An exchange economy problem with transport costs

X. Bacon, G. Carlier, B. Nazaret

Séminaire du SAMM

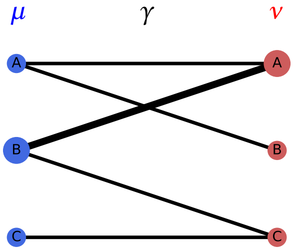
21 octobre 2022

Table of Contents

A digression through Optimal Transport

- $\mu, \nu \in \mathcal{M}_+(X)$ are respectively the source and the target distributions.
- $c = c(x, y)$ is the non-negative transport cost satisfying $c(x, x) = 0$ for all $x \in X$.

The Optimal Transport problem is defined as follows



$$\mathcal{T}_c(\mu, \nu) = \inf \left\{ \int_{X^2} c(x, y) \, \mathrm{d}\gamma(x, y) \text{ s.t. } \gamma \in \Pi(\mu, \nu) \right\}$$

where $\Pi(\mu, \nu)$ denotes the set of all $\gamma \in \mathcal{M}_+(X^2)$ having μ and ν as marginals, called **transference plans**.

Remark : if $\mu(X) \neq \nu(X)$, $\mathcal{T}_c(\mu, \nu) = +\infty$.

A digression through Optimal Transport

Duality : as a convex minimization problem, $\mathscr{T}_c(\mu,\nu)$ admits the following concave maximization formulation

$$\mathscr{T}_c(\mu,\nu) = \sup\left\{\int_X \psi(x)\,\mathrm{d}\mu(x) + \int_X \varphi(y)\,\mathrm{d}\nu(y) \text{ s.t. } \psi(x)+\varphi(y) \leqslant c(x,y)\right\}$$

where ψ,φ are two bounded and continuous functions.

A substitution : $\psi(x) \longleftarrow \boldsymbol{\varphi^c(x) \overset{\text{def.}}{=} \inf_y c(x,y) - \varphi(y)}$ improves the dual cost and satisfies the contraints. As a consequence the Optimal Transport problem can be rewritten as follows

$$\mathscr{T}_c(\mu,\nu) = \sup_{\varphi}\left\{\int_X \varphi^c(x)\,\mathrm{d}\mu(x) + \int_X \varphi(y)\,\mathrm{d}\nu(y)\right\}.$$

Primal problem

- X is a compact metric space.

Primal problem

$$(\mathscr{P}) = \sup \left\{ \qquad \qquad \qquad \qquad \qquad \qquad \right\}.$$

Primal problem

- X is a compact metric space.
- $\mu = (\mu_1, \ldots, \mu_N) \in \mathscr{M}_+(X)^N$ are the source distributions of goods in region X.

Primal problem

$$(\mathscr{P}) = \sup \left\{ \qquad \right\}.$$

Primal problem

- X is a compact metric space.
- $\mu = (\mu_1, \dots, \mu_N) \in \mathscr{M}_+(X)^N$ are the source distributions of goods in region X.
- $\nu = (\nu_1, \dots, \nu_N) \in \mathscr{M}_+(X)^N$ are the target distributions of goods in region X.

Primal problem

$$(\mathscr{P}) = \sup_{\nu} \left\{ \qquad\qquad \text{s.t. } \mu_i(X) = \nu_i(X) \right\}.$$

Primal problem

- X is a compact metric space.
- $\mu = (\mu_1, \dots, \mu_N) \in \mathscr{M}_+(X)^N$ are the source distributions of goods in region X.
- $\nu = (\nu_1, \dots, \nu_N) \in \mathscr{M}_+(X)^N$ are the target distributions of goods in region X.
- $\mathscr{U} = \mathscr{U}(\nu)$ is the average utility.

