Topics in Equilibrium Transportation

Alfred Galichon (NYU and Sciences Po)

GSI, Ecole polytechnique, October 29, 2015
This talk is based on the following two papers:

Agenda:

1. Economic motivation
2. The mathematical problem
3. Computation
4. Estimation
Agenda:

1. Economic motivation
2. The mathematical problem
3. Computation
4. Estimation
Agenda:

1. Economic motivation
2. The mathematical problem
3. Computation
4. Estimation
Agenda:

1. Economic motivation
2. The mathematical problem
3. Computation
4. Estimation
Section 1

ECONOMIC MOTIVATION
Consider a very simple model of labour market. Assume that a population of workers is characterized by their type $x \in \mathcal{X}$, where $\mathcal{X} = \mathbb{R}^d$ for simplicity. There is a distribution $P$ over the workers, which is assumed to sum to one.

A population of firms is characterized by their types $y \in \mathcal{Y}$ (say $\mathcal{Y} = \mathbb{R}^d$), and their distribution $Q$. It is assumed that there is the same total mass of workers and firms, so $Q$ sums to one.

Each worker must work for one firm; each firm must hire one worker. Let $\pi (x, y)$ be the probability of observing a matched $(x, y)$ pair. $\pi$ should have marginal $P$ and $Q$, which is denoted

$$
\pi \in \mathcal{M} (P, Q).
$$
In the simplest case, the utility of a worker $x$ working for a firm $y$ at wage $w(x, y)$ will be

$$\alpha (x, y) + w(x, y)$$

while the corresponding profit of firm $y$ is

$$\gamma (x, y) - w(x, y).$$

In this case, the total surplus generated by a pair $(x, y)$ is

$$\alpha (x, y) + w + \gamma (x, y) - w = \alpha (x, y) + \gamma (x, y) =: \Phi (x, y)$$

which does not depend on $w$ (no transfer frictions). A central planner may thus like to choose assignment $\pi \in M(P, Q)$ so to

$$\max_{\pi \in M(P,Q)} \int \Phi (x, y) d\pi (x, y).$$

But why would this be the equilibrium solution?
The equilibrium assignment is determined by an important quantity: the wages. Let $w(x, y)$ be the wage of employee $x$ working for firm of type $y$.

Let the indirect surpluses of worker $x$ and firm $y$ be respectively

$$u(x) = \max_y \{ \alpha(x, y) + w(x, y) \}$$
$$v(y) = \max_x \{ \gamma(x, y) - w(x, y) \}$$

so that $(\pi, w)$ is an equilibrium when

$$u(x) \geq \alpha(x, y) + w(x, y) \text{ with equality if } (x, y) \in \text{Supp}(\pi)$$
$$v(y) \geq \gamma(x, y) - w(x, y) \text{ with equality if } (x, y) \in \text{Supp}(\pi)$$

By summation,

$$u(x) + v(y) \geq \Phi(x, y) \text{ with equality if } (x, y) \in \text{Supp}(\pi).$$
One can show that the equilibrium outcome \((\pi, u, v)\) is such that \(\pi\) is solution to the primal Monge-Kantorovich Optimal Transportation problem

\[
\max_{\pi \in \mathcal{M}(P,Q)} \int \Phi(x, y) \, d\pi(x, y)
\]

and \((u, v)\) is solution to the dual OT problem

\[
\min_{u,v} \int u(x) \, dP(x) + \int v(y) \, dQ(y)
\]

\[
s.t. \ u(x) + v(y) \geq \Phi(x, y)
\]

Feasibility + Complementary slackness yield the desired equilibrium conditions

\[
\pi \in \mathcal{M}(P, Q)
\]

\[
u(x) + \Phi(y) \geq \Phi(x, y)
\]

\[
(x, y) \in \text{Supp}(\pi) \implies u(x) + v(y) = \Phi(x, y)
\]

“Second welfare theorem”, “invisible hand”, etc.
Is equilibrium always the solution to an optimization problem?

