Théorie du transport optimal

Racine Florian

December 7, 2022

Table des matières

Chapter 1

Cours

1.1 Introduction

1.1.1 Formulation du problème

Question 1

Quelle est la façon optimal de transporter un tas de sable dans un trou ?

Question 2

Comment constuire un chateau de sable d'une forme données à partir d'un tas de sable ?

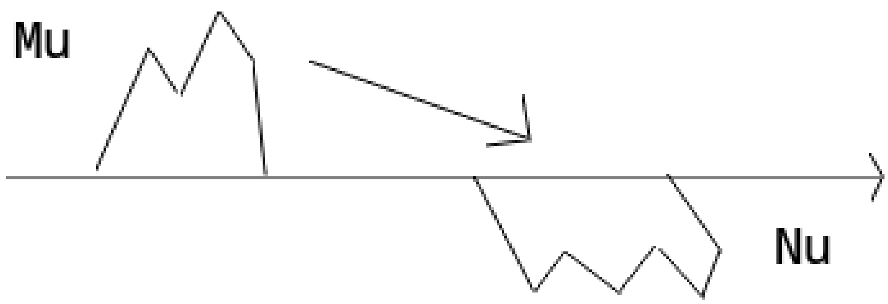

Figure 1.1: Transporter un tas de sable dans un trou

Note:-
Avec le même nombre de grain de sable et la même masse.

1.2 Modélisation

$\nu \in \mathcal{P}(\mathbb{R})$; $\mu \in \mathcal{P}(\mathbb{R})$

Definition 1.2.1

$\forall A \in \mathcal{P}(\mathbb{R}), \mu[A]$ décrit quelle quantité de sable est dans A.

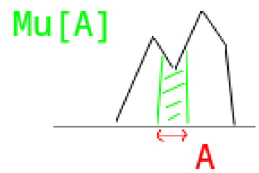

Figure 1.2: $\mu[A]$

Definition 1.2.2

Cout infinitésimal :

$$C : \left| \begin{array}{ccc} \mathbb{R} * \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & C(x, y) \end{array} \right.$$

Cout de transporter un grain de sable de x vers y.

Question 3

Comment transporter un tas de sable avec un cout global minimal ?

Definition 1.2.3

Un plan de transport entre les mesures μ et ν est une mesure de probabilité :
$\Pi \in \mathcal{P}(\mathbb{R} * \mathbb{R})$ à pour marginale μ et ν.

Note:-

$\Pi \in \mathcal{P}(\mathbb{R} * \mathbb{R})$ à pour marginal μ et ν

$\Leftrightarrow$ $\forall$A,B enssemble mesurable avec A$\subset$ $\mathbb{R}$ et B $\subset \mathbb{R}$ $\begin{cases} \Pi[A \times \mathbb{R}] = \mu[A] \\ \Pi[\mathbb{R} \times B] = \mu[B] \end{cases}$

$\Leftrightarrow \forall \varphi \in C^0(\mathbb{R}), \Psi \in C^0(\mathbb{R}) : \displaystyle\int_{\mathbb{R}\times\mathbb{R}} \varphi(x) + \Psi(y)\, d\Pi(x, y) = \int \varphi(x)\, d\mu(x)$

Note:-

On notera, $\Pi(\mu, \nu) = \{ \ \Pi \in \mathcal{P}(\mathbb{R} \times \mathbb{R}) | \Pi$ a pour marginal, $\mu, \nu \ \}$
On remarquera que, $\Pi(\mu, \nu) \neq \emptyset$
$I[\Pi] = \displaystyle\int_{\mathbb{R}^2} C(x, y)\, d\Pi(x, y)$ Le cout total assocé au plan de transport optimal.
On cherche, $\tau_c(\mu, \nu) = INF_{\Pi \in \Pi(\mu,\nu)}(I[\Pi])$

Definition 1.2.4

S'il existe $\Pi_0 \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ tel que $I[\Pi_0] = \tau_c(\mu, \nu)$
Π_0 est appelé un **plan de transfert optimal**

Exemple 1.2.1 (Exemple trivial (Kotorovitch))

$a < b$
$c < d$
$C(x, y) = |x - y|^2$
$\mu = \frac{1}{2}(\delta_a + \delta_b)$
$\nu = \frac{1}{2}(\delta_c + \delta_d)$

Question 4

$\Pi(\mu, \nu) = ?$

Solution: $\Pi_\alpha = \frac{1}{2}(\alpha\delta_{(a,c)} + (1-\alpha)\delta_{(a,d)} + (1-\alpha)\delta_{(b,c)} + \alpha\delta_{(b,d)})$ $\boldsymbol{\Pi(\mu, \nu) = \{\Pi_\alpha | \alpha \in [0, 1]\}}$

Question 5

Calculer : $I[\Pi] \forall \Pi \in \Pi(\mu, \nu)$

Solution: $I[\Pi_\alpha] = \int_{\mathbb{R}^2} C(x, y)\, d\Pi_\alpha(x, y)$
$I[\Pi_\alpha] = \frac{1}{2}(\alpha C(a, c) + (1-\alpha)C(a, d) + (1-\alpha)C(b, c) + \alpha C(b, d))$
$I[\Pi_\alpha] = \frac{1}{2}(a^2 + b^2 + c^2 + d^2) - \alpha(ac + bd) - (1-\alpha)(ad + cb)$