Primal problem

$$(\mathscr{P}) = \sup_{\nu} \left\{ \mathscr{U}(\nu) \qquad \text{s.t. } \mu_i(X) = \nu_i(X) \right\}.$$

Primal problem

- X is a compact metric space.
- $\mu = (\mu_1, \dots, \mu_N) \in \mathcal{M}_+(X)^N$ are the source distributions of goods in region X.
- $\nu = (\nu_1, \dots, \nu_N) \in \mathcal{M}_+(X)^N$ are the target distributions of goods in region X.
- $\mathcal{U} = \mathcal{U}(\nu)$ is the average utility.
- $\mathcal{T}(\mu, \nu) = \sum_{i=1}^{N} \mathcal{T}_{c_i}(\mu_i, \nu_i)$ is the transport cost between μ and ν.

Primal problem

$$(\mathcal{P}) = \sup_{\nu} \left\{ \mathcal{U}(\nu) - \mathcal{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X) \right\}.$$

Utility

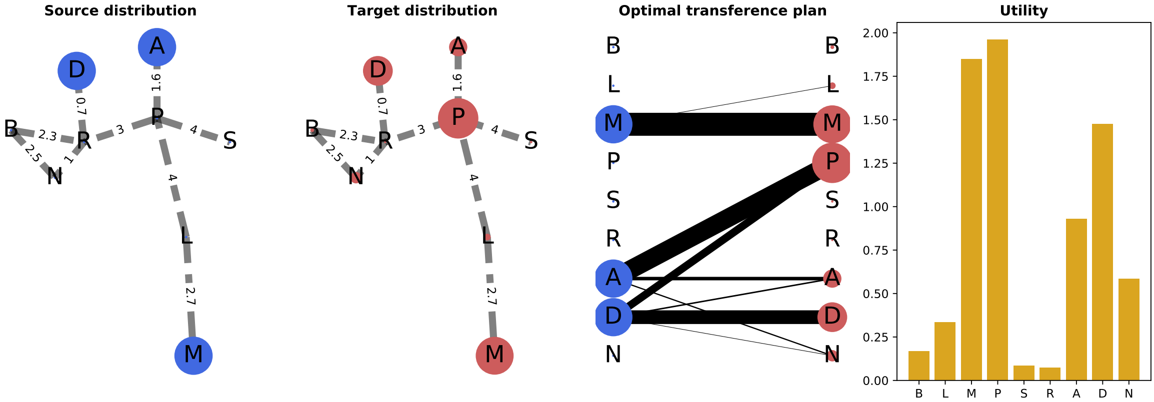
The average utility $\mathscr{U}$ is given by

$$\mathscr{U}(\nu) = \begin{cases} \int_X U(x, \beta_1(x), \ldots, \beta_N(x))\, \mathrm{d}m(x) \text{ if } \nu_i = \beta_i \cdot m \text{ for i = 1,...,N} \\ -\infty \textit{ otherwise} \end{cases}$$

where

- m is a reference measure *i.e.* $\mu_i \ll m$ for $i = 1, \ldots, N$.
- $U : (x, \beta) \in X \times \mathbb{R}_+^N \mapsto U(x, \beta) \in \mathbb{R} \cup \{-\infty\}$ *is the preference of the agent located in* x.

An example



Source distribution
A
D
1.6
0.7
B
2.3
R
3
P
4
S
2.5
1
N
4
L
2.7
M
Target distribution
A
D
1.6
0.7
B
2.3
R
3
P
4
S
2.5
1
N
4
L
2.7
M
Optimal transference plan
B
L
M
P
S
R
A
D
N
Utility
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
B L M P S R A D N

Existence of a minimizer

Technical assumptions on U and c

1. $\forall i = 1, \dots, N$, c_i is continuous, nonnegative and $\boldsymbol{c_i(x,x)} = 0$ for all $\boldsymbol{x} \in \boldsymbol{X}$.
2. for m-a.e. $x \in X$, U is upper semi-continuous, concave, **nondecreasing**.
3. for every $\beta \in \mathbb{R}_+^N$, $x \in X \mapsto U(x,\beta)$ is m-measurable.
4. $\beta \in L^1(X,m)^N \mapsto \int_X U(x,\beta(x))\,\mathrm{d}m(x)$ is not identically equals to $-\infty$.
5. $(x,\beta) \mapsto U(x,\beta)$ is sublinear with respect to β uniformly in $x \in X$, that is for any $\delta > 0$, it exists C_δ s.t. for m-a.e. $x \in X$,

$$U(x,\beta) \leqslant \delta \sum_{i=1}^{N} \beta_i + C_\delta.$$

Existence of a minimizer

$$(\mathscr{P}) = \sup_{\nu} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X) \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

If the assumptions above are satisfied, then the maximization problem $(\mathscr{P})$ admits at least one solution.

