

Notes de cours sur le Transport Optimal

Nathael Gozlan, Paul-Marie Samson, Pierre-André Zitt

21 octobre 2022

Table des matières

1	Le problème de Monge-Kantorovich	3
1.1	Le problème historique de Monge	3
1.2	Un cadre général pour le problème de Monge	4
1.3	Divers exemples élémentaires	5
1.4	Le problème de Monge-Kantorovich	7
a.	Permission de casser les cailloux	7
b.	Noyaux et couplages	8
c.	Le problème de Monge-Kantorovich	10
1.5	Plan du cours	11
2	Existence de plans de transport optimaux	12
2.1	Quelques rappels de topologie	12
a.	Semi-continuité inférieure	12
b.	Topologie faible sur l'espace des mesures de probabilité	14
2.2	Le théorème d'existence	15
3	Transport optimal en dimension un	17
3.1	Quelques généralités sur les fonctions de répartition	17
3.2	Transport optimal en dimension un	20
4	Monotonie cyclique	24
4.1	Monotonie cyclique	24
4.2	Fonctions c -convexes et ensembles c -cycliquement monotones	26
4.3	Caractérisation des plans de transport optimaux	29
4.4	Le théorème de dualité de Kantorovich	30
5	Existence d'applications de transport optimal	32
5.1	Préliminaires sur les fonctions convexes	32
5.2	Le théorème de Brenier	34
5.3	Le cas des coûts strictement convexes	36
6	Algorithmique du transport optimal	39
6.1	Diverses solutions	39
6.2	Différents problèmes	40
6.3	Dualité de Kantorovich et « lâcheté complémentaire »	43
6.4	L'algorithme hongrois	46
a.	Un premier exemple	46
b.	Un exemple complet	48

c.	Synthèse	51
6.5	L'algorithme des enchères	51
6.6	Remarques bibliographiques	54
7	Distance de Wasserstein	55
7.1	Définition, premières propriétés	55
7.2	Caractérisation duale de W_1	57
7.3	Lien avec la topologie faible	58
a.	Convergence faible : rappels	58
b.	Uniforme intégrabilité	60
c.	Deux faits élémentaires	62
d.	Preuve du théorème	63
7.4	Remarques bibliographiques	64
8	Applications en probabilités	65
8.1	Une preuve du théorème limite central	65
a.	Le résultat	65
b.	Distance à la normalité	66
c.	Preuve du TLC.	68
d.	Preuve des lemmes.	69
8.2	Convergence pour les chaînes de Markov par couplage	71
a.	Le résultat classique	71
b.	La chaîne M/M/ ∞	71
8.3	Remarques bibliographiques	73
9	Équation de Monge-Ampère et transport de Brenier	74
9.1	Introduction	74
9.2	L'équation de Monge-Ampère dans un cas simple	74
9.3	Outils d'analyse convexe : transformée de Legendre et dualité convexe	76
9.4	Outils d'analyse convexe : une généralisation de la Hessienne	81
9.5	Outils de théorie de la mesure	85
9.6	Équation de Monge-Ampère, formules de changement de variables	87
9.7	Remarques bibliographiques	91
10	Inégalités de Brunn-Minkowski et de Prékopa-Leindler	92
10.1	Inégalité de Brunn-Minkowski	92
10.2	Inégalité de Prékopa-Leindler	94
11	Inégalité isopérimétrique et inégalités de Sobolev sur \mathbb{R}^n.	96
11.1	Inégalité isopérimétrique sur \mathbb{R}^n	96
11.2	Inégalités de Sobolev	97
a.	Inégalité de Sobolev L^1	97
b.	Inégalité de Sobolev L^p , $1 < p < n$	100
12	Inégalités de Young optimales et de Brascamp-Lieb	102
12.1	Inégalités de Young optimale	102
12.2	Inégalités de Brascamp-Lieb et de Brascamp-Lieb inverse	103
12.3	Démonstration des inégalités de Brascamp-Lieb	107

Chapitre 1

Le problème de Monge-Kantorovich

1.1 Le problème historique de Monge

Dans son *Mémoire sur la théorie des déblais et des remblais* de 1781, le mathématicien Gaspard Monge posait la question pratique suivante : quelle est la manière la plus économique de remplir un trou avec un tas de sable ? Précisons que, dans la question de Monge, le coût engendré par un transport d'une unité de masse d'un point x vers un point y est égal à la distance euclidienne entre ces points.

