A glance on optimal transport and (some of) its (many) applications

Filippo Santambrogio

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud
http://www.math.u-psud.fr/~santambr/

July 6th, 2017, Manchester

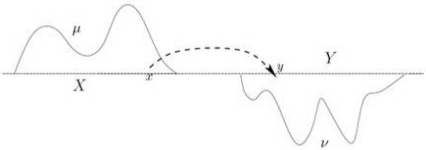
Outline

1. Monge and Kantorovich's theory
   - Duality
   - Stable marriages, prices and contracts
   - Brenier's theorem and convex functions
2. Wasserstein distances
   - Vertical vs horizontal distances
   - Geodesics and dynamic formulation
   - Wasserstein barycenters
   - Geodesic convexity
3. Optimization and equilibria in measure spaces
   - Continuous Nash equilibria
   - Examples in urban land use
   - Cournot-Nash equilibria
4. Variants
   - Congested transport (Wardrop equilibria)
   - Mean Field Games
   - Martingale optimal transport, option pricing and superhedging

Monge and Kantorovich's theory

Duality, existence, and economic interpretation

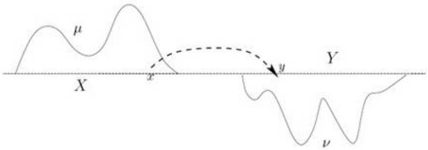
The Monge problem



If two distributions of mass $\mu, \nu \in \mathcal{P}(\Omega)$ are given on a compact domain of $\mathbb{R}^d$, the Monge problem reads:

$$\inf\left\{ \int |x - T(x)| d\mu(x) \ : \ T : \Omega \to \Omega,\ T_{\#}\mu = \nu \right\}.$$

The Monge problem



If two distributions of mass $\mu, \nu \in \mathcal{P}(\Omega)$ are given on a compact domain of $\mathbb{R}^d$, the Monge problem reads:

$$\inf\left\{ \int |x - T(x)| d\mu(x) \;:\; T : \Omega \to \Omega,\; T_{\#}\mu = \nu \right\}.$$

The problem can be generalized with $\mu \in \mathcal{P}(X), \nu \in \mathcal{P}(Y)$ and $c : X \times Y \to \mathbb{R}$, thus becoming

$$\inf\left\{ \int c(x, T(x)) d\mu(x) \;:\; T : X \to Y,\; T_{\#}\mu = \nu \right\}.$$

This problem, proposed in the 18th century, has stayed with no solution for more than 200 years...

G. Monge, Mémoire sur la théorie des déblais et des remblais, 1781

The Kantorovich problem

L. Kantorovich proposed to reformulate the same problem by describing the "transport" from X to Y via a different language:

$$\inf\Big\{ \int_{X\times Y} c(x,y)d\gamma(x,y) \ : \ \gamma\in\Pi(\mu,\nu)\Big\},$$

where $\Pi(\mu,\nu):=\{\gamma\in\mathcal{P}(X\times Y)\ :\ (\pi_x)_\#\gamma=\mu, (\pi_y)_\#\gamma=\nu\}$. We have now a convex, infinite-dimensional, linear programming problem.

The Kantorovich problem

L. Kantorovich proposed to reformulate the same problem by describing the "transport" from X to Y via a different language:

$$\inf\Big\{\int_{X\times Y} c(x,y)d\gamma(x,y)\ :\ \gamma\in\Pi(\mu,\nu)\Big\},$$

where $\Pi(\mu,\nu) := \{\gamma\in\mathcal{P}(X\times Y)\ :\ (\pi_x)_\#\gamma=\mu, (\pi_y)_\#\gamma=\nu\}$. We have now a convex, infinite-dimensional, linear programming problem. Existence is easy by weak compactness of probability measures. And there is also a dual formulation, obtained via inf-sup exchange

The Kantorovich problem

L. Kantorovich proposed to reformulate the same problem by describing the "transport" from X to Y via a different language:

$$\inf\Big\{\int_{X\times Y} c(x,y)d\gamma(x,y)\ :\ \gamma\in\Pi(\mu,\nu)\Big\},$$

where $\Pi(\mu,\nu) := \{\gamma\in\mathcal{P}(X\times Y)\ :\ (\pi_x)_\#\gamma=\mu, (\pi_y)_\#\gamma=\nu\}$. We have now a convex, infinite-dimensional, linear programming problem. Existence is easy by weak compactness of probability measures. And there is also a dual formulation, obtained via inf-sup exchange

$$\inf_{\gamma\geq 0}\sup_{\varphi,\psi}\int c d\gamma+\left(\int\varphi d\mu-\int\varphi(x)d\gamma\right)+\left(\int\psi d\nu-\int\psi(y)d\gamma\right)$$