Existence of a minimizer

$$(\mathscr{P}) = \sup_{\nu} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X) \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

If the assumptions above are satisfied, then the maximization problem $(\mathscr{P})$ admits at least one solution.

Proof : $N = 1$, using the direct method in the calculus of variations, starting with a maximizing sequence $\nu^n = \beta^n \cdot m$, we extract a subsequence (not relabed) which admits the following convergences

$$\beta^n \xrightarrow{m-a.e.} \beta \quad \text{and} \quad \nu^n \xrightarrow{*} \nu$$

Existence of a minimizer

$$(\mathscr{P}) = \sup_{\nu} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X) \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

If the assumptions above are satisfied, then the maximization problem $(\mathscr{P})$ admits at least one solution.

Proof : $N = 1$, using the direct method in the calculus of variations, starting with a maximizing sequence $\nu^n = \beta^n \cdot m$, we extract a subsequence (not relabed) which admits the following convergences

$$\beta^n \xrightarrow{m-a.e.} \beta \quad \text{and} \quad \nu^n \stackrel{*}{\rightharpoonup} \nu$$

⚠ β may violate the mass constraint and ν may not belong to $L^1(m)$ ⚠

Existence of a minimizer

To overcome this difficulty, starting from ν, build a admissible $\tilde{\boldsymbol{\nu}}$ which increases $\mathscr{U}$ and decreases $\mathscr{T}(\mu,.)$. Let $\beta^a \in L^1$ and $\nu^s \in \mathscr{M}_+(X)$ such that $\nu^s \perp m$ and

$$\nu = \beta^a \cdot m + \nu^s \qquad \text{(Radon-Nikodym)}$$

Let $\gamma \in \Pi(\mu,\nu)$ be optimal and decompose it into

$$\gamma = \underbrace{\gamma_{|X\times A}}_{\gamma^a} + \underbrace{\gamma_{|X\times A^c}}_{\gamma^s}$$

where A satisfies $\nu^s(A) = m(A^c) = 0$. Set

$$\tilde{\boldsymbol{\gamma}} = \gamma^a + (\mathrm{Id},\mathrm{Id})\#[\mathrm{proj}_1\#\gamma^s]\cdot m.$$

Then the second marginal of $\tilde{\boldsymbol{\gamma}}$ solves $(\mathscr{P})$.

Table of Contents

Duality

Problem $(\mathscr{P})$ appears naturally as the dual of a convex minimization problem

$$(\mathscr{D}) = \inf\left\{\mathscr{K}(\varphi) + \mathscr{V}(\varphi) \text{ s.t. } \varphi \in C(X,\mathbb{R})^N\right\}$$

where

$$\mathscr{K}(\varphi) = -\sum_{i=1}^N \int_X \varphi_i^{c_i}\,\mathrm{d}\mu_i, \quad \text{and} \quad \mathscr{V}(\varphi) = \int_X V(x,\varphi_1(x),\ldots,\varphi_N(x))\,\mathrm{d}m(x)$$

with $V(x,\varphi) = \sup\{U(x,\beta) - \sum_{i=1}^N \beta_i\varphi_i : \beta \in \mathbb{R}_+^N\}$.

Theorem (Strong duality)

Under the assumptions above, the following equality is satisfied

$$(\mathscr{D}) = (\mathscr{P}).$$

Economic interpretation

Let (β,φ) be optimal in $(\mathscr{P})$ and $(\mathscr{D})$. Then we have an **equilibrium** for the initial monetary endowment $\boldsymbol{w} = \langle \boldsymbol{\varphi^c} \mid \boldsymbol{\alpha}\rangle + \langle \boldsymbol{\varphi} \mid \boldsymbol{\beta}\rangle (= \mathscr{T}_c(\alpha,\beta))$ in the sense that:

- **Sellers** at x maximize their profits by exporting their goods $\alpha(x)$:

$$\text{profits}_i(x) = \max_y \varphi_i(y) - c_i(x,y) \quad (= -\varphi_i^{c_i}(x))$$
$$\text{total profits}(x) = \langle \text{profits}(x) \mid \alpha(x)\rangle$$

- **Consumers** in y have an initial endowment $w(y)$ and buy $\beta_i(y)$ in such a way:

$$\beta_i(y) = \operatorname*{argmax}_\beta \left\{ \boldsymbol{U(y,\beta)} \text{ s.t. } \langle \varphi \mid \beta \rangle \leqslant \underbrace{\overbrace{\langle -\varphi^c(y) \mid \alpha(y)\rangle}^{\text{export profit}} + w(y)}_{\text{total revenue}} \right\}$$

- **Markets are clear** : it exists an optimal transport plan for all target endowments.

Table of Contents

Entropy regularization for Optimal Transport

X finite and $\alpha, \beta \in \mathbb{R}^X$ denotes respectively the source and the target distributions.

- **Entropic OT** : a popular and efficient tool in Computational OT since Cuturi's paper (2013) : for a fixed $\boldsymbol{\varepsilon} > 0$, the regularized OT problem is given by

$$\mathcal{T}^{\boldsymbol{\varepsilon}}(\alpha,\beta) = \inf \left\{ \langle c \mid \gamma \rangle + \boldsymbol{\varepsilon}\,\mathbf{Entropy}(\boldsymbol{\gamma}) \text{ s.t. } \gamma \in \Pi(\alpha,\beta) \right\}$$

  whose unique solution converges to the solution of $\mathcal{T}(\alpha,\beta)$ with maximal entropy.

- **Computation of regularized OT** : optimality conditions reads here

$$\gamma_{ij} = u_i K_{ij} v_j \qquad \text{(NSC)}$$

  where $u_i = \exp(\psi_i/\varepsilon)$, $v_j = \exp(\varphi_i/\varepsilon)$ and $K = \exp(-c_{ij}/\varepsilon)$. We use the following scheme to compute it:

$$u^{k+1} = \frac{\alpha}{K \cdot v^k} \text{ and } v^{k+1} = \frac{\beta}{K^T \cdot u^{k+1}} \qquad \text{(Sinkhorn's iterates)}$$

An entropic approximation algorithm

- m is the counting measure.

The algorithm below is based on a variant introduced by G. Peyré (2015) in the context of Wasserstein gradient flows.

An entropic approximation algorithm

- m is the counting measure.

The algorithm below is based on a variant introduced by G. Peyré (2015) in the context of Wasserstein gradient flows.

$$(\mathscr{D}) = \inf_{\varphi \in \mathbb{R}^{X \times N}} \sum_{y \in X} V(y, \varphi(y)) + \sum_{i=1}^{N} \sum_{x \in X} \alpha_i(x) \overbrace{\max_{y \in X} \{\varphi_i(y) - c_i(x, y)\}}^{= -\varphi^c(x)}.$$

An entropic approximation algorithm

- m is the counting measure.

The algorithm below is based on a variant introduced by G. Peyré (2015) in the context of Wasserstein gradient flows.

$$(\mathscr{D}) = \inf_{\varphi \in \mathbb{R}^{X \times N}} \sum_{y \in X} V(y, \varphi(y)) + \sum_{i=1}^{N} \sum_{x \in X} \alpha_i(x) \overbrace{\max_{y \in X} \{\varphi_i(y) - c_i(x,y)\}}^{= -\varphi^c(x)}.$$

- **From $(\mathscr{D})$ to $(\mathscr{D}_\varepsilon)$:** $\max_{y \in X} \{\varphi_i(y) - c_i(x,y)\} \leftarrow \underbrace{\boldsymbol{\varepsilon} \log \left(\sum_{y \in X} e^{\frac{\varphi_i(y) - c_i(x,y)}{\varepsilon}} \right)}_{\text{soft-max}}$.