Modélisons la situation de la manière suivante¹ : on considère deux sous-ensembles boréliens $A, B \subset \mathbb{R}^2$ tels que $0 < |A|, |B| < \infty$ (où $|\cdot|$ désigne la mesure de Lebesgue) et on leur associe les mesures de probabilité μ et ν définies par

$$d\mu(x) = \frac{\mathbf{1}_A(x)}{|A|} dx \quad \text{et} \quad d\nu(x) = \frac{\mathbf{1}_B(x)}{|B|} dx.$$

Les ensembles A et B délimitent respectivement l'emplacement du tas de sable et celui du trou. La mesure μ est la distribution *source* et représente la répartition de masse dans le tas de sable (ici, pour fixer les idées, la répartition est uniforme). La mesure *cible* ν représente quant à elle la répartition de masse dans le trou une fois celui-ci bouché (là encore on imagine une répartition uniforme pour fixer les idées).

Donnons quelques exemples :

Exemple 1.1. A est le carré de diagonale $(0, 0) - (-1, 1)$ et B le carré de diagonale $(0, 0) - (-1, 1)$.

Exemple 1.2. A est le carré unité et B le carré de diagonale $(0, 0) - (-1/2, -1/2)$.

Exemple 1.3. A est le carré unité et B un carré infinitésimal en 0

Le problème de Monge est de construire une application mesurable $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ transportant μ sur ν à coût minimal. Expliquons ce qu'on entend par là. On demande tout d'abord à T de transporter μ sur ν , c'est-à-dire de vérifier

$$\mu(\{x : T(x) \in C\}) = \nu(C)$$

pour tout sous-ensemble borélien $C \subset \mathbb{R}^2$. Autrement dit, la *mesure image* de μ par l'application T doit être égale à la mesure cible ν . En transport optimal, on utilise souvent la notation $T_{\#}\mu$ pour désigner la mesure image de μ par T . Rappelons qu'elle est définie par

$$T_{\#}\mu(C) = \mu(T^{-1}(C)) = \mu(\{x \in A : T(x) \in C\}).$$

1. La modélisation historique est dans \mathbb{R}^3 , mais les dessins sont plus faciles à faire dans \mathbb{R}^2 ...

Rappelons aussi qu'en terme d'intégration, on a le « théorème du transfert » : pour toute fonction borélienne bornée,

$$\int f(y)d\nu(y) = \int f(y)d(T_{\#}\mu)(y) = \int f(T(x))d\mu(x),$$

la définition précédente étant la spécialisation de cette formule pour les indicatrices d'ensembles.

Sur les exemples précédents, on peut donner plusieurs applications qui effectuent ce transport. Dans le premier exemple, la translation, la symétrie axiale et la symétrie centrale marchent toutes les trois. Dans le second exemple on peut faire l'homothétie de rapport $-1/2$, ou celle de rapport $(1/2)$ centrée en $(-1, -1)$. Dans le troisième cas, l'application $T : (x_1, x_2) \mapsto (0, 0)$ envoie bien μ sur δ_0 :

$$\int f(T(x_1, x_2))d\mu(x_1, x_2) = \int f(0, 0)d\mu(x_1, x_2) = f(0, 0) = \int f(y_1, y_2)d\delta_0(y_1, y_2).$$

Parmi les applications effectuant le transport, certaines peuvent coûter plus que d'autres : on cherche donc à minimiser le coût de cette opération de transport. Dans l'approche historique de Monge, le coût pour transporter une unité de masse de x à $y = T(x)$ est donné simplement par la distance euclidienne entre x et y . Le coût engendré par une application de transport donnée T vaut donc

$$C_1(T) = \int |x - T(x)| d\mu(x),$$

on cherche donc à déterminer

$$C_1(\mu, \nu) = \inf\{C_1(T); T : A \rightarrow B, T_{\#}\mu = \nu\}$$

et, si possible, une application optimale T^* réalisant cet infimum.