L. Kantorovich, On the transfer of masses, 1942.

The Kantorovich problem

L. Kantorovich proposed to reformulate the same problem by describing the "transport" from X to Y via a different language:

$$\inf\Big\{\int_{X\times Y} c(x,y)d\gamma(x,y)\ :\ \gamma\in\Pi(\mu,\nu)\Big\},$$

where $\Pi(\mu,\nu) := \{\gamma\in\mathcal{P}(X\times Y)\ :\ (\pi_x)_\#\gamma=\mu, (\pi_y)_\#\gamma=\nu\}$. We have now a convex, infinite-dimensional, linear programming problem. Existence is easy by weak compactness of probability measures. And there is also a dual formulation, obtained via inf-sup exchange

$$\sup_{\varphi,\psi}\inf_{\gamma\geq 0}\int c d\gamma+\left(\int\varphi d\mu-\int\varphi(x)d\gamma\right)+\left(\int\psi d\nu-\int\psi(y)d\gamma\right)$$

L. Kantorovich, On the transfer of masses, 1942.

The Kantorovich problem

L. Kantorovich proposed to reformulate the same problem by describing the "transport" from X to Y via a different language:

$$\inf\Big\{\int_{X\times Y} c(x,y)d\gamma(x,y)\ :\ \gamma\in\Pi(\mu,\nu)\Big\},$$

where $\Pi(\mu,\nu) := \{\gamma \in \mathcal{P}(X\times Y)\ :\ (\pi_x)_\#\gamma = \mu, (\pi_y)_\#\gamma = \nu\}$. We have now a convex, infinite-dimensional, linear programming problem. Existence is easy by weak compactness of probability measures. And there is also a dual formulation, obtained via inf-sup exchange

$$\sup_{\varphi,\psi}\inf_{\gamma\geq 0}\int c d\gamma + \left(\int \varphi d\mu - \int\varphi(x)d\gamma\right) + \left(\int\psi d\nu - \int\psi(y)d\gamma\right)$$

L. Kantorovich, On the transfer of masses, 1942.

The Kantorovich problem

L. Kantorovich proposed to reformulate the same problem by describing the "transport" from X to Y via a different language:

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

where $\Pi(\mu,\nu) := \{\gamma\in\mathcal{P}(X\times Y) \;:\; (\pi_x)_\#\gamma=\mu, (\pi_y)_\#\gamma=\nu\}$. We have now a convex, infinite-dimensional, linear programming problem. Existence is easy by weak compactness of probability measures. And there is also a dual formulation, obtained via inf-sup exchange

$$\sup_{\varphi,\psi} \int \varphi d\mu + \int \psi d\nu \;:\; \varphi(x)+\psi(y)\le c(x,y)$$

L. Kantorovich, On the transfer of masses, 1942.

Existence of solutions

Theorem

If X and Y are compact and c continuous, then

- *there exists a solution γ to the primal problem* $\min\left\{\int c\, d\gamma \;:\; \gamma \in \Pi(\mu,\nu)\right\}$ **(KP)**,
- *there exist a solution $(\varphi,\psi) \in C(X) \times C(Y)$ to the dual problem* $\max\left\{\int \varphi d\mu + \int \psi d\nu \;:\; \varphi(x)+\psi(y) \le c(x,y)\right\}$ **(DP)**,
- min**(KP)**$=$ max**(DP)**,
- *given γ and (φ,ψ), they are optimal in the primal and dual problems, respectively, if and only if we have $\varphi(x)+\psi(y) = c(x,y)$ on $\mathrm{supp}(\gamma)$.*

If μ has no atoms, the infimum in the Monge problem equals min**(KP)**

$$\inf\left\{\int c(x,T(x))d\mu(x) \;:\; T_{\#}\mu = \nu\right\} = \min\left\{\int c\, d\gamma \;:\; \gamma \in \Pi(\mu,\nu)\right\}$$

and if the optimal γ is of the form $\gamma = (id, T)_{\#}\mu$ (i.e. it is concentrated on the graph of a map $T : X \to Y$), then T solves the Monge problem.