$$(\mathscr{D}_\varepsilon) = \inf_{\varphi \in \mathbb{R}^{X \times N}} \sum_{y \in X} V(y, \varphi(y)) + \boldsymbol{\varepsilon} \sum_{i=1}^{N} \sum_{x \in X} \alpha_i(x) \log \left(\sum_{y \in X} e^{\frac{\varphi_i(y) - c_i(x,y)}{\varepsilon}} \right).$$

Optimality conditions

The optimality conditions for $(\mathscr{D}_\varepsilon)$ writes:

$$-\beta(y) \in \partial V(y, \varphi(y)), \forall y \in X$$

where β is given by

$$\beta_i(y) = \sum_{x \in X} \boldsymbol{\alpha_i(x)} \frac{\boldsymbol{e^{\frac{\varphi_i(y) - c_i(x,y)}{\varepsilon}}}}{\boldsymbol{\sum_{z \in X} e^{\frac{\varphi_i(z) - c_i(x,z)}{\varepsilon}}}}$$

Optimality conditions

The optimality conditions for $(\mathscr{D}_\varepsilon)$ writes:

$$-\beta(y) \in \partial V(y, \varphi(y)), \forall y \in X$$

where β is given by

$$\beta_i(y) = \sum_{x \in X} \boldsymbol{\alpha_i(x)} \frac{\boldsymbol{e^{\frac{\varphi_i(y) - c_i(x,y)}{\varepsilon}}}}{\boldsymbol{\sum_{z \in X} e^{\frac{\varphi_i(z) - c_i(x,z)}{\varepsilon}}}} \stackrel{\text{def.}}{=} \sum_{x \in X} \boldsymbol{\gamma_i(x,y)}$$

Optimality conditions

The optimality conditions for $(\mathscr{D}_\varepsilon)$ writes:

$$-\beta(y) \in \partial V(y, \varphi(y)), \forall y \in X$$

where β is given by

$$\beta_i(y) = \sum_{x \in X} \boldsymbol{\alpha_i(x)} \frac{\boldsymbol{e^{\frac{\varphi_i(y) - c_i(x,y)}{\varepsilon}}}}{\boldsymbol{\sum_{z \in X} e^{\frac{\varphi_i(z) - c_i(x,z)}{\varepsilon}}}} \stackrel{\text{def.}}{=} \sum_{x \in X} \boldsymbol{\gamma_i(x,y)}$$

- $\boldsymbol{\gamma_i}$ solves the Entropic OT problem $\mathscr{T}_{c_i}^{\varepsilon}(\alpha_i, \beta_i)$.
- β solves the dual problem of $(\mathscr{D}_\varepsilon)$:

$$(\mathscr{P}_\varepsilon) = \sup_{\beta \in \mathbb{R}_+^{X \times N}} \sum_{y \in X} U(y, \beta(y)) - \sum_{i=1}^{N} \mathscr{T}_{c_i}^{\varepsilon}(\alpha_i, \beta_i).$$

Table of Contents

Rewritting the dual

$(\mathscr{D}_\varepsilon)$ can be reformulated by considering the convex problem

$$(\widetilde{\mathscr{D}}_\varepsilon) = \inf_{\varphi,\psi} \ \Phi_\varepsilon(\varphi,\psi)$$

where $\Phi_\varepsilon(\varphi,\psi) = \sum_{y\in X} V(y,\varphi(y)) - \sum_{i=1}^{N}\sum_{x\in X} \alpha_i(x)\psi_i(x) + \varepsilon \sum_{i=1}^{N}\sum_{(x,y)\in X^2} e^{\frac{\psi_i(x)+\varphi_i(y)-c_i(x,y)}{\varepsilon}}.$

Proof : for fixed φ, the minimizer of $\psi \mapsto \Phi_\varepsilon(\varphi,\psi)$ is explicitly given by

$$\psi_i(x) = \varepsilon \log(\alpha_i(x)) - \varepsilon \log\left(\sum_{y\in X} e^{\frac{\varphi_i(y)-c_i(x,y)}{\varepsilon}}\right)$$

so replacing in Φ_ε, we get

$$\inf_{\psi} \Phi_\varepsilon(\varphi,\psi) = C + \sum_{y\in X} V(y,\varphi(y)) + \varepsilon \sum_{i=1}^{N}\sum_{x\in X} \alpha_i(x) \log\left(\sum_{y\in X} e^{\frac{\varphi_i(y)-c_i(x,y)}{\varepsilon}}\right).$$