On peut se convaincre, par un raisonnement heuristique, que pour être optimale l'application T^* doit au moins vérifier le principe de base suivant : si x_1 est envoyé sur y_1 et x_2 est envoyé sur y_2 , alors ou bien tous les points sont alignés, ou bien les segments $[x_1, y_1]$ et $[x_2, y_2]$ ne se coupent pas sauf si $x_1 = x_2$ ou $y_1 = y_2$. Raisonnons par l'absurde et supposons que les points ne sont pas alignés et $x_1 \neq x_2$ et $y_1 \neq y_2$ mais que les segments se coupent en un point c . Par l'inégalité triangulaire,

$$|x_1 - y_2| + |x_2 - y_1| < |x_1 - c| + |c - y_2| + |x_2 - c| + |c - y_1| = |x_1 - y_1| + |x_2 - y_2|.$$

L'inégalité stricte vient du fait que c ne peut pas appartenir à la fois aux segments $[x_1, y_1]$ et $[x_1, y_2]$ et aux segments $[x_2, y_2]$ et $[x_2, y_1]$. Par conséquent, il est moins coûteux d'envoyer x_1 sur y_2 et x_2 sur y_1 . On pourrait donc modifier T^* et obtenir un transport moins coûteux de μ sur ν , ce qui contredit l'optimalité de T^* .

Remarquons que pour le deuxième exemple, ceci exclut la première homothétie ; pour le premier exemple la symétrie centrale ne peut pas non plus être optimale.

1.2 Un cadre général pour le problème de Monge

Donnons une définition du problème de Monge dans un cadre général. On se place sur un espace métrique (\mathcal{X}, d) polonais (c'est-à-dire séparable et complet) et on se donne une *fonction de coût*

$$c : \mathcal{X} \times \mathcal{X} \rightarrow [0, +\infty]$$

définissant le coût de transport d'un point x à un point y . On note $\mathcal{P}(\mathcal{X})$ l'ensemble des mesures de probabilité boréliennes sur \mathcal{X} . On dira qu'une application mesurable $T : \mathcal{X} \rightarrow \mathcal{X}$ est une *application de transport* d'une probabilité μ sur une probabilité ν si $T_{\#}\mu = \nu$.

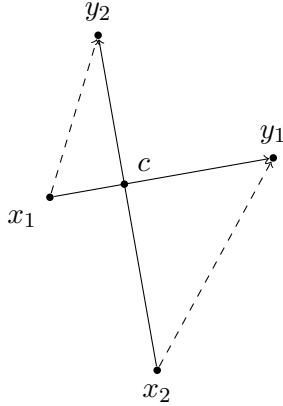


Illustration du raisonnement heuristique : dans la situation de la figure, il est plus économique d'envoyer (x_1, x_2) sur (y_1, y_2) en suivant les flèches en pointillés, dont la longueur totale est inférieure à celle des flèches en traits pleins.

FIGURE 1.1 – Pas de croisement pour les trajectoires

Définition 1.1. Soit $\mu, \nu \in \mathcal{P}(\mathcal{X})$; pour toute application de transport T de μ sur ν , on définit

$$C(T) = \int c(x, T(x)) d\mu(x).$$

Problème de Monge

Étant donné : Deux mesures de probabilité μ et ν , une fonction de coût c
 minimiser : le coût de transport $C(T) = \int c(x, T(x))$
 s. l. c. : l'application T transporte μ sur ν .

On notera $C(\mu, \nu)$ la valeur du problème, c'est à dire $\inf\{C(T); T : \mathcal{X} \rightarrow \mathcal{X}, T_{\#}\mu = \nu\}$. Une application T^* est dite *solution du problème de Monge* pour la fonction de coût c si $C(T^*) = C(\mu, \nu)$.

Dans le cas où $c(x, y) = c_p(x, y) = d(x, y)^p$, pour $p > 0$, on utilisera les notations $C_p(T)$ et $C_p(\mu, \nu)$.