Few references - monographs

C. Villani, *Topics on Optimal Transportation*, 2003

Few references - monographs

C. VILLANI, *Topics on Optimal Transportation*, 2003

C. VILLANI, *Optimal Transport, Old and New*, 2008

Few references - monographs

C. Villani, *Topics on Optimal Transportation*, 2003

C. Villani, *Optimal Transport, Old and New*, 2008

F. Santambrogio, *Optimal Transport for Applied Mathematicians*, 2015

Few references - monographs

C. VILLANI, *Topics on Optimal Transportation*, 2003

C. VILLANI, *Optimal Transport, Old and New*, 2008

F. SANTAMBROGIO, *Optimal Transport for Applied Mathematicians*, 2015

A. GALICHON, *Optimal Transport Methods in Economics*, 2016

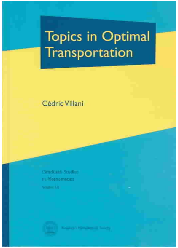
Few references - monographs

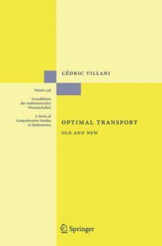
C. VILLANI, *Topics on Optimal Transportation*, 2003

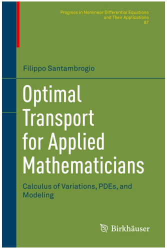
C. VILLANI, *Optimal Transport, Old and New*, 2008

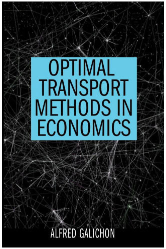
F. SANTAMBROGIO, *Optimal Transport for Applied Mathematicians*, 2015

A. GALICHON, *Optimal Transport Methods in Economics*, 2016









Stable Marriages

X = types of women, Y = types of men, μ and ν their distributions.
$u_w(x, y)$ = the interest of Ms x for Mr y, $u_m(x, y)$ = that of Mr y for Ms x.
Problem: **finding a stable set of marriages**, i.e. a measure $\gamma \in \Pi(\mu, \nu)$ (who marries whom), such that no new couple (x, y) will decide to divorce (each one from his/her current partner) to go together.
Case of transferable utility: once x and y get married, they decide how to split their total utility $u_w(x, y) + u_m(x, y)$, into a quantity $\varphi(x)$ (utility surplus for Ms x - now Mrs y), and $\psi(y)$ for Mr y. Only the sum $U(x, y) := u_w(x, y) + u_m(x, y)$ really plays a role.
A stable marriage is a triple (γ, φ, ψ) such that

- $U(x, y) = \varphi(x) + \psi(y)$ γ-a.e.,
- $U(x, y) \leq \varphi(x) + \psi(y)$ for all (x, y),
- $\gamma \in \Pi(\mu, \nu)$.

Just solve **(KP)** and **(DP)** for $c = -U$ and change the sign to φ, ψ.

D. Gale, L. S. Shapley, College Admissions and the Stability of Marriage, *Amer. Math. Month.* 1962
P.-A. Chiappori, R. J. McCann, L. P. Nesheim Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness, *Economic Theory*, 2010.

Prices

$X =$ types of goods on the market, $Y =$ consumers, μ, ν their distributions.
$u(x, y) =$ the utility of the consumer y when he buys the good x.
The goal is to **determine the prices of the goods and who buys what**.
Suppose that the price $p(x)$ of each good is known; then, each y will choose what to buy by solving $\max_x u(x, y) - p(x)$. Let us call $p^{(u)}(y)$ the value of the max. We look for (γ, p) such that

- $\gamma \in \Pi(\mu, \nu)$.
- $p^{(u)}(y) = u(x, y) - p(x)$ for $(x, y) \in \text{supp}(\gamma)$.

Just solve **(KP)** and **(DP)** for $c = -u$ and use $p = -\varphi$ and $p^{(u)} = \psi$.

Prices

X = types of goods on the market, Y = consumers, μ, ν their distributions. $u(x, y)$ = the utility of the consumer y when he buys the good x.
The goal is to **determine the prices of the goods and who buys what**.
Suppose that the price $p(x)$ of each good is known; then, each y will choose what to buy by solving $\max_x u(x, y) - p(x)$. Let us call $p^{(u)}(y)$ the value of the max. We look for (γ, p) such that

- $\gamma \in \Pi(\mu, \nu)$.
- $p^{(u)}(y) = u(x, y) - p(x)$ for $(x, y) \in \mathrm{supp}(\gamma)$.

Just solve **(KP)** and **(DP)** for $c = -u$ and use $p = -\varphi$ and $p^{(u)} = \psi$.

Wait, wait... prices are only obtained up to an additive constant??
Consider a special good x_0 corresponding to "not buying anything at all", and impose $p(x_0) = 0$ (if $\mu(\{x_0\}) > 0$ then there are not enough goods for everybody and some consumers will stay out of the market).

Prices

$X =$ types of goods on the market, $Y =$ consumers, μ, ν their distributions. $u(x, y) =$ the utility of the consumer y when he buys the good x.
The goal is to **determine the prices of the goods and who buys what.**
Suppose that the price $p(x)$ of each good is known; then, each y will choose what to buy by solving $\max_x u(x, y) - p(x)$. Let us call $p^{(u)}(y)$ the value of the max. We look for (γ, p) such that

- $\gamma \in \Pi(\mu, \nu)$.
- $p^{(u)}(y) = u(x, y) - p(x)$ for $(x, y) \in \operatorname{supp}(\gamma)$.