Coordinate descent/Sinkhorn

$(\widetilde{\mathscr{D}}_\varepsilon)$ can be solved by coordinate descent: starting from (ψ^0, φ^0), updates are computed as follows:

$$\psi^{k+1} = \underset{\psi \in \mathbb{R}^{X \times N}}{\operatorname{argmin}} \Phi(\varphi^k, \psi) \quad \text{and} \quad \varphi^{k+1} = \underset{\varphi \in \mathbb{R}^{X \times N}}{\operatorname{argmin}} \Phi(\varphi, \psi^{k+1}).$$

- The first update is explicitly given by

$$\psi_i^{k+1}(x) = \varepsilon \log(\alpha_i(x)) - \varepsilon \log\Big(\sum_{y \in X} e^{\frac{\varphi_i^k(y) - c_i(x,y)}{\varepsilon}}\Big)$$

- The second update is (for fixed i and y) the same as solving a one-dimensional strictly convex minimization problem.

Remark : if V is smooth and locally strongly convex on its domain, this scheme convergences linearly (Beck, Tetruashvili, 2013).

The Cobb-Douglas case

If the utility is of the form

$$U(x,\beta(x)) = w(x)\prod_{i=1}^{N}\beta_i(x)^{a_i} \qquad \text{(Cobb-Douglas utility)}$$

where $a_i > 0$ and $a = \sum_{i=1}^{N} a_i < 1$, a direct computation of its Fenchel conjugate gives

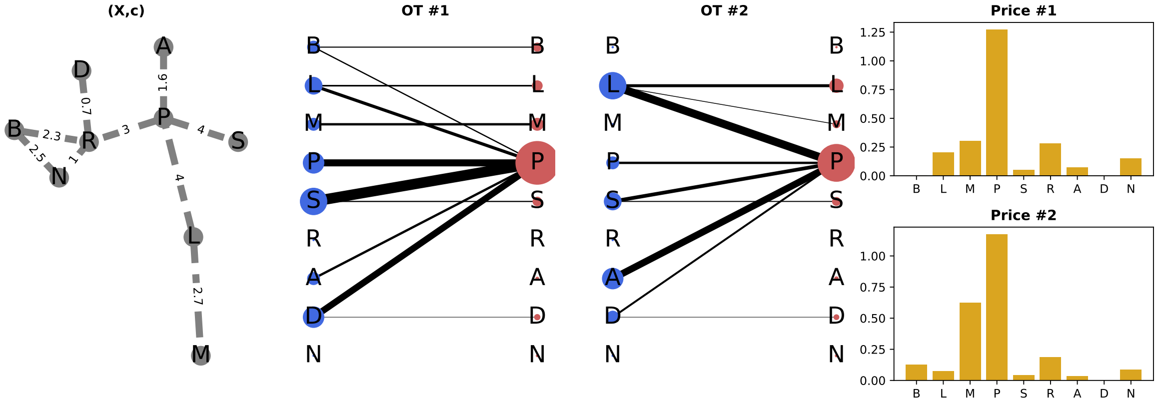
$$V(x,\varphi(x)) = w(x)^{\frac{1}{1-a}}\prod_{i=1}^{N}[a_i\varphi_i(x)]^{\frac{a_i}{a-1}}$$

and the second minimization step is reduced to find the root t of the strictly monotone equation (for some A and b)

$$e^t t^b = A$$

which can be solved using Newton's or dichotomy methods.

Simulation



(X,c)
A
D
B
R
P
S
N
L
M
1.9
0.7
2.3
3
4
2.5
1
4
2.7
OT #1
OT #2
B
L
M
P
S
R
A
D
N
Price #1
Price #2
1.25
1.00
0.75
0.50
0.25
0.00
B L M P S R A D N

Perspectives

- Extend this problem to other Transport-like models.
- Dynamic model?

Merci pour votre attention.