On equilibrium on quality markets: optimal transport and beyond

Guillaume Carlier [a]

Workshop on Optimal Transportation, Equilibrium, and applications to Economics, NYU, April 2016.

[a]CEREMADE, Université Paris Dauphine et MOKAPLAN (Inria-Dauphine).

Introduction

Aims of this talk:

- survey some links between optimal transport and equilibrium issues on quality good markets,
- introduce a more realistic framework for consumers' preferences (work in progress with I. Ekeland and A. Galichon), still related to (a variant of) optimal transport.

Outline

① Kantorovich duality as an equilibrium problem

② Matching

③ Adding noise and using it for numerics

④ Non quasi-linearity of preferences on the demand side

Kantorovich duality as an equilibrium problem

Source Y a compact metric space endowed with a probability, $\nu \in \mathcal{P}(Y)$, target, Z, another compact metric space with $\eta \in \mathcal{P}(Z)$, *transport* cost c: $Y \times Z \to \mathbf{R}$ (continuous, say), Monge-Kantorovich optimal transport problem

$$\inf_{\sigma \in \Pi(\nu,\eta)} \int_{Y \times Z} c(y,z) \mathrm{d}\sigma(y,z) \tag{1}$$

where $\Pi(\nu,\eta)$ is the set of probability measures on $Y \times Z$ having ν and η as marginals.

This is an a priori infinite-dimensional linear programming problem, whose dual reads

$$\sup_{\varphi \in C(Z)} \int_Y \varphi^c(y) \mathrm{d}\nu(y) + \int_Z \varphi(z) \mathrm{d}\eta(z)$$

with

$$\varphi^c(y) = \min_{z \in Z} \{c(y,z) - \varphi(z)\}.$$

Kantorovich duality: the least transport cost coincides with the value of the dual, both values are attained. Optimality conditions (complementary slackness): an optimal transport plan σ is concentrated on the set of pairs (y, z) for which

$$\varphi^c(y) = c(y,z) - \varphi(z)$$

Interpretation: fixed demand distribution η for goods of type z, producers have type y, and production cost c, Kantorovich potential φ: price system. A Kantorovich potential φ is an equilibrium price system, it equalizes the (fixed, here) demand distribution η with the supply distribution resulting from producers each chosing a profit maximizing quality z.

Semi-discrete case $Z = \{z_1, \cdots, z_J\}$, $\eta = \sum_{j=1}^{J} \eta_j \delta_{z_j}$, equilibrium/Kantorovich price system $\varphi(z_j) = p_j$ such that

$$\eta_j = \mu(\{y \in Y \ : \ c(y, z_j) - p_j \leq c(y, z_k) - p_k, \ \forall k\})$$

which is the optimality condition for

$$\sup_p \sum_j p_j \eta_j + \int_Y \min_j (c(y, z_j) - p_j) \mathrm{d}\mu(y).$$

I'm cheating a little ignoring ties...

Variant (somehow similar to discrete choice models): both Y and Z are finite, $\nu = \sum_i \nu_i \delta_{y_i}$, $\eta = \sum_{j=1} \eta_j \delta_{z_j}$, there is some random term ε_{ij} on the cost c_{ij}, equilibrium:

$$\eta_j = \sum_i \mu_i \mathbf{P}\Big(c_{ij} - p_j + \varepsilon_{ij} = V_i(p)\Big)$$

where $V_i(p)$ is the (random) least net cost

$$V_i(p) := \min_j \{c_{ij} - p_j + \varepsilon_{ij}\}.$$

Equilibrium conditions= Euler-Lagrange equations for

$$\max_p \sum_j p_j \eta_j + \sum_i \mu_i \mathbf{E}(V_i(p)).$$

Matching

Previously, the demand side was exogenous. Now we shall recall, following the seminal works of Gretsky, Ostroy and Zame, Chiappori, McCann and Nesheim, Ekeland, that introducing the demand side and assuming quasi-linear utility for consumers, the equilibrium issue is again an OT problem (in the tradition of the LP approach of Shapley-Shubik to the core of assignment games).