Just solve **(KP)** and **(DP)** for $c = -u$ and use $p = -\varphi$ and $p^{(u)} = \psi$.

Wait, wait... prices are only obtained up to an additive constant??
Consider a special good x_0 corresponding to "not buying anything at all", and impose $p(x_0) = 0$ (if $\mu(\{x_0\}) > 0$ then there are not enough goods for everybody and some consumers will stay out of the market).

The happy ending of free market economy: the stable (**equilibrium**) solution is also the one which maximizes the social utility (**optimal**).

Monopolist pricing, principal-agent, contract theory...

A company has the monopoly on a given sector and can decice prices p and production μ. Now X =of feasible goods. We also have a production cost $k : X \to \mathbb{R}$. Y, ν, u and x_0 are as before (with $k(x_0) = 0$).

Goal: find the optimal pricelist p. Every y selects its optimal good $X_p(y) \in \mathrm{argmax}_x\, u(x, y) - p(x)$, and that the total income of the company is

$$I(p) := \int (p - k)(X_p(y))d\nu(y) = \int (p - k)d\mu,$$

where $\mu = X_{\#}\nu$ (measure of the real production of goods). The price p will be chosen so as to maximize $I(p)$, with $p(x_0) = 0$.

Of course the formulation should be revisited in case the optimal point $X_p(y)$ is not unique...

M. Armstrong, Multiproduct nonlinear pricing, *Econometrica,* 1996
M. Ghisi and M. Gobbino The monopolist's problem: existence, relaxation and approximation, *Calc. Var. PDE*, 2005

The Rochet-Choné formulation of the principal-agent problem

Consider $u(x,y) = x \cdot y$, with $x, y \in \mathbb{R}^d$ (which means: goods x have d characteristics, consumers are classified by their interest for each of them). Look at $\max_x x \cdot y - p(x)$: we are touching p from below by affine functions and p can be replaced by its convex envelop $\overline{p}$ (if $p(x) > \overline{p}(x)$ then no y will ever buy x, so it's better to decrease its price to $\overline{p}(x)$). The maximizer x is characterized by $y = \nabla\overline{p}(x)$, i.e. $X_p(y) = (\nabla\overline{p})^{-1}(y) = \nabla p^*(y)$. Using $\overline{p}(\nabla p^*(y)) = y \cdot \nabla p^*(y) - p^*(y)$, the income maximization problem becomes a problem on p^*. The condition $p(0) = 0$ will bring $p^* \geq 0$:

$$\min\left\{\int (k(\nabla p^*(y)) + p^*(y) - y \cdot \nabla p^*(y))\, d\nu(y) \;:\; p^* \text{ convex},\, p^* \geq 0\right\},$$

which is a calculus of variations problem with (non-standard) convexity constraints.

J.-C. Rochet, P. Choné Ironing, Sweeping, and Multidimensional Screening. *Econometrica*, 1998.

Optimal transport maps, Brenier's theorem, and gradients of convex functions

Let's consider $X = Y \subset \mathbb{R}^d$ and c smooth. Take γ, φ, ψ optimal and $(x_0, y_0) \in \mathrm{supp}(\gamma)$. We get that $x \mapsto c(x, y_0) - \varphi(x)$ is maximal at $x = x_0$, hence $\nabla_x c(x_0, y_0) = \nabla\varphi(x_0)$. If c satisfies the **twist condition** ($\nabla_x c$ is injective in y for every x_0), then $y_0 = (\nabla_x c(x_0, \cdot))^{-1}(\nabla\varphi(x_0)) := T(x_0)$ is uniquely defined, and γ is unique and concentrated on the graph of T (of course, differentiability of φ must be guaranteed).

Theorem

Suppose $c(x, y) = -x \cdot y$ and μ absolutely continuous. Then, given ν, the optimal γ in **(KP)** *is unique and concentrated on the graph of $\nabla\varphi$, where φ is a convex function and solves* **(DP)**.
A map T is optimal for the Monge problem if and only if it is the gradient of a convex function (which is differentiable a.e., hence μ-a.e.).

Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, *CRAS*, 1987.