Question 6

Trouver : $\tau_c(\mu, \nu)$

Solution: $P(\alpha) = \frac{\partial I[\Pi_\alpha]}{\partial \alpha} \implies P(\alpha) = -ac - bd + ad + cb; \implies P(\alpha) = (d - c)(a - b) < 0$
Donc, $I[\Pi_\alpha]$ atteint son min en $\alpha = 1$
$\Pi_0 = \Pi_{\alpha=1}$
Donc, $a \longrightarrow c$ et, $b \longrightarrow d$

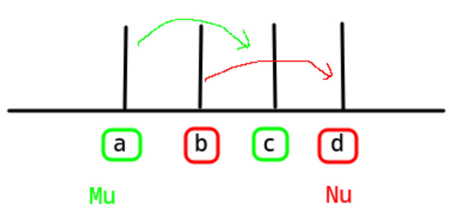

Figure 1.3: Solution

1.3 La formulation du problème de transfert optimal de Monge

Note:-
On autorise pas le fait de couper les masses. A chaque x est associé une unique y.
On dit que T envoie μ sur ν et on note : $T\#\mu = \nu$

Proposition 1.3.1
$\forall A \subset \mathbb{R}$ partie mesurable : $\nu(A) = \mu(T^{-1}(A))$

$\Leftrightarrow$

Proposition 1.3.2

$\forall \varphi$ continue : $\int_{\mathbb{R}} \varphi(y)\, d\nu(y) = \int_{\mathbb{R}} (\varphi o T)(x)\, d\mu(x)$

$\tau_c^M(\mu,\nu) = INF_{TtqT\#f=\nu} I[T]$

$I(T) = \int_{\mathbb{R}} C(x,T(x))\, d\mu(x)$

Note:-

Solution de cout optimal d'après Kantorovitch $\leq$ Solution de cout optimal d'après Monge
Dans le première exemple ils coincident.

Note:-

Kantorovitch définit un problème linéare en Π.
Monge définit un problème non linéare en T.

Note:-

Problème de kantorovitch admet toujours une solution Π_0.
Problème de Monge n'admet pas toujours de solution n'y même d'application qui envoi μ sur ν.

Exemple 1.3.1

$\begin{cases} \mu \in \mathcal{P}(\mathbb{R}) \\ \nu = \delta_a \end{cases}$

Kantorovitch : $\Pi(\mu,\nu) = \{\mu \otimes \delta_a\}$
Monge : Quelles sont les T tel que $T\#\mu = \nu$?
Il en existe une seule:

$$\forall x | T : \left| \begin{array}{ccc} x & \longrightarrow & a \\ \mathbb{R} & \longmapsto & \mathbb{R} \end{array} \right.$$

$\tau_c^M(\mu,\nu) = \tau_c(\mu,\nu)$

D'une part :

$\tau_c^M(\mu,\nu) = \int_{\mathbb{R}} C(T(x),x)\, d\mu(x)$

$\tau_c^M(\mu,\nu) = \int_{\mathbb{R}} C(0,x)\, d\mu(x)$

D'autre part :

$\tau_c(\mu,\nu) = \int_{\mathbb{R}} C(x,y)\, d\Pi(x,y)$

$\tau_c(\mu,\nu) = \int_{\mathbb{R}} C(x,y)\, d(\mu \otimes \delta_a)(x,y)$

$\tau_c(\mu,\nu) = \int_{\mathbb{R}} C(x,y)\, d\mu(x) d\delta_a$

$\tau_c(\mu,\nu) = \int_{\mathbb{R}} C(x,0)\, d\mu(x)$

Exemple 1.3.2

$\begin{cases} \mu = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i} \\ \nu = \frac{1}{n}\sum_{i=1}^{n} \delta_{y_i} \end{cases}$

Les plans de transports Π entre μ et ν peuvent être représenté par des matrices bistochastiques de tailles n.

$0 \leq \Pi_{i,j} \leq 1$

$\sum_{i=1}^{n} \Pi_{i,j} = 1$

$\sum_{j=1}^{n} \Pi_{i,j} = 1$

Note:-

On note $\mathcal{B}_n$ l'enssemble des matrices bisctochastiques.

Soit $\Pi \in \mathcal{B}_n$: $I[\Pi] = \frac{1}{n}\sum_{i=1}^{n} C(x_i, y_i)\Pi_{i,j}$
$\tau_c(\mu,\nu) = INF_{\Pi\in\mathcal{B}_n}\{\frac{1}{n}\sum_{i=1}^{n} \Pi_{i,j}C(x_i, y_i)$
Il s'agit d'un problème linéaire de minimisation sur un enssemble convexe.

Proposition 1.3.3 Enssemble convexe

$\mathcal{B}_n$ est convexe $\Leftrightarrow A, B \in \mathcal{B}_n$ alors $\forall\theta \in [0,1]|\theta A + (1-\theta)B \in \mathcal{B}_n$

Definition 1.3.1: Points extremaux

L'enssemble des points extremaux de E convexe est l'enssemble des $e \in E$ tel que :
si $e = \theta e_1 + (1-\theta)e_2$ avec $\theta \in [0,1], e_1 \in E, e_2 \in E$
Alors $\theta = 0$ ou $\theta = 1$

Théoreme 1.3.1 Théorème de Choquet

F est linéaire sur un domaine K convexe et compact, alors F admet au moin un minimum. Parmi les minimums de F au moin l'un d'eux est un extrema de K.

