An exchange economy problem with transport costs

Xavier Bacon

Joint work with Guillaume Carlier and Bruno Nazaret

GT CalVa, Université Paris Dauphine-PSL

26 septembre 2022

1. Primal and dual problems
    - An existence result
    - Duality
    - Economic interpretation

2. Algorithm and simulations
    - Entropy regularization
    - Coordinate descent/Sinkhorn

Table of Contents

- X is a compact metric space.

Primal problem

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

Primal problem

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

Primal problem

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

Primal problem

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

Primal problem

$$(\mathscr{P}) = \sup_{\nu \in \mathscr{M}_+(X)^N} \Big\{ \qquad \text{s.t. } \mu_i(X) = \nu_i(X), i = 1, \ldots, N \Big\}$$

Primal problem

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

Primal problem

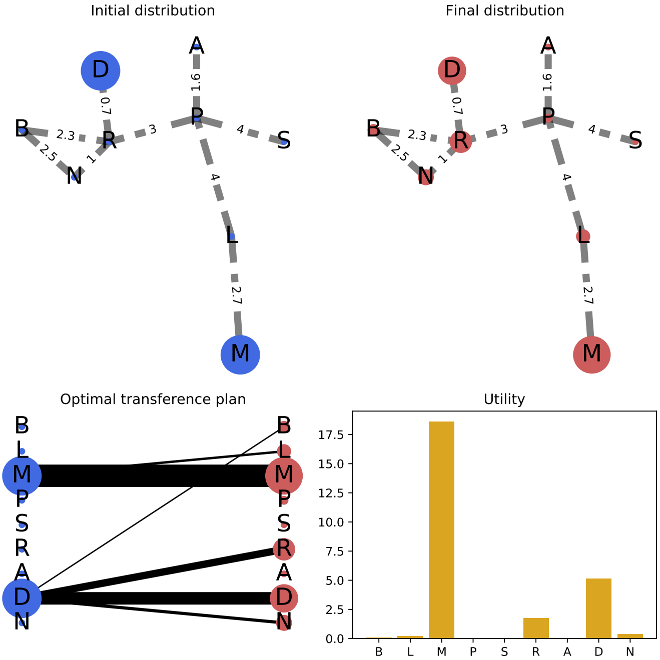
$$(\mathscr{P}) = \sup_{\nu \in \mathcal{M}_+(X)^N} \left\{ \mathcal{U}(\nu) \qquad \text{s.t. } \mu_i(X) = \nu_i(X), i = 1, \dots, N \right\}$$

Primal problem

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

Primal problem

$$(\mathcal{P}) = \sup_{\nu \in \mathcal{M}_+(X)^N} \left\{ \mathcal{U}(\nu) - \mathcal{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X), i = 1, \ldots, N \right\}$$


Initial distribution
Final distribution
A
D
P
B
R
S
N
L
M
1.6
0.7
2.3
3
4
2.5
1
4
2.7
Optimal transference plan
B
L
M
P
S
R
A
D
N
Utility
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
B L M P S R A D N

What is $T_c(\mu, \nu)$?

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

$$c(x,y) \geqslant 0 \text{ and } c(x,x) = 0.$$

What is $T_c(\mu,\nu)$?

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

$$c(x,y) \geqslant 0 \text{ and } c(x,x) = 0.$$

Transference plan and related cost

$\gamma \in \mathcal{M}_+(X \times Y)$ is a **transference plan between** μ **and** ν if

$$\int_Y \mathrm{d}\gamma(x,y) = \mu(x) \quad \text{and} \quad \int_X \mathrm{d}\gamma(x,y) = \nu(y)$$

Its related cost with respect to c is given by

$$\langle c \,|\, \gamma \rangle \overset{\text{def.}}{=} \iint_{X \times Y} c(x,y)\,\mathrm{d}\gamma(x,y) \in [0,\infty].$$

$\Pi(\mu, \nu)$ denotes the set of all transference plans between μ and ν.

$\Pi(\mu, \nu)$ denotes the set of all transference plans between μ and ν.

Optimal Transport problem

The optimal transport cost between μ and ν is given by

$$T_c(\mu, \nu) = \inf\left\{\langle c \mid \gamma \rangle : \gamma \in \Pi(\mu, \nu)\right\}.$$

Utility

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

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

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.