Now the quality line measure $\eta \in \mathcal{P}(Z)$ on the quality space Z is unknown. Supply side is as before and given by the producers's type space Y, the distribution $\nu \in \mathcal{P}(Y)$ and the cost function $c : Y \times Z \to \mathbf{R}$, given a price system $\varphi\colon Z \to \mathbf{R}$, y-type producers solve

$$\varphi^c(y) = \min_{z \in Z}\{c(y,z) - \varphi(z)\}.$$

Demand side: consumers have a type space X (compact metric), type distribution is given by $\mu \in \mathcal{P}(X)$, quasi-linear utility

$$U(x,z) - \varphi(z)$$

with $U \in C(X \times Z)$. Quasi-linearity is convenient but it is of course a strong (and actually questionable assumption). Here for simplicity $\mu(X) = \nu(Y) = 1$, total demand equals total demand.

Equilibrium: a quality line measure $\eta \in \mathcal{P}(Z)$, together with plans

- $\gamma \in \Pi(\mu, \eta)$, $\sigma \in \Pi(\nu, \eta)$ (same marginal on the quality space),
- a price system $\varphi : Z \to \mathbf{R}$ such that:

  for σ-a.e. (y, z):

$$\varphi^c(y) = c(y, z) - \varphi(z) \tag{2}$$

  and for γ-a.e. (x, z)

$$(-\varphi)^{-U}(x) = \varphi(z) - U(x, z) \tag{3}$$

  where

$$(-\varphi)^{-U}(x) = \min_{z \in Z}\{\varphi(z) - U(x, z)\}.$$

For $\eta \in \mathcal{P}(Z)$ define

$$F_c(\eta) := \inf_{\sigma \in \Pi(\nu,\eta)} \int_{Y \times Z} c(y,z)\mathrm{d}\sigma(y,z)$$

and

$$G_U(\eta) := \sup_{\gamma \in \Pi(\mu,\eta)} \int_{X \times Z} U(x,z)\mathrm{d}\gamma(y,z)$$

and

$$\inf_{\eta \in \mathcal{P}(Z)} J(\eta) := F_c(\eta) - G_U(\eta) \tag{4}$$

as well as (its dual)

$$\sup_{\varphi \in C(Z)} \int_Y \varphi^c \mathrm{d}\nu + \int_X (-\varphi^{-U})\mathrm{d}\mu. \tag{5}$$

Chiappori, McCann and Nesheim, Ekeland made it clear that this is equivalent to OT.

Then $\eta, \varphi, \gamma, \sigma$ is an equilibrium if and only if

- η solves (4),
- φ solves (5),
- σ is optimal between ν and η i.e.

$$F_c(\eta) = \int_{Y \times Z} c(y, z) \mathrm{d}\sigma(y, z)$$

- γ is optimal between μ and η i.e.

$$G_U(\eta) = \int_{X \times Z} U(x, z) \mathrm{d}\gamma(x, z).$$

The convex program (4) is actually equivalent to a classical OT problem, for the indirect cost function (negative of the surplus):

$$b(x,y) := \min_{z\in Z}\{c(y,z) - U(x,z)\}$$

namely

$$\inf_{\theta\in\Pi(\mu,\nu)} \int_{X\times Y} b(x,y)\mathrm{d}\theta(x,y)$$

Note that (4) can also be written in LP form as

$$\inf_{(\sigma,\gamma)} \left\{ \int_{Y\times Z} c(y,z)\mathrm{d}\sigma(y,z) - \int_{X\times Z} U(x,z)\mathrm{d}\gamma(x,z) \right\}$$

subject to the constraints that the first marginal of σ is ν, the first marginal of γ is μ and the second marginals of σ and γ coincide.