Théoreme 1.3.2 Théorème de Birkhoff

$\mathcal{B}_n$ est convexe et compact.
$\mathcal{B}_n$ admet n points extremaux qui sont les matrices de permutations
Ainsi, le min pour le problème de Kantorovitch est atteint pour $\begin{cases} \Pi_{i,j} = 1 | sij = \sigma(i) \\ \Pi_{i,j} = 0 | sinon \end{cases}$

1.4 La dualité de Kantorovitch

1.4.1 La théorie

Théoreme 1.4.1 Dualité de Kantorovitch

$\mu \in \mathcal{P}(\mathbb{R}^n)$; $\nu \in \mathcal{P}(\mathbb{R}^n)$
C semi continue inférieurement (par exemple C continue)

Note:-

Pour $\Pi \in \Pi(\mu,\nu), I[\Pi] = \displaystyle\int C(x,y)\, d\Pi(x,y)$

Soit $(\varphi,\psi) \in \phi_c$
$J(\varphi,\psi) = \displaystyle\int_{\mathbb{R}^n} \varphi(x)\, d\mu(x) + \int_{\mathbb{R}^n} \psi(y)\, d\nu(y)$
$\phi_c = \{(\varphi,\psi) \in (C^0(\mathbb{R}^n))^2$ tq $\varphi(x) + \psi(y) \le C(x,y)$ presque partout $\}$
Alors, $INF_{\Pi\in\Pi(\mu,\nu)} I[\Pi] = SUP_{(\varphi,\psi)\in\phi_c} J(\varphi,\psi)$

Note:-

Interprétation :

1. On embauche un transporteur.
2. Il achète de la masse située en x au pris $\varphi(x)$.
3. Il vous débarasse au prix $\int \varphi(x)\,d\mu(x)$
4. Il vous revend de la masse en y au prix $\psi(y)$
5. On rachète ν au prix $\int \psi(y)\,d\nu(y)$

On embauche le transporteur sous la condition : $\varphi(x)+\psi(y)\leqslant C(x,y)$

Proof: D'une part :
Soit $(\varphi,\psi)\in\phi_c$
Soit $\Pi\in\Pi(\mu,\nu)|I[\Pi])\tau(\mu,\nu)$

$$J(\varphi,\psi)=\int_{\mathbb{R}^n}\varphi(x)\,d\mu(x)+\int_{\mathbb{R}^n}\psi(y)\,d\nu(y)\le\int C(x,y)\,d\Pi(x,y)=I[\Pi]=\tau(\mu,\nu)=INF_{\Pi}I[\Pi]$$

Ainsi, $J(\varphi,\psi)\le INF_{\Pi}I[Pi]$

D'autre part :

😊

Definition 1.4.1: Les fonction C-concaves (relatif au coût)

Soit, $\varphi:\mathbb{R}^n\to\mathbb{R}\cup\{-\infty\}$
On définit sa fonction C-conjuguée par :

$$\varphi^c:\left|\begin{array}{rcl}\mathbb{R}^n & \longrightarrow & \mathbb{R}\cup-\{\infty\}\\ y & \longmapsto & INF_{x\in\mathbb{R}^n}(C(x,y)-\varphi(x))\end{array}\right.$$

On dit que φ est C-concave si $\exists\psi:\mathbb{R}^n\to\mathbb{R}\cup\{-\infty\}$ tq $\forall x\in\mathbb{R}^n,\varphi(x)=\psi^c(x)$

Lemme 1.4.1

1. $\forall(x,y)\in\mathbb{R}^n\times\mathbb{R}^n,\varphi(x)+\varphi^c(y)\le C(x,y)$
2. $\varphi^c=\varphi^{ccc}$
3. $\varphi=\varphi^{cc}\Leftrightarrow\varphi$ est C-concave

Proof:

1. Par définition, $\forall y,\varphi^c(y)=INF_{x\in\mathbb{R}^n}(C(x,y)-\varphi(x))\le C(x,y)-\varphi(x)\forall x\in\mathbb{R}^n$
   Ainsi, $\forall(x,y)\in\mathbb{R}^n\times\mathbb{R}^n,\varphi(x)+\varphi^c(y)\le C(x,y)$

2. Par définition, $\varphi^{cc}(y)=INF_{x\in\mathbb{R}^n}(C(x,y)-\varphi^c(x))\ge\varphi(y)$
   En effet comme démontré ci-dessus, $\forall(x,y)\in\mathbb{R}^n\times\mathbb{R}^n,\varphi(y)\le C(x,y)-\varphi^c(x)$

   $\Rightarrow$ D'une part, $\varphi^c(x)=INF_y(C(x,y)-\varphi(y))\ge INF_y(C(x,y)-\varphi^{cc}(y))=\varphi^{ccc}(x)$
   Ainsi, $\boldsymbol{\varphi^c(x)\ge\varphi^{ccc}(x)}$

$\Rightarrow$ Et d'autre part, $INF_{x\in\mathbb{R}^n}(C(x,y) - \varphi^c(x)) \leq C(x,y) - \varphi^c(x) \forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n$
Donc, $\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n, \varphi^c(x) \leq C(x,y) - \varphi^{cc}(x)$
Donc, $\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n, \varphi^c(x) \leq INF_{x\in\mathbb{R}^n}(C(x,y) - \varphi^{cc}(x)) = \varphi^{ccc}(x))$
Ainsi, $\boldsymbol{\varphi^c(x) \leq \varphi^{ccc}(x))}$

Ce qui montre bien que,$\boldsymbol{\varphi^c(x) = \varphi^{ccc}(x)}$

3. $\Rightarrow$ Si φ est C-concave alors $\exists \psi | \varphi = \psi^c$
Donc, $\varphi^{cc} = \psi^{ccc} = \psi^c = \varphi$
$\Leftarrow$ Si $\varphi = \varphi^{cc}, alors \varphi = (\varphi^c)^c$ Donc φ est C-concave.