Wasserstein spaces

Distances, curves, geodesics, and barycenters

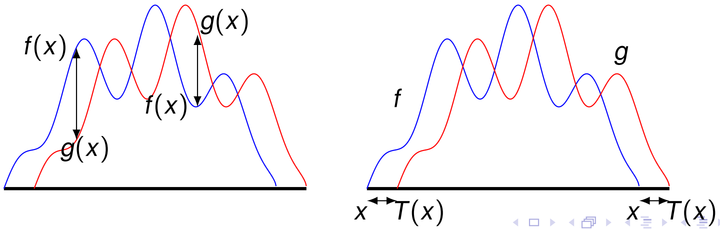
Wasserstein distances and Wasserstein spaces - 1

Starting from the values **(KP)** we can define a set of distances over $\mathcal{P}(X)$, in the following way: for any $p \in [1, +\infty[$ set

$$W_p(\mu,\nu) = \Big(\min \ \textbf{(KP)} \text{ with } c(x,y) = |x-y|^p\Big)^{1/p}.$$

Compared to L^p distances between densities we can say that they are "horizontal" instead of "vertical".

Topology and functional analysis: if X is compact, then the convergence for any W_p is equivalent to the weak convergence in the dual of $C(X)$, the space of continuous functions on X.



Wasserstein distances and Wasserstein spaces - 2

There is also a dynamical formulation whenever $X \subset \mathbb{R}^d$ is convex:

$$\begin{aligned} W_p^p(\mu,\nu) &= \inf\left\{ \int_0^1 \int \rho_t |v_t|^p \, dxdt \ : \ \partial_t\rho + \nabla\cdot(\rho v) = 0, \ \rho_0 = \mu, \rho_1 = \nu \right\} \\ &= \inf\left\{ \int_0^1 \int \frac{|w_t|^p}{\rho_t^{p-1}} \, dxdt \ : \ \partial_t\rho + \nabla\cdot w = 0, \quad \rho_0 = \mu, \rho_1 = \nu \right\} \end{aligned}$$

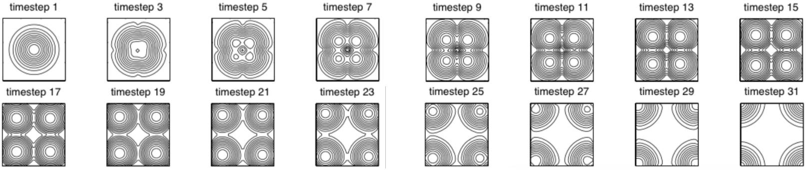
This kinetic energy minimization is the so-called *Benamou-Brenier* formulation, which amounts to a convex optimization problem, solvable by Augmented Lagrangian methods.

In the case $p = 1$ the last problem gives an alternative formulation, which is dynamic but stationary, and is known as the Beckmann's **minimal flow** formulation of the Monge problem:

$$W_1(\mu,\nu) = \min\left\{ \int |w| \ : \ \nabla\cdot w = \mu - \nu \right\}.$$

J.-D. Benamou, Y. Brenier A computational fluid mechanics solution to the Monge- Kantorovich mass transfer problem, *Numer. Math.*, 2000.
M. Beckmann, A continuous model of transportation, *Econometrica*, 1952.

Wasserstein distances and Wasserstein spaces - 3

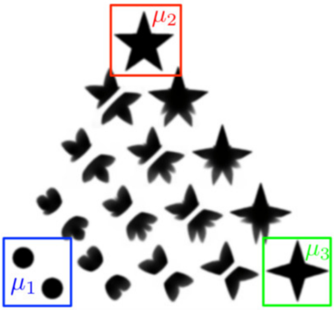


The measures ρ_t which minimize the Benamou-Brenier formulation are indeed constant-speed geodesics connecting μ to ν and they have an explicit expression

$$\rho_t = ((1-t)id + tT)_{\#}\mu,$$

where T is the optimal transport map from μ to ν. Such a map exists and is unique, since $|x-y|^p$ satisfies the twist condition (for $p > 1$).

Wasserstein barycenters



Picture and numerics by G. Peyré

Given some distributions of mass μ_i on a metric space X, how can we "average" them, to get a typical distribution out of them? We can apply different weights λ_i with $\sum_i \lambda_i = 1$, but want to preserve qualitative properties (concentration,...) which would disappear if we took $\sum_i \lambda_i \mu_i$. In a Hilbert space, a barycenter $\sum_i \lambda_i x_i$ is the solution of $\min_y \sum_i \lambda_i ||y - x_i||^2$.
Here solve $\min_\rho \sum_i \lambda_i W_2^2(\rho, \mu_i)$.

M. Agueh, G. Carlier, Barycenters in the Wasserstein space, *SIAM J. Math. An.*, 2011.
J. Rabin, G. Peyré, J. Delon, M. Bernot. Wasserstein Barycenter and Its Application to Texture Mixing, *Scale Space and Variational Methods in Computer Vision*, 2012.