Technical assumptions on U and c

1. $\forall i = 1, \dots, N$, c_i is continuous, nonnegative and $c_i(x,x) = 0$ for all $x \in X$.
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 : $\forall \delta > 0, \exists C_\delta$ s.t. for m-a.e. $x \in X$

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

$$(\mathscr{P}) = \sup_{\nu \in \mathscr{M}_+(X)^N} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X), i = 1, \dots, N \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

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

$$(\mathscr{P}) = \sup_{\nu \in \mathscr{M}_+(X)^N} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X), i = 1, \dots, N \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

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

Proof : $N = 1$, let $\nu^n = \beta^n \cdot m$ be a maximizing sequence.

$$(\mathscr{P}) = \sup_{\nu \in \mathscr{M}_+(X)^N} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X), i = 1, \dots, N \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

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

Proof : $N = 1$, let $\nu^n = \beta^n \cdot m$ be a maximizing sequence.

- **Mass constraints :** $\forall n \in \mathbb{N}, \nu^n(X) = \int \beta^n \, \mathrm{d}m = \mu(X)$

$$(\mathscr{P}) = \sup_{\nu \in \mathscr{M}_+(X)^N} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu, \nu) \text{ s.t. } \mu_i(X) = \nu_i(X), i = 1, \dots, N \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

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

Proof : $N = 1$, let $\nu^n = \beta^n \cdot m$ be a maximizing sequence.

- **Mass constraints :** $\forall n \in \mathbb{N}, \nu^n(X) = \int \beta^n \, \mathrm{d}m = \mu(X)$

$$\Longrightarrow (\beta^n) \text{ is bounded in } L^1(m)$$

$$(\mathscr{P}) = \sup_{\nu \in \mathscr{M}_+(X)^N} \left\{ \mathscr{U}(\nu) - \mathscr{T}(\mu,\nu) \text{ s.t. } \mu_i(X) = \nu_i(X), i = 1,\dots,N \right\}$$

Proposition (B.,Carlier,Nazaret, 2021)

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

Proof : $N = 1$, let $\nu^n = \beta^n \cdot m$ be a maximizing sequence.

- **Mass constraints :** $\forall n \in \mathbb{N}, \nu^n(X) = \int \beta^n \,\mathrm{d}m = \mu(X)$

$$\Longrightarrow (\beta^n) \text{ is bounded in } L^1(m)$$

$$\overset{\text{Komlós}}{\Longrightarrow} \exists \beta \in L^1(m), \exists \text{ subseq. (not relabed) s.t.}$$

$$\frac{1}{n}\sum_{i=1}^{n} \beta^i \overset{m-a.e.}{\longrightarrow} \beta.$$

$\mathscr{U}$ and $-\mathscr{T}(\mu,\cdot)$ concave $\Rightarrow$ (β^n) is a maximizing sequence.

- Fatou's Lemma, sublinearity and concavity of U implies

$$\overline{\lim} \int U(x, \beta^n(x))\, \mathrm{d}m(x) \leqslant \int U(x, \beta(x))\, \mathrm{d}m(x).$$

- Fatou's Lemma, sublinearity and concavity of U implies

$$\overline{\lim} \int U(x, \beta^n(x))\, \mathrm{d}m(x) \leqslant \int U(x, \beta(x))\, \mathrm{d}m(x).$$

- $\nu^n = \beta^n \cdot m$ bounded in $\mathscr{M}_+(X) \Rightarrow$ (not relabed) $\exists \nu \in \mathscr{M}_+(X)$ s.t.

$$\nu^n \overset{*}{\rightharpoonup} \nu$$

- Fatou's Lemma, sublinearity and concavity of U implies

$$\overline{\lim} \int U(x, \beta^n(x)) \, \mathrm{d}m(x) \leqslant \int U(x, \beta(x)) \, \mathrm{d}m(x).$$

- $\nu^n = \beta^n \cdot m$ bounded in $\mathscr{M}_+(X) \Rightarrow$ (not relabed) $\exists \nu \in \mathscr{M}_+(X)$ s.t.

$$\nu^n \xrightarrow{*} \nu$$

since $\mathscr{T}(\mu, .)$ is seq. weakly-$*$ l.s.c. then

$$\overline{\lim} -\mathscr{T}(\mu, \nu^n) \leqslant -\mathscr{T}(\mu, \nu)$$

$$\Rightarrow (\mathscr{P}) = \int U(x, \beta) \, \mathrm{d}m(x) - \mathscr{T}(\mu, \nu)$$