😊

Théoreme 1.4.2

La dualité de Kantorovitch peut être restreinte à des couples de fonction C-conjuguées.
$SUP_{(\varphi,\ \psi)\in(C(\mathbb{R}^n)^2} J(\varphi,\psi) = MAX_{(\psi^c,\psi)} J(\psi^c,\psi)$

Proof: On montre que le sup est un max. 😊

Corollaire 1.4.1 Les plans de transferts optimals sont caractérisés par leur support

Si (φ,ψ) est un maximiseur pour le problème de Kantorovitch dual, alors
$\Pi \in \Pi(\mu,\nu)$ est un minimiseur pour le problème de Kantorovitch primal si et seulement si Π est concentrée sur $\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n | \varphi(x) + \psi(y) = C(x,y)\}$

1.4.2 Application(s)

Proposition 1.4.1 On considère un objet indexé par j présent en quantité ν_j.
(type(j), quantjté(j)) = (j, ν_j)
On considère un consomateur indexé par i présent en quantité μ_i.
(type(i), quantité(i)) = (i, μ_i)
Utilité de l'objet j pour l'agent i.
Hypothèse L'utilité est tranférable.
L'objet j a une utilité nette $U_{r,j} - P$.
Pour un système de prix P_j, l'agent i choisit l'objet j_p qui maximise $U_{ij} - P_j$.
Transfert optimal: $SUP_\Pi \sum_{ij} U_{ij}\Pi_{ij}$ sous la contrainte $\sum_j \Pi_{ij} = \mu_i$; $\sum_i \Pi_{ij} = \nu_j$

Note:-

Explication :
$\mu = \sum_i \mu_i \delta_{x_i}$
$\nu = \sum_j \nu_j \delta_{y_j}$
$C(i,j) = -u_{i,j}$
$I[\Pi] = -\sum_{ij} u_{ij}\Pi_{ij}$
Utilité maximale : $INF_\Pi I[\Pi] = SUP_\Pi\{-\sum_i \varphi_i\mu_i - \sum_j \psi_j \nu j\}$
Problème dual : $(D) : INF_{P_j}\{\sum_j \nu_j P_j + \sum_i \mu_i MAX_j(\nu ij - Pj)\}$

Definition 1.4.2: Prix d'équilibre

Un système de prix qui satisfait (D) est un prix d'équilibre du problème. Un tel système de prix permet d'atteindre l'optimum global

Note:-

Pour le mar. 29 nov. 2022
Lire Guillaume Carlier, Teaching, Transfert Optimal (Chapitre 3 Matching equilibre)

Faire exercice 5, 6, 8 et 9.

1.5 La distance de Wasserstein

Definition 1.5.1: Distance de Wasserstein

On munit $\mathbb{R}^n$ d'une distance d. On considère la fonction cout : $C(x, y) = [d(x, y)]^p$ Par convention, si $p = 0$, par convention, $(d(x, y))^0 = \begin{cases} = 0 | six = y \\ = 1 | sinon \end{cases}$

Definition 1.5.2

$\mathcal{P}_p(\mathbb{R}^n)$

Note:-

Si $p \in \mathcal{P}_p(\mathbb{R}^n)$ alors $\forall y \in \mathbb{R}^n$, $\displaystyle\int d(y, x)^p \, d\mu(x) < \infty$

Si d est borné $\mathcal{P}_p(\mathbb{R}^n) = \mathcal{P}(\mathbb{R}^n)$

Definition 1.5.3

Soit $p \geq 1$

On définit, $W_p^p(\mu, \nu) = INF_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^n \times \mathbb{R}^n} d(x, y)^p \, d\pi(x, y)$

Note:-

On remarque que, $W_p(\mu, \nu) = [\tau_p(\mu, \nu)]^{\frac{1}{p}}$

Lemme 1.5.1 Lemme de

Théoreme 1.5.1

W^p est une mesure

Definition 1.5.4: Inégalité Triangulaire

Soient $\mu_1, \mu_2, \mu_3 \in \mathcal{P}_p(\mathbb{R}^n)$ $\pi_{1,2}(\mu_1) = \mu_2$ $\pi_{2,3}(\mu_2) = \mu_3$

Des plans de transports optimaux. On note π la mesure de proba sur $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ définie dans le lemme définit ci-dessous.

On note $\pi_{1,3}$ sa marginale sur $\mathbb{R}^n \times \mathbb{R}^n$

On a $\pi_{1,3} \in \Pi(\mu_1, \mu_3)$ mais pas (forcément) optimal.