On peut alors se poser les questions usuelles pour les problèmes d'optimisation :

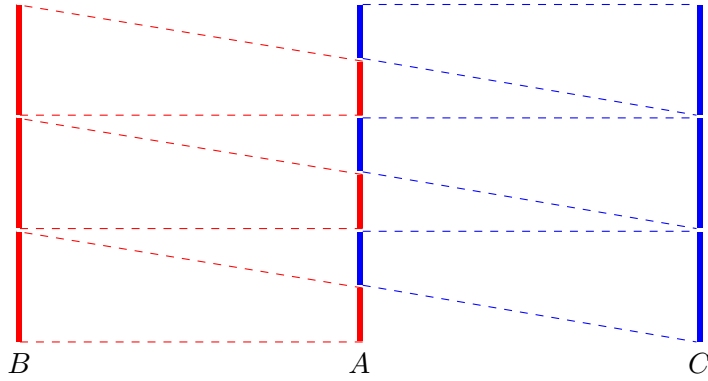
1. Le problème a-t-il une solution ?
2. Si oui, cette solution est-elle unique ?
3. Peut-on trouver une bonne caractérisation de cette solution ? Des conditions nécessaires, suffisantes, d'optimalité ?
4. En pratique comment peut-on résoudre le problème (numériquement) ?

1.3 Divers exemples élémentaires

Pour illustrer les difficultés des questions d'existence et d'unicité de solution, commençons par donner une batterie d'exemples, avec les coûts C_p dans \mathbb{R}^2 .

Exemple 1.4 (Un exemple où il n'y a pas d'application de transport). Reprenons le troisième exemple en inversant mesure source et mesure cible. Pour toute fonction T , l'image par T de la mesure de Dirac en 0 est la mesure de Dirac en $T(0)$. Autrement dit l'ensemble sur lequel on minimise est vide, et le problème n'a pas de solution.

Le problème de Monge n'est donc *pas symétrique*.



La mesure source est uniforme sur le segment du milieu A , la mesure cible uniforme sur l'union des deux segments extrêmes $B \cup C$.

FIGURE 1.2 – Construction d'applications quasi-optimales

Exemple 1.5 (Un exemple où une ou plusieurs applications optimales existent, suivant le coût). Pour le coût de Monge, dans le premier exemple (deux carrés côte à côte), on peut calculer le coût de la translation, le coût de la symétrie axiale. Ils sont égaux (à 1). On peut montrer que c'est le coût optimal, qui est donc réalisé par (au moins...) deux applications différentes.

Pour le coût quadratique, la symétrie coûte $4/3$ et la translation 1. On montrera que la translation est en réalité l'unique optimum.

Exemple 1.6 (Un exemple où il y a toujours plusieurs applications de transport optimales). On se place dans \mathbb{R}^2 et on désigne par A, B, C et D les sommets du carré $[0, 1] \times [0, 1]$: $A = (0, 0)$, $B = (0, 1)$, $C = (1, 1)$, $D = (1, 0)$. On considère $\mu = \frac{1}{2}\delta_A + \frac{1}{2}\delta_C$ et $\nu = \frac{1}{2}\delta_B + \frac{1}{2}\delta_D$. Dans ce cas, il y a exactement deux applications de transport : T_1 qui envoie A sur B et C sur D et T_2 qui envoie A sur D et C sur B . Les deux applications sont optimales pour tous les coûts C_p .

Exemple 1.7 (Un exemple où il y a des applications de transport, mais pas d'application optimale). On se place dans \mathbb{R}^2 muni de la fonction de coût $c_p(x, y) = |x - y|^p$, $x, y \in \mathbb{R}^2$ et on considère $A = \{0\} \times [0, 1]$, $B = \{-1\} \times [0, 1]$ et $C = \{1\} \times [0, 1]$. On considère ensuite les mesures de probabilité

$$\mu = \delta_0 \otimes m \quad \text{et} \quad \nu = \frac{1}{2}\delta_{-1} \otimes m + \frac{1}{2}\delta_1 \otimes m,$$

où m est la probabilité uniforme sur $[0, 1]$. Autrement dit, si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction borelienne bornée, on a

$$\int f(x) \mu(dx) = \int f(0, x_2) m(dx_2)$$

et

$$\int f(y) \nu(dy) = \frac{1}{2} \int f(-1, y_2) m(dy_2) + \frac{1}{2} \int f(1, y_2) m(dy_2).$$

D'un point de vue probabiliste, μ est la loi de $(0, X)$ où X suit la loi uniforme sur $[0, 1]$, et ν est la loi de (ε, X) où ε est une variable aléatoire indépendante de X , de loi de Rademacher $\mathbb{P}[\varepsilon = -1] = \mathbb{P}[\varepsilon = 1] = 1/2$. Montrons que $C_p(\mu, \nu) = 1$ mais qu'aucune application T ne réalise l'infimum. Tout d'abord, si T est une application de transport de μ sur ν , on a pour tout $x_2 \in [0, 1]$, $T(0, x_2) \in B \cup C$ et donc $|(0, x_2) - T(0, x_2)|^p \geq 1$. Par conséquent, $C_p(T) \geq 1$ et donc $C_p(\mu, \nu) \geq 1$.