Adding noise and its use for numerics

An old idea (goes back to Schrödinger in the early 1930's, simulated annealing, Gibbs measures, soft max, entropic penalization of LP...) very nicely *remise au goût du jour* recently by Alfred Galichon and Bernard Salanié in a matching context and by Marco Cuturi for computational purposes in machine learning.

The rough idea is the following: in OT one looks for a plan $\theta \in \Pi(\mu, \nu)$ concentrated on the minimal set of $b(x, y) - f(x) - g(y)$, approximation, small $\varepsilon > 0$ and rather look for a plan $\theta \in \Pi(\mu, \nu)$ having a density proportional to the Gibbs density

$$e^{-\frac{1}{\varepsilon}(b(x,y) - f(x) - g(y))}.$$

Amounts to adding some noise in the cost (see Alfred and Bernard's papers which to my knowledge are the first to propose a tractable theory for inverse problems based on OT, should be useful in machine learning...).

We want to solve numerically a discrete (say) OT problem like

$$\inf\Big\{\sum_{i,j} b_{ij}\theta_{ij} \ : \ \theta_{ij} \geq 0, \sum_j \theta_{ij} = \mu_i, \ \sum_i \theta_{ij} = \nu_j\Big\} \qquad (6)$$

or a discrete version of (4)

$$\inf_{\gamma_{jk},\sigma_{ik}} \Big\{\sum_{j,k} c_{jk}\sigma_{jk} - \sum_{i,k} U_{ik}\gamma_{ik}\Big\}$$

subject to $\gamma_{jk} \geq 0$, $\sigma_{ik} \geq 0$,

$$\sum_k \sigma_{jk} = \nu_j, \ \sum_k \gamma_{ik} = \mu_i, \ \sum_j \sigma_{jk} = \sum_i \gamma_{ik}.$$

Approximate (6) by adding ε times the Boltzmann entropy of the plan:

$$\inf_{\theta\in\Pi(\mu,\nu)}\Big\{\sum_{ij} b_{ij}\gamma_{ij}+\varepsilon\sum_{ij}\theta_{ij}\log(\theta_{ij})\Big\}$$

this becomes a Kullback-Leibler projection problem

$$\inf_{\theta\in\Pi(\mu,\nu)}\mathrm{KL}(\theta|g_\varepsilon)$$

where $(g_\varepsilon)_{ij} := e^{-\frac{b_{ij}}{\varepsilon}}$ and

$$\mathrm{KL}(\theta|g) := \sum_{i,j}\theta_{ij}\log\left(\frac{\theta_{ij}}{g_{ij}}\right)$$

Consider the KL projection problem

$$\inf_{\theta\in\Pi(\mu,\nu)} \mathrm{KL}(\theta|g)$$

it has a unique solution which is characterized by

$$\theta_{ij} = a_i b_j g_{ij}$$

with a_i, b_j positive and such that the marginal constraints are met i.e.

$$a_i = \frac{\mu_i}{\sum_j b_j g_{ij}} := R_i(b),\ b_j = \frac{\nu_j}{\sum_i a_i g_{ij}} := S_j(a)$$

which can be rewritten as a fixed point-problem on a only: find a in the positive cone such that $a = Ta$ (with $T = R \circ S$).

The IPFP procedure (Iterative Proportional Fitting Procedure aka Sinkhorn) consists in iterating the map T (it can also be seen as an alternating KL projection algorithm where one projects alternatively on the first and second marginals constraints, such projections being explicit). Why does it converge? It can be seen as a particular case of Bauschke and Borwein convergence result for alternating projections but there is another powerful viewpoint (which also works in the continuous case as shown by Chen, Georgiou and Pavon) based on the Hilbert projective metric

$$d_H(a, a') := \log\Big(\frac{\max_i \frac{a_i}{a'_i}}{\min_i \frac{a_i}{a'_i}}\Big), \ (a, a') \in (0, +\infty)^N.$$

for which T is a contraction.