Note:-

Π à pour marginale $\pi_{1,3}$ sur $\mathbb{R}_1^n \times \mathbb{R}_2^n$

$\forall \varphi \in (\mathbb{R}^n \times \mathbb{R}^n) \int \varphi(x, y) \, d\Pi(x, y, z) = \int \varphi(x, y) \, d\pi_{1,3}(x, y)$

Proof:

😄

Exemple 1.5.1

$p = 2$

W_2 est la distance de Wasserstein quadratique.
Si $\mu \in \mathcal{P}_2(\mathbb{R}^n), a \in \mathbb{R}^n$
$W_2(\mu, \delta_a) = \int_{\mathbb{R}^n} (x-a)^2 \, d\mu(x)$
Et la moyenne de μ (son esperance) est définie comme : $m = \int x d\mu(x) \, dx$
$m = INF_{a\in\mathbb{R}^n} W_2(\mu,\ \delta_a)$

Note:-
Cela permet de définir des espérances sur des espaces compliqués dans lesquels on ne peut pas définir d'intégrale.

Exemple 1.5.2
$p = 1$
τ_1 s'apppelle la distance de Rubinstein-Kantorovitch.

Proposition 1.5.1

Exemple 1.5.3
Le cas $\mathbb{R}^n = \mathbb{R}$

Definition 1.5.5: CDF

$\mu\ \in \mathcal{P}(\mathbb{R})$ On définit sa CDF par : $F(x) = \int\ d$

Definition 1.5.6

On appelle F^{-1} l'inverse généralisé de F définie sur [0,1] par : $F^{-1}(t) = INF_{x\in\mathbb{R}|F(x)>t}$

Proposition 1.5.2
$W_1(\mu, \nu) = \int_0^1 F^{-1}(t) - G^{-1}(t) \, dt$
$W_1(\mu, \nu) = \int_{\mathbb{R}} |F(x) - G(x)| \, dx$
$W_1(\mu, \nu) = ||F - G||_{L_1(\mathbb{R})}$

Note:-
On note, $\mu_1, \ldots, \mu_n \in \mathcal{P}_2(\mathbb{R}^n)$
$\lambda_1, \ldots, \lambda_n \in \mathbb{R}^+$ tel que $\sum_{i=1}^n \lambda_i = 1$

Definition 1.5.7

$\mu^* = INF_{\mu\in\mathcal{P}_2(\mathbb{R})} \sum_{i=1}^n \lambda_i W_2^2(\mu, \mu_i)$
est le barycentre de $(\mu_1, \nu_1)\ldots(\mu_n, \lambda_n)$

Note:-
Voir cours : Peyré / Cuturi : Computational Transfert optimal.

Chapter 2

TD

2.1 TD1

Exercice

Exercice 1. On considère le coût $c(x,y) = |x-y|$. Dans chacun des cas, donner les solutions des problèmes de Monge et de Kantorovitch.
1. $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1, \quad \nu = \frac{1}{3}\delta_{-1} + \frac{1}{3}\delta_2 + \frac{1}{3}\delta_3,$
2. $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1, \quad \nu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1,$
3. $\mu = \frac{1}{3}\delta_0 + \frac{1}{3}\delta_1 + \frac{1}{3}\delta_2, \quad \nu = \frac{1}{3}\delta_{-1} + \frac{1}{3}\delta_0 + \frac{1}{3}\delta_3.$

Solution:

1. $0 \geq \alpha \geq \frac{1}{3}; 0 \geq \beta \geq \frac{1}{3}$
   $\frac{1}{6} \geq \alpha + \beta \geq \frac{1}{2}$
   $\Pi_{\alpha,\beta} = \alpha\delta_{(0,-1)} + \beta\delta_{(0,2)} + (\alpha+\beta)\delta_{(0,3)}$

2. Equivalent à l'exemple du cours.

Exercice

Exercice 2. Soit $T : \mathbb{R} \to \mathbb{R}$ définie par $T(x) = x+1, S : \mathbb{R} \to \mathbb{R}$ définie par $S(x) = 2x$ et $Z : \mathbb{R} \to \mathbb{R}$ définie par $Z(x) = 2-x$. On définit $\mu = \mathbb{1}_{[0,1]}$ et $\nu = \mathbb{1}_{[1,2]}$. A-t-on $T\#\mu = \nu$? $S\#\mu = \nu? Z\#\mu = \nu?$

Solution: Soit $\varphi \in C^0(\mathbb{R})$,

$$\int_{[0,1]} (\varphi o T)(x)\, d\mu(x) = \int_{[0,1]} \varphi(1+x)\, dx = \int_{[1,2]} \varphi(y)\, dy = \int \varphi(y)\, d\nu(y)$$

$\Rightarrow T\#\mu = \nu$

$$\int_{[0,1]} (\varphi o S)(x)\, d\mu(x) = \int_{[0,1]} \varphi(2x)\, dx = \frac{1}{2}\int_{[0,2]} \varphi(y)\, dy$$

$\Rightarrow S\#\mu = \frac{1}{2}\mathbb{1}_{[0,2]} \Rightarrow S\#\mu = \frac{1}{2}(\mu+\nu) \Rightarrow S\#\mu \neq \nu$

$$\int_{[0,1]} (\varphi o Z)(x)\, d\mu(x) = \int_{[0,1]} \varphi(2-x)\, dx = \int_{[1,2]} \varphi(y)\, dy = \int \varphi(y)\, d\nu(y)$$

$\Rightarrow Z\#\mu = \nu$

Exercice

Exercice 3. (Non-unicité pour un coût convexe - Book shifting). On définit $\mu = \frac{1}{2}\mathbb{1}_{[0,2]}$ et $\nu = \frac{1}{2}\mathbb{1}_{[1,3]}$ et le coût $c(x,y) = |x-y|$. Soit $T_1(x) = x+1$ et

$$T_2(x) = \begin{cases} x+2, \text{ si } x \in [0,1], \\ x, \text{ si } x \in (1,2]. \end{cases}$$

Montrer que T_1 et T_2 sont deux applications optimales.