It is not. This is why this talk is about “Equilibrium Transportation,” which contains, but is strictly more general than “Optimal Transportation”.
Is equilibrium always the solution to an optimization problem?

It is not. This is why this talk is about “Equilibrium Transportation,” which contains, but is strictly more general than “Optimal Transportation”.
Consider the same setting as above, but instead of assuming that workers’ and firm’s payoffs are linear in surplus, assume

$$u(x) = \max_y \{U_{xy}(w(x, y))\}$$

$$v(y) = \max_x \{V_{xy}(w(x, y))\}$$

where $U_{xy}(w)$ is nondecreasing and continuous, and $V_{xy}(w)$ is nonincreasing and continuous.

Motivation: taxes, decreasing marginal returns, risk aversion, etc. Of course, Optimal Transportation case is recovered when

$$U_{xy}(w) = \alpha_{xy} + w$$

$$V_{xy}(w) = \gamma_{xy} - w.$$
For \((u, v) \in \mathbb{R}^2\), let

\[
\Psi_{xy}(u, v) = \min \{ t \in \mathbb{R} : \exists w, u - t \leq U_{xy}(w) \text{ and } v - t \leq V_{xy}(w) \}
\]

so that \(\Psi\) is nondecreasing in both variables and \((u, v) = (U_{xy}(w), V_{xy}(w))\) for some \(w\) if and only if \(\Psi_{xy}(u, v) = 0\). Optimal Transportation case is recovered when

\[
\Psi_{xy}(u, v) = \frac{u + v - \Phi_{xy}}{2}.
\]

As before, \((\pi, w)\) is an equilibrium when

\[
\begin{align*}
u(x) &\geq U_{xy}(w(x, y)) \text{ with equality if } (x, y) \in \text{Supp}(\pi) \\
v(y) &\geq V_{xy}(w(x, y)) \text{ with equality if } (x, y) \in \text{Supp}(\pi)
\end{align*}
\]

We have therefore that \((\pi, u, v)\) is an equilibrium when

\[
\Psi_{xy}(u(x), v(y)) \geq 0 \text{ with equality if } (x, y) \in \text{Supp}(\pi).
\]
Section 2

THE MATHEMATICAL PROBLEM
We have therefore that \((\pi, u, v)\) is an equilibrium outcome when

\[
\begin{cases}
\pi \in \mathcal{M}(P, Q) \\
\Psi_{xy}(u(x), v(y)) \geq 0 \\
(x, y) \in \text{Supp}(\pi) \implies \Psi_{xy}(u(x), v(y)) = 0
\end{cases}
\]

Problem: existence of an equilibrium outcome? This paper: yes in the discrete case \((\mathcal{X} \text{ and } \mathcal{Y} \text{ finite}), \text{ via entropic regularization.}\)
As soon as $\Psi_{xy}$ is strictly increasing in both variables, $\Psi_{xy}(u, v) = 0$ expresses as

$$u = G_{xy}(v) \text{ and } v = G_{xy}^{-1}(u)$$

where the generating functions $G_{xy}$ and $G_{xy}^{-1}$ are decreasing and continuous functions. In this case, relations

$$u(x) = \max_{y \in Y} G_{xy}(v(y)) \text{ and } v(y) = \max_{x \in X} G_{xy}^{-1}(u(x))$$

generalize the Legendre-Fenchel conjugacy. This pair of relations form a Galois connection; see Singer (1997) and Noeldeke and Samuelson (2015).
Remark 2: Trudinger’s local theory of prescribed Jacobians

Assuming everything is smooth, and letting $f_P$ and $f_Q$ be the densities of $P$ and $Q$ we have under some conditions that the equilibrium transportation plan is given by $y = T(x)$, where mass balance yields

$$|\det DT(x)| = \frac{f(x)}{g(T(x))}$$

and optimality yields

$$\partial_x G_{xT(x)}^{-1}(u(x)) + \partial_u G_{xT(x)}^{-1}(u(x)) \nabla u(x) = 0$$

which thus inverts into

$$T(x) = e(x, u(x), \nabla u(x)).$$

Trudinger (2014) studies Monge-Ampere equations of the form

$$|\det De(., u, \nabla u)| = \frac{f}{g(e(., u, \nabla u))}.$$

(more general than Optimal Transport where no dependence on $u$).
Our work (GKW 2015a and b) focuses on the discrete case, when $P$ and $Q$ have finite support. Call $p_x$ and $q_y$ the mass of $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ respectively.