Comments on IPFP:

- flexible and easy to implement, parallelizable, each step only stores a vector with size N (and not N^2 for mesaures discretized with N points),
- in practice for really large scale problems, still much too costly, needs to combine IPFP with other ideas (e.g. MCMCs or grid refinement, Adam Oberman)
- IPFP easily adapts to unbalanced problems, multi-marginals, OT with capacity constraints (McCann, Korman)...

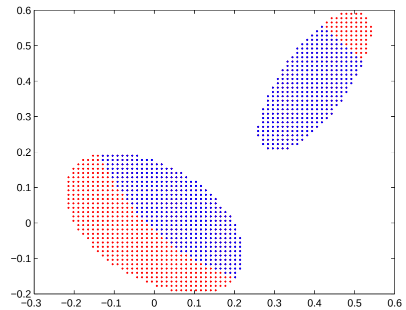
Partial OT, two measures, μ and ν (with possibly diferent masses), $m \leq \min(\mu(X), \nu(Y))$ consider

$$\inf_{\gamma \in \Pi_m^-(\mu,\nu)} \langle c, \gamma \rangle$$

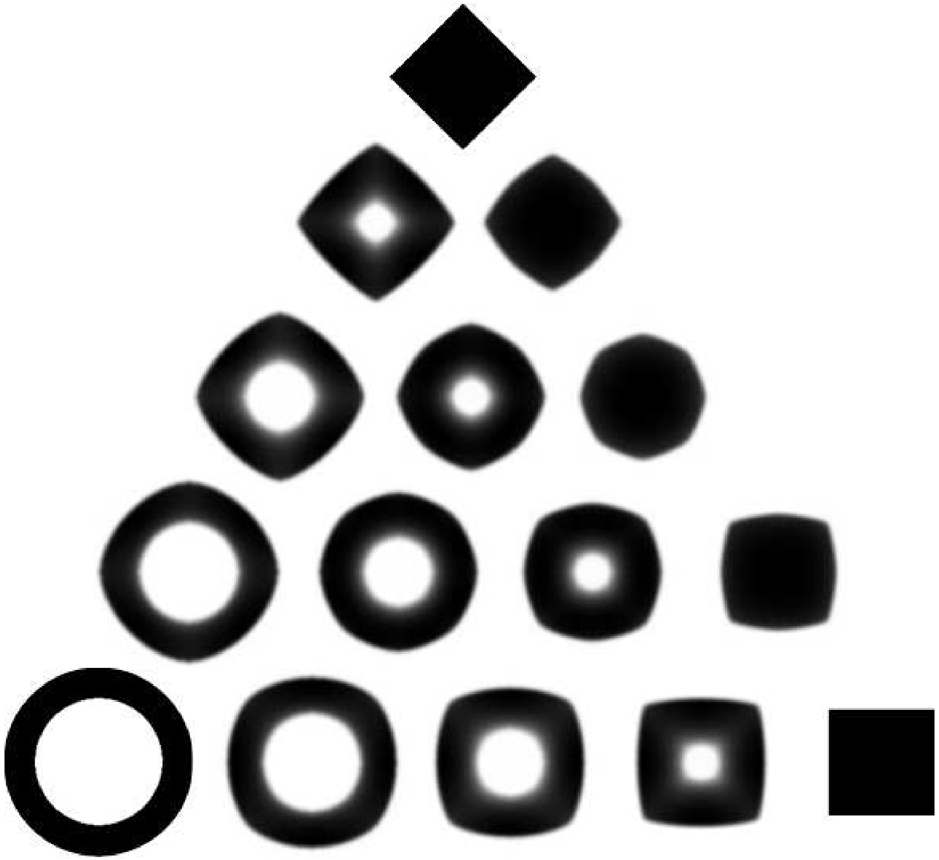
with

$$\Pi_m^-(\mu,\nu) := \{\gamma \in \mathcal{P}(X \times Y) \ : \ \pi_{1\#}\gamma \leq \mu, \ \pi_{2\#}\gamma \leq \nu, \ \gamma(X \times Y) = m\}$$

intersection of three convex sets on which KL projections are explicit:



Computing barycenters by IPFP:



Non separability on the demand side

Not very realistic to assume that consumers have quasi-linear utility. The assumptions on the supply side are as before: producers type space Y, cost function $c(z, y)$ ($z \in Z$, quality space) and $\nu \in \mathcal{P}(Y)$ distribution of producer's type. Again quasi-linear specification for producers, given prices $z \mapsto \varphi(z)$, y-producers chose to minimize $c(y, z) - \varphi(z)$ as before.