Solution: Vérifions que :

$\boldsymbol{T_1 \# \mu = \nu}$

$$\int_{[0,2]} (\varphi o T_1)(x)\, d\mu(x) = \frac{1}{2}\int_{[0,2]} \varphi(x+1)\, dx = \frac{1}{2}\int_{[1,3]} \varphi(y)\, dy = \int \varphi(y)\, d\nu(y)$$

$$I(T_1) = \int_{\mathbb{R}} C(x, T_1(x))\, d\mu(x) = \frac{1}{2}\int_{[0,2]} |x - T_1(x)|\, dx = \frac{1}{2}\int_{[0,2]} 1\, dx = 1$$

$\boldsymbol{T_2 \# \mu = \nu}$

$$\int_{[0,2]} (\varphi o T_2)(x)\, d\mu(x) = \frac{1}{2}(\int_{[0,1]} \varphi(x+2)\, dx + \int_{[1,2]} \varphi(x)\, dx =)\ \frac{1}{2}(\int_{[2,3]} \varphi(x)\, dx + \int_{[1,2]} \varphi(x)\, dx) = \frac{1}{2}\int_{[1,3]} \varphi(x)\, dx$$

$$I(T_2) = \int_{[0,2]} C(x, T_2(x))\, d\mu(x) = \frac{1}{2}\int_{[0,1]} |x - T_2(x)|\, dx + 0 = \frac{1}{2}\int_{[0,1]} 2\, dx = 1$$

De façon générale, $\boldsymbol{T \# \mu = \nu}$ alors :

$$I(T) = \frac{1}{2}\int_{[0,2]} |x - T(x)|\, dx \geq \frac{1}{2}|\int_{[0,2]} |x|\, dx - \int_{[0,2]} |T(x)|\, dx| = \frac{1}{2}|2-4| = 1$$

Exercice

Exercice 4. (Non existence d'une application de transport). On prend μ la mesure uniforme sur $[0,1]$ et ν la mesure uniforme sur $[-1,1]$. On considère le coût $c(x,y) = \left(x^2 - y^2\right)^2$. 1. Pour tout entier n on définit l'application

$$T_n(x) = \begin{cases} 2x - \frac{k}{2n}, & \text{pour } x \in \left[\frac{k}{2n}, \frac{k+1}{2n}\right] \text{ si } k \text{ est pair,} \\ -2x + \frac{k+1}{2n}, & \text{pour } x \in \left[\frac{k}{2n}, \frac{k+1}{2n}\right] \text{ si } k \text{ est impair.} \end{cases}$$

Monter que $T_n \# \mu = \nu$ et montrer que

$$\lim_{n\to\infty} \int_0^1 c\,(x, T_n(x))\, d\mu(x) = 0.$$

2. En déduire qu'il n'esxite pas d'application de transport qui soit optimale. 3. Construire un plan de transport optimal.

Exercice

Exercice 5. (Transport quadratique et translation). On considère le coût $c(x,y) = (x-y)^2$ sur $\mathbb{R}^2$ Pour $a \in \mathbb{R}$, on définitit la translation $\tau_a(x) = x - a$. Soit f et g deux fonctions continues Le but est de montrer que

$$\mathcal{T}_c\,(f \circ \tau_a, g \circ \tau_b) = \mathcal{T}\,c(f,g) + (b-a)^2 + 2(b-a)\left(m_g - m_f\right),$$

où

$$m_f = \int_{\mathbb{R}} x f(x) dx, \quad m_g = \int_{\mathbb{R}} x g(x) dx.$$

1. Soit T une application optimale qui envoie f sur g. On définit S par $S(x) = T(x-a) + b$. Montrer que $S\#\,(f \circ \tau_a) = g \circ \tau_b$.
2. Montrer que

$$\mathcal{T}_c\,(f \circ \tau_a, g \circ \tau_b) \leqslant \int_{\mathbb{R}} |S(x) - x|^2 f\,(\tau_a(x))\, dx$$

3. En déduire que

$$\mathcal{T}_c\,(f \circ \tau_a, g \circ \tau_b) \leqslant \mathcal{T}\,c(f,g) + (b-a)^2 + 2(b-a)\left(m_g - m_f\right),$$

4. De même montrer que

$$\mathcal{T}_c\,(f \circ \tau_a, g \circ \tau_b) \geqslant \mathcal{T}\,c(f,g) + (b-a)^2 + 2(b-a)\left(m_g - m_f\right),$$

et en déduire (1)
5. En déduire que $\mathcal{T}_c\left(\mathbb{1}_{[0,1]}, \mathbb{1}_{[1,2]}\right) = 1$.