Geodesic convexity

What does it mean that a function $\mathcal{F} : X \to \mathbb{R}$ is convex, when X is a metric space? we can say that $\mathcal{F}$ is **geodesically convex** if $t \mapsto \mathcal{F}(x(t))$ is convex for every constant-speed geodesics connecting arbitrary points $x(0), x(1)$. In the case $X = W_2(\mathbb{R}^d)$ this was introduced by McCann with the name of **displacement convexity** and the geodesics ρ_t are known. Three main functionals are considered

$$\mathcal{H}(\rho) := \int H(\rho(x))dx, \; \mathcal{V}(\rho) := \int V d\rho, \; \mathcal{W}(\rho) := \int W(x-y)d\rho(x)d\rho(y).$$

Theorem

$\mathcal{V}$ is geodesically convex if and only if V is convex, $\mathcal{W}$ if W is convex, and $\mathcal{H}$ if $s \mapsto s^d H(s^{-d})$ is convex nondecreasing (this is satisfied by $H(s) = s^p$, $p > 1$, and $H(s) = s \log s$).

R. J. McCann A convexity principle for interacting gases. *Adv. Math.*, 1997.

Equilibria and optimization with measures

Non-atomic games and urban equilibria

Nash equilbria with a continuum of players

Consider a game with infinitely many identical players, each one negligible compared to the others (*non-atomic games*), and a common space S of strategies; players choose their strategies and we look for the realized measure, $\rho \in \mathcal{P}(S)$, which induces a payoff function $f_\rho : S \to \mathbb{R}$ and we want: $f_\rho(x) = \min f_\rho$ for every $x \in \mathrm{supp}(\rho)$.

Definition/notation: given a functional $\mathcal{F} : \mathcal{P}(S) \to \mathbb{R}$ we define its first variation as the function $\frac{\delta \mathcal{F}}{\delta \rho}$, if it exists, such that

$$\frac{d}{d\varepsilon}\mathcal{F}(\rho + \varepsilon\chi)_{|\varepsilon=0} = \int \frac{\delta \mathcal{F}}{\delta \rho}(\rho) d\chi.$$

In the particular case where $f_\rho = \frac{\delta \mathcal{F}}{\delta \rho}$, then the above game is a **potential game**, and equilibria can be found by minimizing $\mathcal{F}(\rho)$ among $\rho \in \mathcal{P}(S)$. If $\mathcal{F}$ is convex, then necessary optimality conditions are also sufficient, and every equilibrium is a minimizer.

J. NASH, Equilibrium points in n-person games, *Proc. Nati. Acad. Sci.*, 1950.
J. NASH, Non-Cooperative Games *Ann. Math.*, 1951.
R. J. AUMANN, L. S. SHAPLEY, *Values of Non-Atomic Games*, 1968.

An example from urban use of land

A number of agents must choose where to live in an urban region $\Omega \subset \mathbb{R}^d$; $\rho \in \mathcal{P}(\Omega)$ is their density. Every agent considers the sum of three costs:

- an exogenous cost, depending on the amenities of x only: $V(x)$ (distance to the points of interest...);
- an interaction cost, depending on the distances with all the others; when living at x the cost is $\int W(x-y)\rho(y)\,dy$ where W is usually an increasing function of the distance;
- a residential cost, increasing in the density at x; at x the cost is $h(\rho(x))$, for $h : \mathbb{R}_+ \to \mathbb{R}$ increasing (where more people live, the price of land is higher or, for the same price, they have less space).

The total cost that we consider is $f_\rho(x) := V(x) + (W * \rho)(x) + h(\rho(x))$.

M.J. Beckmann. Spatial equilibrium and the dispersed city, *Mathematical Land Use Theory,* 1976.

M. Fujita and J. F. Thisse. *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth.* 2002.

Equilibrium and optimality conditions

Consider the following quantity

$$\mathcal{F}(\rho) := \int_\Omega V(x)\rho(x)dx + \frac{1}{2}\int_\Omega\int_\Omega W(x-y)\rho(x)\rho(y)dxdy + \int_\Omega H(\rho(x))dx,$$

where H is defined through $H' = h$.
We can see that $f_\rho := V + W * \rho + h(\rho) = \frac{\delta\mathcal{F}}{\delta\rho}$.

Warning: the energy $\mathcal{F}$ is not the total cost for all the agents, which should be $\int_\Omega f_\rho(x)\rho(x)dx$.
Minimizers are equilibria. What about the converse? In general $\mathcal{W}$ is not convex (think at $W(z) = |z|^2$, so that $\frac{1}{2}\mathcal{W}(\rho) = \int |x|^2 d\rho(x) - (\int x d\rho(x))^2$).

Theorem

If $\mathcal{F}$ is geodesically convex, then minimizers of $\mathcal{F}$ and equilibria for f_ρ coincide.