The demand side is now a bit more complicated, consumers type consists of two parameters $x = (\theta, w) \in X$ with $w \in [a, b] \subset \mathbf{R}_+$, revenue, and θ is a preference parameter, now consumer of type $x = (\theta, w)$ solves

$$V_\varphi(x) := \max_{z \in Z} \{U(\theta, z) \ : \ \varphi(z) \leq w\}$$

and the distribution of $x = (\theta, w)$ is given and again denoted by $\mu \in \mathcal{P}(X)$. Setting $w = w(x)$, we see that this is a special case of

$$V_\varphi(x) := \max_{z \in Z} \{U(x, z) \ : \ \varphi(z) \leq w(x)\}$$

with w a given revenue function from X to $[a, b]$.

In this context an equilibrium consists of a price $\varphi \in C(Z, \mathbf{R}_+)$ such that $\min_Z \varphi = \min_X w = a$, a quality line $\eta \in \mathcal{P}(Z)$ as well as consumer-quality and producer-quality couplings γ and σ such that

1. $\gamma \in \Pi(\mu, \eta)$, $\sigma \in \Pi(\nu, \eta)$,
2. for γ-a.e. (x, z), one has
$$\varphi(z) \leq w(x); \quad \text{and } V_\varphi(x) = U(x, z) \tag{7}$$
3. for σ-a.e. (y, z), one has
$$\varphi(z) + \varphi^c(y) = c(y, z). \tag{8}$$

Existence of an equilibrium, can be established under quite general assumptions on the data (C., Ekeland, Galichon), V_φ might be discontinuous wrt w but it is nondecreasing. We are currently investigating uniqueness of the equilibrium price. But, I would like to emphasize now in a heuristic way, some OT-like problem related to the demand side. On the supply side it is obvious that if η is an equilibrium quality line then the coupling σ should be optimal between ν and η. The question is whether the consumer/quality coupling γ also solves an OT-like problem.

Given distribution of consumers $\mu \in \mathcal{P}(X)$ and distribution of goods $\eta \in \mathcal{P}(Y)$, can we find a plan $\gamma \in \Pi(\mu, \eta)$ and a price systeml φ such that

$$\varphi(z) \leq w(x),\ V_\varphi(x) = U(x,z) \text{ for } \gamma\text{-a.e. } (x,z)? \tag{9}$$

If such a pair exists it has to minimize

$$\int_X V_\varphi(x) \mathrm{d}\mu(x) - \int_{X \times Z} U(x,z) \mathrm{d}\gamma(x,z)$$

with respect to $\gamma \in \Pi(\mu, \eta)$ and φ subject to the constraint that

$$\varphi(z) \leq w(x) \ \gamma\text{-a.e.}$$

Since V_φ is non increasing with respect to φ, the best to do to minimize the first integral is to take

$$\varphi(z) = \operatorname{essinf}_{\gamma^z} w := \varphi_\gamma$$

This becomes a problem in γ only:

$$\inf_{\gamma\in\Pi(\mu,\eta)} \int_X V_{\varphi_\gamma}(x)\mathrm{d}\mu(x) - \int_{X\times Z} U(x,z)\mathrm{d}\gamma(x,z)$$

where (formally) $V_{\varphi_\gamma}(x)$ is the sup of $U(x,z)$ over all z's for which there is an x' with $w(x') \leq w(x)$ and (x',z) belongs to the support of γ.