Solution:

1. Montrons que : $S\#(f\circ\tau_a)=g\circ\tau_b$
On sait que T est une appication optimal de f vers g. Donc, : $T\#f=g$
Donc, $\forall\psi$ continue : $\displaystyle\int(\psi\circ T)(x)\,df(x)=\int\psi(x)\,dg(x)$
Et, $\displaystyle\tau^c(f,g)=\int_{\mathbb{R}}|x-T(x)|^2\,df(x)$
D'autre part, $\forall\varphi$ continue : $\displaystyle\int(\varphi\circ S)(x)\,d(f\circ\tau_a)(x)=\int\varphi(T(x-a)+b)f(x-a)\,dx$
On pose,$u=x-a$
$\displaystyle\int\varphi(T(u)+b)f(u)\,du=\int\varphi\circ\tau_{-b}\circ T(u)f(u)\,du=\int\varphi\circ\tau_{-b}\circ T(u)g(u)\,du=\int\varphi(u+b)g(u)\,du$
On pose, $z=u+b$
$\displaystyle\int\varphi(z)g(z-b)\,dz=\int\varphi(z)g\circ\tau_b\,dz$
D'où : $S\#(f\circ\tau_a)=g\circ\tau_b$

2. Montrons que : $\mathcal{T}_c\left(f\circ\tau_a,g\circ\tau_b\right)\leqslant\int_{\mathbb{R}}|S(x)-x|^2f(\tau_a(x))dx$
$\mathcal{T}_c\left(f\circ\tau_a,g\circ\tau_b\right)=\int_{\mathbb{R}}|T(x)-x|^2f(\tau_a(x))dx\leqslant\int_{\mathbb{R}}|S(x)-x|^2f(\tau_a(x))dx$

3. Montrons que : $\mathcal{T}_c\left(f\circ\tau_a,g\circ\tau_b\right)\leqslant\mathcal{T}c(f,g)+(b-a)^2+2(b-a)\left(m_g-m_f\right),$
$\displaystyle\int_{\mathbb{R}}|S(x)-x|^2f(x-a)\,dx=\int_{\mathbb{R}}|T(x-a)+b-x|^2f(x-a)\,dx=\int_{\mathbb{R}}|T(y)+b-y-a|^2f(y)\,dy$
$\displaystyle=\int_{\mathbb{R}}|T(y)-y|^2f(y)\,dy+\int_{\mathbb{R}}|b-a|^2f(y)\,dy+\int_{\mathbb{R}}2|T(y)-y|\times|b-a|f(y)\,dy$
$=\mathcal{T}_c(f,g)+(b-a)^2+2(b-a)\left(m_g-m_f\right),$
D'où, $\mathcal{T}_c\left(f\circ\tau_a,g\circ\tau_b\right)\leqslant\mathcal{T}c(f,g)+(b-a)^2+2(b-a)\left(m_g-m_f\right),$

4. $\mathcal{T}_c\left(f\circ\tau_a,g\circ\tau_b\right)=\int_{\mathbb{R}}|T(x)-x|^2f(\tau_a(x))dx$
On a, $\mathcal{T}_c\left(f,g\right)=\mathcal{T}_c\left((f\circ\tau_a)\circ\tau_{-a},(g\circ\tau_b)\circ\tau_{-b}\right)\leqslant\mathcal{T}_c(f\circ\tau_a,g\circ\tau_b)+(b-a)^2+2(a-b)\left(m_{g\circ\tau_a}-m_{f\circ\tau_b}\right),$
Avec, $m_{f\circ\tau_a}=\int_{\mathbb{R}}xf(x-a)dx=\int_{\mathbb{R}}(y+a)f(y)dy,=a+m_f$
et, $\quad m_{g\circ\tau_b}=\int_{\mathbb{R}}xg(x-b)dx.=\int_{\mathbb{R}}(y+b)g(y)dy.=b+m_g$
Donc, $\mathcal{T}_c(f,g)\leq\mathcal{T}_c(f\circ\tau_a,g\circ\tau_b)+(b-a)^2+2(a-b)\left(b+m_g-a-m_f\right),$
Donc, $\mathcal{T}_c(f,g)\leq\mathcal{T}_c(f\circ\tau_a,g\circ\tau_b)+(b-a)^2-2(b-a)^2+2(a-b)\left(m_g-m_f\right),$
D'ou $\mathcal{T}c(f,g)+(b-a)^2+2(b-a)\left(m_g-m_f\right),\leqslant\mathcal{T}_c\left(f\circ\tau_a,g\circ\tau_b\right)$
Ainsi, $\mathcal{T}c(f,g)+(b-a)^2+2(b-a)\left(m_g-m_f\right),=\mathcal{T}_c\left(f\circ\tau_a,g\circ\tau_b\right)$

5. Pour, $f=g=\mathbb{1}_{[0,1]}$ et $\tau_a=\tau_0$ et $\tau_b=\tau_1$
$\tau_c(\mathbb{1}_{[0,1]},\mathbb{1}_{[1,2]})=\tau_c(\mathbb{1}_{[0,1]},\mathbb{1}_{[0,1]})+(1-0)^2+2(1-0)(m_g-m_f)=1$

Exercice

Exercice 6. (Non unicité des potentiels de Kantorovitch). Montrer que si (φ,ψ) est une paire optimale de potentiel de Kantorovitch, alors, pour tout $a\in\mathbb{R}$, la paire $(\varphi+a,\psi-a)$ l'est aussi.