A. Blanchet, P. Mossay, F. Santambrogio Existence and uniqueness of equilibrium for a spatial model of social interactions, *Int. Econ. Rev.*, 2015.

About the residential cost

Suppose that agents have a certain budget to be divided into land consumption and money consumption, and that they have a concave and increasing utility function U for land. They solve a problem of the form

$$\max\{U(L) + m \,:\, pL + m \leq B\},$$

where p represents the price for land, L is the land consumption, m is the left-over of the money, and B the budget constraint. The optimal land consumption will be such that $U'(L) = p$. The optimal utility is $B + U(L) - U'(L)L$ (relation between L and utility).
The land consumption is the reciprocal of the density, hence $L = \frac{1}{\rho}$, and the residential cost $h(\rho)$, which is the opposite of the utility, is

$$h(\rho) = \frac{1}{\rho}U'\left(\frac{1}{\rho}\right) - U\left(\frac{1}{\rho}\right) - B.$$

$\frac{1}{t}U'(\frac{1}{t}) - U(\frac{1}{t}) = \left(-tU(\frac{1}{t})\right)'$, hence $h = H'$ with $H(t) = -tU(\frac{1}{t}) - Bt$.
H satisfies McCann's condition if and only $s \mapsto U(s^d)$ is concave and increasing.

Cournot-Nash urban equilibria

Cournot-Nash equilibria concern nonatomic games where agents are not indistinguishable. Consider an extra variable $z \in Z$, the origin of agents, and $\mu \in \mathcal{P}(Z)$ its distribution. Agents must choose a location $x \in \Omega$, and their payoff is given by $c(z,x) + f_\rho(x)$, where ρ is the distribution of their choices on Ω. We look for a triple (γ, ρ, φ) with $\gamma \in \Pi(\mu,\rho), \varphi(z) = \min_x c(z,x) + f_\rho(x)$ and $\varphi(z) = c(z,x) + f_\rho(x)$ on $\mathrm{spt}(\gamma)$.
This condition may be obtained by solving

$$\min\left\{\int c\, d\gamma + \mathcal{F}((\pi_x)_\#\gamma) \ :\ (\pi_z)_\#\gamma = \mu\right\}.$$

In the case $c(x,y) = |x-y|^p$ this means $\min_\rho W_p^p(\mu,\rho) + \mathcal{F}(\rho)$. The Kantorovich potentials will be of the form $(\varphi, -f_\rho)$.
Convexity: the extra term $W_p^p(\mu,\rho)$ is convex in ρ. It is in general non geodesically convex (!!) but it is convex along generalized geodesics $\rho_t = ((1-t)T_0 + tT_1)_\#\mu$ (with T_i the optimal map from μ to ρ_i, $i = 0, 1$).

A. Blanchet, G. Carlier, Optimal transport and Cournot-Nash Equilibria, *Math. Op. Res.*, 2016
L. Ambrosio, N. Gigli, G. Savaré *Gradient Flows*, 2005

Variants

Congestion problems – Martingale transport

Congestion and continuous Wardrop equilibria

Consider $\min\{\int |w| \ : \ \nabla \cdot w = \mu - \nu\} = \min$**(KP)**. If we insert a weight, $\min\{\int k(x)|w| \ : \ \nabla \cdot w = \mu - \nu\}$ becomes equivalent to **(KP)** with the cost c given by the distance d_k $\left(d_k(x,y) := \min\left\{\int_0^1 k(\omega)|\omega'| : \omega(0) = x, \omega(1) = y\right\}\right)$.

Congestion and continuous Wardrop equilibria

Consider $\min\{\int|w| \,:\, \nabla\cdot w = \mu-\nu\} = \min$**(KP)**. If we insert a weight, $\min\{\int k(x)|w| \,:\, \nabla\cdot w = \mu-\nu\}$ becomes equivalent to **(KP)** with the cost c given by the distance d_k ($d_k(x,y) := \min\left\{\int_0^1 k(\omega)|\omega'| : \omega(0)=x, \omega(1)=y\right\}$). But in congested urban regions, k mainly depends on traffic itself, i.e. on w. Hence, we should consider the superlinear Beckmann's problem

$$\min\left\{\int H(|w|) \,:\, \nabla\cdot w = \mu-\nu\right\},$$

which is connected to degenerate elliptic equations $\nabla\cdot\left(H'(|\nabla\phi|)\frac{\nabla\phi}{|\nabla\phi|}\right) = \mu-\nu$, and to equilibrium problems where the cost per unit length of a path is $H'(i)$ and i is the traffic intensity.

