Thank you for your attention !

- Optimal Transport for Applied Mathematicians, F. Santambrogio.
- Computational Optimal Transport, M. Cuturi, G. Peyré.
- Optimal Transport, Old and New, C. Villani.

References I