Exercice

Exercice 7. (Potentiels de Kantorovitch). Donner au moins une paire optimale de potentiel de Kantorovitch pour le book-shifting problem.

Exercice

Exercice 8. Pour $x\in\mathbb{R}$ et $y\in\mathbb{R}$, on considère le cout $c(x,y)=\Psi(|x-y|)$ avec

$$\Psi(z)=\begin{cases}1-z, & 0\leqslant z\leqslant 1\\ z-1, & z\geqslant 1.\end{cases}$$

On considère les mesures $\mu = \frac{1}{2}(\delta_{-1} + \delta_2)$ et $\nu = \frac{1}{3}(\delta_{-1} + \delta_0 + \delta_1)$. Calculer le coût global du transport optimal entre μ et ν, et donner l'ensemble des plans de transport optimaux.

Solution: $0 \geq \alpha \geq \frac{1}{3}$; $0 \geq \beta \geq \frac{1}{3}$
$\frac{1}{6} \geq \alpha + \beta \geq \frac{1}{2}$
$\Pi_{\alpha,\beta} = \alpha\delta_{(-1,0)} + \beta\delta_{(-1,1)} + (\frac{1}{2} - \alpha - \beta)\delta_{(-1,-1)} + (\frac{1}{3} - \ \alpha)\delta_{(2,0)} + (\frac{1}{3} - \ \beta)\delta_{(2,1)} + (\alpha + \ \beta - \frac{1}{6})\delta_{(2,-1)}$
$I[\Pi_{\alpha\beta}] = \beta + \frac{1}{2} - \alpha - \beta + \frac{1}{3} - \alpha + 2\alpha + 2\beta - \frac{1}{3} = \frac{1}{2} + 2\ \beta$
$\Pi(\mu, \nu) = \{\Pi_{\alpha\beta=0} | \frac{1}{6} \geq \alpha \geq \frac{1}{3}\}$

Exercice

Exercice 9. Soit $A \in M_{n,m}(\mathbb{R}), b \in \mathbb{R}^m, c \in \mathbb{R}^n$. Montrer que

$$\sup_{Ax \leqslant b} c.x = \inf_{\substack{y \geqslant 0 \\ tAy = c}} b.y$$

Notation: on dit que $x \geqslant 0$ si toutes ses composantes sont positives, et ${}^tA$ est la transposée de la matrice A. Pour $c \in \mathbb{R}^n$ et $x \in \mathbb{R}^n$, on note $x.c$ le produit scalaire entre c et x. Indice: S'inspirer de la preuve de la dualité de Kantorovitch.

Solution: $SUP_{Ax \leq b} < C, x >= SUP_{x \in \mathbb{R}^n} < C, x > + \begin{cases} 0 | si Ax \leq b \\ -\infty | sinon \end{cases}$

Or, $\begin{cases} 0 | si Ax \leq b \\ -\infty | sinon \end{cases} = INF_{y \geq 0}[- < Ax - b, y >]$

Si $Ax \leq b$

Toutes les composante de $-(Ax - b)$ sont positives ou nulles. $\mathcal{L}y \geq 0$ (si toutes les composantes de y sont positives ou nulles) $< -(Ax + b), y >\geq 0$ Pour $y = 0, < -(Ax + b), y >= 0$

Si $Ax \leq b$ est faux
$\exists i | (Ax - b)_i \geq 0$

On prend : $y = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ n \\ \vdots \\ 0 \\ 0 \end{bmatrix}$

Exercice

Exercice 10. On définit pour $x \in \mathbb{R}$

$$G_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}.$$

Calculer $W_1(G_\sigma, \delta_0)$.

Exercice

Exercice 11 (Interpolation de McCann). Soient μ_0 et μ_1 des mesures à densité ρ_0 et ρ_1 par rapport à la mesure de Lebesgue. On considère le coût quadratique $c(x, y) = \|x - y\|_2^2$. Soit T l'application optimale de Monge. 1. Quel théorème garantit l'existence de T ? 2. On suppose que T est inversible. Montrer que T^{-1} envoie μ_1 sur μ_0 de façon optimale. 3. On définit pour $t \in [0, 1]$

$$T_t(x) = (1 - t)x + tT(x),$$

et

$$\mu_t = T_t \# \mu_0.$$

Montrer que $\left(T^{-1}\right)_{1-t}$ envoie μ_1 sur μ_t. 4. Montrer que

$$W_2\left(\mu_0, \mu_t\right) \leqslant t W_2\left(\mu_0, \mu_1\right).$$

5. Montrer que

$$W_2\left(\mu_1, \mu_t\right) \leqslant (1-t) W_2\left(\mu_0, \mu_1\right).$$

6. Calculer $W_2\left(\mu_0, \mu_t\right)$. 7. On prend

$$\rho_0(x) = \begin{cases} x+1, & x \in [-1,0] \\ 1-x, & x \in [0,1] \\ 0, & \text{sinon,} \end{cases} \qquad \rho_1(x) = \begin{cases} 1, & x \in [3,4], \\ 0, & \text{sinon.} \end{cases}$$

Calculer la densité ρ_t de μ_t.