J. G. Wardrop, Some theoretical aspects of road traffic research, *Proc. Inst. Civ. Eng.*, 1952.
G. Carlier, C. Jimenez, F. Santambrogio, Optimal transportation with traffic congestion and Wardrop equilibria, *SIAM J. Contr. Optim.* 2008.
L. Brasco, G. Carlier, F. Santambrogio, Congested traffic dynamics, weak flows and very degenerate elliptic equations, *J. Math. Pures et Appl.*, 2010.

MFG with density penalization

A vairant of the Benamou-Brenier problem: for $G(x,\cdot)$ convex, minimize

$$\mathcal{A}(\rho, v) := \int_0^T \int_\Omega \left(\frac{1}{2}\rho_t |v_t|^2 + G(x, \rho_t) \right) + \int_\Omega \Psi \rho_T$$

among pairs (ρ, v) such that $\partial_t \rho + \nabla\cdot(\rho v) = 0$, with given ρ_0. The solution is given by $v = -\nabla\phi$, where (ρ, ϕ) solves a coupled system:

$$\begin{cases} -\partial_t \varphi + \frac{|\nabla\varphi|^2}{2} = G'(x,\rho), & \varphi(T,x) = \Psi(x), \\ \partial_t \rho - \nabla\cdot(\rho\nabla\varphi) = 0, & \rho(0,x) = \rho_0(x). \end{cases}$$

Every agent minimizes $\int_0^T \left[\frac{1}{2}|\omega'(t)|^2 + G'(\omega(t), \rho_t(\omega(t))) \right] dt + \Psi(\omega(T))$: the running cost depends on the density ρ_t realized by the agents themselves.

J.-M. Lasry, P.-L. Lions, Mean-Field Games, *Japan. J. Math.* 2007
P.-L. Lions, courses at Collège de France, 2006/12, videos available online
P. Cardaliaguet, lecture notes, www.ceremade.dauphine.fr/~cardalia/
G. Buttazzo, C. Jimenez, E. Oudet An optimization problem for mass transportation with congested dynamics *SICON*, 2009
J.-D. Benamou, G. Carlier, F. Santambrogio, Variational Mean Field Games, *Active Particles I,* 2016

Option pricing

Let $Z = f(S_T^1, S_T^2, \dots, S_T^N)$ be an option with payoff depending on several assets S_T^i at a same time T. Its value is $\mathbb{E}^{\mathbb{Q}}[Z]$ for a probability Q which makes each price a martingale. Suppose that Z is new on the market, but that the European Calls on each S^i already exist for each strike price K. We know the law μ_i of each of the S_T^i, but not their joint law (under Q): estimating the price of Z is a **multi-marginal transport problem**

$$\min / \max \left\{ \int f(x_1, x_2, \dots, x_N)\, d\gamma \ : \ (\pi_i)_\# \gamma = \mu_i \right\}$$

If on the contrary, $Z = f(S_1, S_2, \dots, S_T)$ depends on the history of a same asset (and the corresponding call options $(S_t - K)_+$ exist on the market for every strike and every maturity time), then the problem is different

$$\min / \max \left\{ \mathbb{E}^{\mathbb{Q}}[f(S_1, \dots, S_N)] \ : \ (S_i)_\# \mathbb{Q} = \mu^i, \ \ (S_t)_t \text{ is a } \mathbb{Q}\text{-martingale} \right\}.$$

Kantorovich problems with the martingale constraints are a new frontier of optimal transport!

M. Beiglböck, N. Juillet On a problem of optimal transport under marginal martingale constraints, *Ann. Prob.*, 2016

Robust super-hedging

The same problem can be read in this way: suppose we want to hedge a possible loss of $f(S_1, \dots, S_N)$, depending on the prices of the asset S. What we can do:

- buy usual options, based on the value of the asset at time t, i.e. buy $\phi(S_t)$, paying $\int \phi \, d\mu_t$;
- buy the asset itself at time $t \le N-1$, pay S_t, and re-sell at price S_{t+1} (but the number of assets to buy can only be chosen according to the information at time t, i.e. it must be of the form $\psi(S_1, \dots, S_t)$, which hedges an amount $\psi(S_1, \dots, S_t)(S_{t+1} - S_t)$.

The optimal hedging problem becomes

$$\min \left\{ \sum_{t=0}^{N} \int \varphi_t \, d\mu_t \, : \, \sum_{t=0}^{N} \varphi_t(x_t) + \sum_{t=0}^{N-1} \psi(x_1, \dots, x_t)(x_{t+1} - x_t) \ge f(x_1, \dots, x_N) \right\}.$$

This problem is exactly the dual of the martingale transport problem (the ϕ_t dualize the marginal constraints, the ψ_t the martingale constraints).

A. Galichon, P. Henry-Labordère, N. Touzi, A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options *Ann. Appl. Prob.*, 2014.

That's all for this short presentation of OT and some of its applications

Thanks for your attention