- Fatou's Lemma, sublinearity and concavity of U implies

$$\overline{\lim} \int U(x, \beta^n(x))\, \mathrm{d}m(x) \leqslant \int U(x, \beta(x))\, \mathrm{d}m(x).$$

- $\nu^n = \beta^n \cdot m$ bounded in $\mathscr{M}_+(X) \Rightarrow$ (not relabed) $\exists \nu \in \mathscr{M}_+(X)$ s.t.

$$\nu^n \overset{*}{\rightharpoonup} \nu$$

since $\mathscr{T}(\mu, .)$ is seq. weakly-$*$ l.s.c. then

$$\overline{\lim} -\mathscr{T}(\mu, \nu^n) \leqslant -\mathscr{T}(\mu, \nu)$$
$$\Rightarrow (\mathscr{P}) = \int U(x, \beta)\, \mathrm{d}m(x) - \mathscr{T}(\mu, \nu)$$

⚠ With the first convergence (m-a.e.), β may violate the mass constraint and with the second (weakly) convergence ν may not belong to $L^1(m)$.

- Weak-$*$ convergence $\Rightarrow \nu \geq \beta \cdot m$

- Weak-$*$ convergence $\Rightarrow \nu \geq \beta \cdot m$
- **Radon-Nikodym Theorem** : let $\beta^a \in L^1$ and $\nu^s \in \mathscr{M}_+(X)$ be s.t.

$$\nu = \beta^a \cdot m + \nu^s$$

with $\nu^s \perp m$

- Weak-$*$ convergence $\Rightarrow \nu \geq \beta \cdot m$
- **Radon-Nikodym Theorem** : let $\beta^a \in L^1$ and $\nu^s \in \mathcal{M}_+(X)$ be s.t.

$$\nu = \beta^a \cdot m + \nu^s$$

with $\nu^s \perp m$ ($\Longrightarrow \exists A$ measurable s.t. $\nu^s(A) = m(A^c) = 0$).

- Weak-$*$ convergence $\Rightarrow \nu \geq \beta \cdot m$
- **Radon-Nikodym Theorem** : let $\beta^a \in L^1$ and $\nu^s \in \mathcal{M}_+(X)$ be s.t.

$$\nu = \beta^a \cdot m + \nu^s$$

with $\nu^s \perp m$ ($\Longrightarrow \exists A$ measurable s.t. $\nu^s(A) = m(A^c) = 0$).

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

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

- Weak-$*$ convergence $\Rightarrow \nu \geq \beta \cdot m$
- **Radon-Nikodym Theorem** : let $\beta^a \in L^1$ and $\nu^s \in \mathcal{M}_+(X)$ be s.t.

$$\nu = \beta^a \cdot m + \nu^s$$

with $\nu^s \perp m$ ($\Longrightarrow \exists A$ measurable s.t. $\nu^s(A) = m(A^c) = 0$).

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

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

Set

$$\widetilde{\gamma} = \gamma^a + (\mathrm{Id}, \mathrm{Id}) \# \alpha^s \cdot m$$

where $\alpha^s = \mathrm{proj} \# \gamma^s$.

$$\text{then } \operatorname{proj}_1 \# \widetilde{\gamma} = \mu \text{ and } \operatorname{proj}_2 \# \widetilde{\gamma} = \underbrace{(\beta^a + \alpha^s)}_{\widetilde{\beta}} \cdot m$$

$$\text{then } \operatorname{proj}_1 \# \widetilde{\gamma} = \mu \text{ and } \operatorname{proj}_2 \# \widetilde{\gamma} = \underbrace{(\beta^a + \alpha^s)}_{\widetilde{\beta}} \cdot m$$

$$\mathscr{T}(\mu, \widetilde{\beta} \cdot m) \leqslant \int c \, \mathrm{d}\widetilde{\gamma} = \int c \, \mathrm{d}\gamma^a \text{ since } c(x,x) = 0$$

then $\mathrm{proj}_1 \# \tilde{\gamma} = \mu$ and $\mathrm{proj}_2 \# \tilde{\gamma} = \underbrace{(\beta^a + \alpha^s)}_{\tilde{\beta}} \cdot m$

$$\begin{aligned} \mathscr{T}(\mu, \tilde{\beta} \cdot m) &\leqslant \int c \, \mathrm{d}\tilde{\gamma} = \int c \, \mathrm{d}\gamma^a \text{ since } c(x,x) = 0 \\ &\leqslant \int c \, \mathrm{d}(\underbrace{\gamma^a + \gamma^s}_{\gamma}) = \mathscr{T}(\mu, \nu) \end{aligned}$$