In the discrete case, problem boils down to looking for $(\pi, u, \nu)$ such that

\[
\begin{align*}
\pi_{xy} &\geq 0, \quad \sum_y \pi_{xy} = p_x, \quad \sum_x \pi_{xy} = q_y \\
\Psi_{xy}(u_x, \nu_y) &\geq 0 \\
\pi_{xy} > 0 &\implies \Psi_{xy}(u_x, \nu_y) = 0
\end{align*}
\]
Section 3

COMPUTATION
Take temperature parameter $T > 0$ and look for $\pi$ under the form

$$\pi_{xy} = \exp \left( - \frac{\Psi_{xy}(u_x, v_y)}{T} \right)$$

Note that when $T \to 0$, the limit of $\Psi_{xy}(u_x, v_y)$ is nonnegative, and the limit of $\pi_{xy} \Psi_{xy}(u_x, v_y)$ is zero.
If $\pi_{xy} = \exp \left( -\Psi_{xy}(u_x, v_y) / T \right)$, condition $\pi \in \mathcal{M}(P, Q)$ boils down to set of nonlinear equations in $(u, v)$

$$
\begin{align*}
\sum_{y \in Y} \exp \left( \frac{- \Psi_{xy}(u_x, v_y)}{T} \right) &= p_x \\
\sum_{x \in X} \exp \left( \frac{- \Psi_{xy}(u_x, v_y)}{T} \right) &= q_y
\end{align*}
$$

which we call the \textit{nonlinear Bernstein-Schrödinger equation}.

In the optimal transportation case, this becomes the classical B-S equation

$$
\begin{align*}
\sum_{y \in Y} \exp \left( \frac{\Phi_{xy} - u_x - v_y}{2T} \right) &= p_x \\
\sum_{x \in X} \exp \left( \frac{\Phi_{xy} - u_x - v_y}{2T} \right) &= q_y
\end{align*}
$$
Note that $F_x : u_x \to \sum_{y \in Y} \exp \left( - \frac{\Psi_{xy}(u_x, v_y)}{T} \right)$ is a decreasing and continuous function. Mild conditions on $\Psi$ therefore ensure the existence of $u_x$ so that $F_x(u_x) = p_x$.

Our algorithm is thus a nonlinear Jacobi algorithm:
- Make an initial guess of $v_y^0$
- Determine the $u_x^{k+1}$ to fit the $p_x$ margins, based on the $v_y^k$
- Update the $v_y^{k+1}$ to fit the $q_y$ margins, based on the $u_x^{k+1}$
- Repeat until $v_y^{k+1}$ is close enough to $v_y^k$.

One can proof that if $v_y^0$ is high enough, then the $v_y^k$ decrease to fixed point. Convergence is very fast in practice.
Section 4

Statistical Estimation
In practice, one observes $\hat{\pi}_{xy}$ and would like to estimate $\Psi$. Assume that $\Psi$ belongs to a parametric family $\Psi^\theta$, so that $\pi^\theta_{xy} = \exp(-\Psi^\theta_{xy}(u^\theta_x, v^\theta_y)) \in M(P, Q)$.

The log-likelihood $l(\theta)$ associated to observation $\hat{\pi}_{xy}$ is

$$l(\theta) = \sum_{xy} \hat{\pi}_{xy} \log \pi^\theta_{xy}$$

$$= -\sum_{xy} \hat{\pi}_{xy} \Psi^\theta_{xy}(u^\theta_x, v^\theta_y)$$

and thus the maximum likelihood procedure consists in

$$\min_{\theta} \sum_{xy} \hat{\pi}_{xy} \Psi^\theta_{xy}(u^\theta_x, v^\theta_y).$$