then $\text{proj}_1 \# \widetilde{\gamma} = \mu$ and $\text{proj}_2 \# \widetilde{\gamma} = \underbrace{(\beta^a + \alpha^s)}_{\widetilde{\beta}} \cdot m$

$$\begin{aligned} \mathcal{T}(\mu, \widetilde{\beta} \cdot m) &\leqslant \int c \, \mathrm{d}\widetilde{\gamma} = \int c \, \mathrm{d}\gamma^a \text{ since } c(x,x) = 0 \\ &\leqslant \int c \, \mathrm{d}(\underbrace{\gamma^a + \gamma^s}_{\gamma}) = \mathcal{T}(\mu, \nu) \end{aligned}$$

$$\Rightarrow \mathcal{T}(\mu, \widetilde{\beta} \cdot m) \leqslant \mathcal{T}(\mu, \nu)$$

$$\text{then } \mathrm{proj}_1 \# \tilde{\gamma} = \mu \text{ and } \mathrm{proj}_2 \# \tilde{\gamma} = \underbrace{(\beta^a + \alpha^s)}_{\tilde{\beta}} \cdot m$$

$$\begin{aligned}\mathscr{T}(\mu, \tilde{\beta} \cdot m) &\leqslant \int c \, \mathrm{d}\tilde{\gamma} = \int c \, \mathrm{d}\gamma^a \text{ since } c(x,x) = 0 \\ &\leqslant \int c \, \mathrm{d}(\underbrace{\gamma^a + \gamma^s}_{\gamma}) = \mathscr{T}(\mu, \nu)\end{aligned}$$

$$\Rightarrow \mathscr{T}(\mu, \tilde{\beta} \cdot m) \leqslant \mathscr{T}(\mu, \nu)$$

Moreover, $\tilde{\beta} \geqslant \beta^a \geqslant \beta$ then $\mathscr{U}(\tilde{\beta} \cdot m) \geqslant \mathscr{U}(\beta \cdot m)$ by monotonicity of $\mathscr{U}$.
Then $\tilde{\beta}$ solves $(\mathscr{P})$.

Table of Contents

Duality

Problem $(\mathscr{P})$ appears naturally as the dual of a convex minimization problem over continuous functions $\varphi \in C(X, \mathbb{R})^N$:

Duality

Problem $(\mathscr{P})$ appears naturally as the dual of a convex minimization problem over continuous functions $\varphi \in C(X,\mathbb{R})^N$:

$$(\mathscr{P}) = \sup_{\nu} \mathscr{U}(\nu) - \sum_{i=1}^{N} T_{c_i}(\mu_i, \nu_i)$$

Duality

Problem $(\mathscr{P})$ appears naturally as the dual of a convex minimization problem over continuous functions $\varphi \in C(X,\mathbb{R})^N$:

$$
\begin{aligned}
(\mathscr{P}) &= \sup_{\nu} \mathscr{U}(\nu) - \sum_{i=1}^{N} T_{c_i}(\mu_i, \nu_i) \\
&= \sup_{\nu} \mathscr{U}(\nu) - \sup_{\varphi} \left(\sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i + \int_X \varphi_i \, \mathrm{d}\nu_i \right)
\end{aligned}
$$

Duality

Problem $(\mathscr{P})$ appears naturally as the dual of a convex minimization problem over continuous functions $\varphi \in C(X,\mathbb{R})^N$:

$$
\begin{aligned}
(\mathscr{P}) &= \sup_{\nu} \mathscr{U}(\nu) - \sum_{i=1}^{N} T_{c_i}(\mu_i, \nu_i) \\
&= \sup_{\nu} \mathscr{U}(\nu) - \sup_{\varphi} \left(\sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i + \int_X \varphi_i \, \mathrm{d}\nu_i \right) \\
&= \sup_{\nu} \inf_{\varphi} \mathscr{U}(\nu) - \sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i - \sum_{i=1}^{N} \int_X \varphi_i \, \mathrm{d}\nu_i
\end{aligned}
$$

Duality

Problem $(\mathscr{P})$ appears naturally as the dual of a convex minimization problem over continuous functions $\varphi \in C(X,\mathbb{R})^N$:

$$
\begin{aligned}
(\mathscr{P}) &= \sup_{\nu} \mathscr{U}(\nu) - \sum_{i=1}^{N} T_{c_i}(\mu_i,\nu_i) \\
&= \sup_{\nu} \mathscr{U}(\nu) - \sup_{\varphi} \left(\sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i + \int_X \varphi_i \, \mathrm{d}\nu_i \right) \\
&= \sup_{\nu} \inf_{\varphi} \mathscr{U}(\nu) - \sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i - \sum_{i=1}^{N} \int_X \varphi_i \, \mathrm{d}\nu_i
\end{aligned}
$$

$$
\text{(no duality gap)} = \inf_{\varphi} \left(\underbrace{- \sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i}_{:=\mathscr{K}(\varphi)} + \underbrace{\sup_{\nu} \left[\mathscr{U}(\nu) - \sum_{i=1}^{N} \int_X \varphi_i \, \mathrm{d}\nu_i \right]}_{:=\mathscr{V}(\varphi)} \right).
$$

Calculations above suggest setting

$$\mathscr{K}(\varphi) = -\sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i,$$

$$\mathscr{V}(\varphi) = \int_X V(x, \varphi_1(x), \ldots, \varphi_N(x)) \, \mathrm{d}m(x)$$

where $V(x, \varphi) := \sup_{\beta \in \mathbb{R}_+^N} U(x, \beta) - \sum_{i=1}^N \beta_i \varphi_i$ and setting the following *dual* problem :

$$(\mathscr{D}) = \inf_{\varphi \in C(X, \mathbb{R})^N} \mathscr{K}(\varphi) + \mathscr{V}(\varphi).$$

Calculations above suggest setting

$$\mathcal{K}(\varphi) = -\sum_{i=1}^{N} \int_X \varphi_i^{c_i} \, \mathrm{d}\mu_i,$$

$$\mathcal{V}(\varphi) = \int_X V(x, \varphi_1(x), \ldots, \varphi_N(x)) \, \mathrm{d}m(x)$$

where $V(x,\varphi) := \sup_{\beta \in \mathbb{R}_+^N} U(x,\beta) - \sum_{i=1}^N \beta_i \varphi_i$ and setting the following *dual* problem :

$$(\mathscr{D}) = \inf_{\varphi \in C(X,\mathbb{R})^N} \mathcal{K}(\varphi) + \mathcal{V}(\varphi).$$

Theorem (Strong duality)

Under the assumptions above, the following equality is satisfied

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

Table of Contents

An **equilibrium** for an initial monetary endowment w is a system of prices φ and a final endowment of the goods β satisfying:

An **equilibrium** for an initial monetary endowment w is a system of prices φ and a final endowment of the goods β satisfying:

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

$$\text{profits}_i(x) = \max \varphi_i(y) - c_i(x,y) \quad (= -\varphi_i^{c_i}(x))$$

$$\text{total profits}(x) = \text{profits}(x) \cdot \alpha(x)$$

An **equilibrium** for an initial monetary endowment w is a system of prices φ and a final endowment of the goods β satisfying:

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

$$\text{profits}_i(x) = \max \varphi_i(y) - c_i(x,y) \quad (= -\varphi_i^{c_i}(x))$$
$$\text{total profits}(x) = \text{profits}(x) \cdot \alpha(x)$$

- For all y, consumers in y have an initial endowment $w(y)$ and buy $\beta_i(y)$ in order to maximize its utility under budget constraint:

$$\beta_i(y) = \underset{\beta}{\text{argmax}} \left\{ U(y,\beta) : \varphi \cdot \beta \leqslant \underbrace{\overbrace{-\varphi^c(y) \cdot \alpha(y)}^{\text{export profit}} + w(y)}_{\text{total revenue}} \right\}$$

An **equilibrium** for an initial monetary endowment w is a system of prices φ and a final endowment of the goods β satisfying:

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

$$\text{profits}_i(x) = \max \varphi_i(y) - c_i(x,y) \quad (= -\varphi_i^{c_i}(x))$$
$$\text{total profits}(x) = \text{profits}(x) \cdot \alpha(x)$$

- For all y, consumers in y have an initial endowment $w(y)$ and buy $\beta_i(y)$ in order to maximize its utility under budget constraint:

$$\beta_i(y) = \underset{\beta}{\text{argmax}} \left\{ U(y,\beta) : \varphi \cdot \beta \leqslant \underbrace{\overbrace{-\varphi^c(y) \cdot \alpha(y)}^{\text{export profit}} + w(y)}_{\text{total revenue}} \right\}$$

- Markets are clear : one can find a plan $\gamma_i \in \Pi(\alpha_i m, \beta_i m)$ such that for every $(x,y) \in \text{supp}(\gamma_i)$ we have

$$\underbrace{-\varphi_i^{c_i}(x)}_{\text{profits of the seller}} = \underbrace{\varphi_i(y)}_{\text{price}} - \underbrace{c_i(x,y)}_{\text{transport cost}}$$

Theorem

Let β and φ solves $(\mathscr{P})$ and $(\mathscr{D})$ respectively and define

$$w = \varphi \cdot \beta + \varphi^c \cdot \alpha$$

then (β, φ) is an equilibrium with monetary endowment w.

Table of Contents

- Entropic optimal transport and Sinkhorn algorithm : popular and efficient tool in computational optimal transport since Cuturi's paper (2013).
- The algorithm below is based on a variant introduced by G. Peyré (2015) in the context of Wasserstein gradient flows.

An entropic approximation algorithm

- X finite, m is the counting measure.
- $\alpha_i(x) > 0$ denotes the initial endowment of location $x \in X$ in the good $i \in \{1, \ldots, N\}$.

An entropic approximation algorithm

- X finite, m is the counting measure.
- $\alpha_i(x) > 0$ denotes the initial endowment of location $x \in X$ in the good $i \in \{1, \ldots, N\}$.

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

- X finite, m is the counting measure.
- $\alpha_i(x) > 0$ denotes the initial endowment of location $x \in X$ in the good $i \in \{1, \dots, N\}$.

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

Let $\varepsilon > 0$ be a regularization parameter,

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

An entropic approximation algorithm

- X finite, m is the counting measure.
- $\alpha_i(x) > 0$ denotes the initial endowment of location $x \in X$ in the good $i \in \{1, \dots, N\}$.

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

Let $\varepsilon > 0$ be a regularization parameter,

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

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

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

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

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

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

Introducing

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

we have $\gamma_i \in \Pi(\alpha_i, \beta_i)$ and γ_i solves the entropic optimal transport problem

$$T^\varepsilon_{c_i}(\alpha_i, \beta_i) := \inf_{\gamma \in \Pi(\alpha_i, \beta_i)} \sum_{(x,y) \in X^2} (c_i(x,y) + \varepsilon \log(\gamma(x,y))\gamma(x,y).$$

In fact, such a β solves the dual problem of $(\mathscr{D}_\varepsilon)$:

$$\sup_{\beta \in \mathbb{R}_+^{X \times N}} \sum_{y \in X} U(y, \beta(y)) - \sum_{i=1}^{N} T_{c_i}^{\varepsilon}(\alpha_i, \beta_i), \tag{2.1}$$

which is the entropic regularization of $(\mathscr{P})$ with regularization parameter ε.

Table of Contents

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

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

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

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

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

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

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

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

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

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

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

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

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

so replacing in Φ_ε, we get

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

Coordinate descent/Sinkhorn

To solve $(\mathscr{D}_\varepsilon)$ it is sufficient to solve $(\widetilde{\mathscr{D}}_\varepsilon)$ that is

$$\inf_{\varphi,\psi} \Phi_\varepsilon(\varphi,\psi)$$

which can be solved by coordinate descent: starting from (ψ^0,φ^0), updates are computed as follows:

$$\psi^{k+1} = \text{argmin}_{\psi\in\mathbb{R}^{X\times N}} \Phi(\varphi^k,\psi)$$
$$\varphi^{k+1} = \text{argmin}_{\varphi\in\mathbb{R}^{X\times N}} \Phi(\varphi,\psi^{k+1})$$

Notice that one of the update is explicity 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 i and y fixed) the same as solving a one-dimensional strictly convex minimization problem.

Notice that one of the update is explicity 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 i and y fixed) the same as solving a one-dimensional strictly convex minimization problem.
Remark : if V is smooth and locally strongly convex ont its domain, this scheme convergences linearly (Beck, Tetruashvili, 2013).

The Cobb-Douglas case : if the utility is of the form

$$U(\beta) = \prod_{i=1}^{N} \beta_i^{a_i}, \quad a_i > 0, \quad a = \sum_{i=1}^{N} a_i < 1$$

then $V(\varphi) = (1-a) \prod_{i=1}^{N} a_i^{\frac{a_i}{1-a}} \varphi_i^{\frac{a_i}{a-1}}$

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 a dichotomy method.

Merci pour votre attention.