Montrons à présent que $C_p(\mu, \nu) \leq 1$. Pour cela nous allons construire une suite d'applications T_n transportant μ sur ν telles que $C_p(T_n) \leq (1 + 1/(2n))^p$, $n \in \mathbb{N}^*$. La construction, illustrée sur la figure 1.2, est la suivante. Divisons le segment A en sous-segments verticaux $\{0\} \times I_1, \dots, \{0\} \times I_{2n}$ (numérotés de haut en bas) de longueur $1/(2n)$. De même, on divise les segments B et C en sous

segments $\{-1\} \times J_1, \dots, \{-1\} \times J_n \subset B$ et $\{1\} \times K_1, \dots, \{1\} \times K_n \subset C$ tous de longueur $1/n$. On définit ensuite une application T_n de la manière suivante : pour $x_2 \in I_{2p}$, $T_n(0, x_2) = (-1, f_{2p}(x_2))$, où f_{2p} est l'unique application affine croissante envoyant I_{2p} sur J_p et pour tout $x_2 \in I_{2p+1}$, $T_n(0, x_2) = (1, f_{2p+1}(x_2))$, où f_{2p+1} est l'unique application affine croissante envoyant I_{2p+1} sur K_p . On vérifie facilement que T_n transporte μ sur ν . Evaluons le coût de transport : on remarque que comme $J_p = K_p = I_{2p} \cup I_{2p+1}$, on a toujours $|x_2 - f_p(x_2)| \leq \frac{1}{2n}$ pour tout $x_2 \in I_p$ et donc $|(0, x_2) - T_n(0, x_2)| = \sqrt{1 + |x_2 - f_p(x_2)|^2} \leq 1 + \frac{1}{2n}$. Par suite, $C_p(T_n) \leq (1 + \frac{1}{2n})^p$, ce qui prouve que $C_p(\mu, \nu) = 1$.

Montrons finalement qu'aucune application T ne peut réaliser cet infimum. Procédons par l'absurde en supposant que T^* est une application de transport optimale. La fonction $x_2 \mapsto |(0, x_2) - T(0, x_2)|$ est minorée par 1, mais par optimalité

$$\int_0^1 |(0, x_2) - T(0, x_2)|^p m(dx_2) = 1,$$

donc elle vaut 1 Lebesgue presque sûrement. Par conséquent il existe une fonction $s : [0, 1] \rightarrow \{-1, 1\}$, mesurable, telle que $T(0, x_2) = (s(x_2), x_2)$ pour presque tout x_2 . Notons I_{\pm} les ensembles où s prend la valeur ± 1 . Au moins l'un de ces ensembles est de mesure positive, supposons que c'est le cas de I_+ . Alors

$$\int \mathbf{1}_{\{-1\} \times I_+} d\nu = \frac{1}{2} m(I_+) \neq 0 = \int \mathbf{1}_{\{-1\} \times I_+} dT_{\#}\mu,$$

et on a la contradiction cherchée.

Exercice 1.1 (Livres sur une étagère). *On considère une rangée de N livres posés sur une étagère dans la remise d'une librairie. Les livres en question sont identiques. On souhaite libérer l'espace occupé par le premier livre de la rangée (le plus à gauche pour fixer les idées). On envisage pour cela deux méthodes. La première consiste à prendre le premier livre et à le déplacer à l'extrémité droite de la rangée. La seconde méthode consiste à déplacer chaque livre d'un cran vers la droite. On modélise la situation sur \mathbb{R} par un problème de transport de $\mu(dx) = \frac{1}{N} \mathbf{1}_{[0, N]}(x) dx$ sur $\nu(dy) = \frac{1}{N} \mathbf{1}_{[1, N+1]}(y) dy$. Écrire les applications de transport correspondant aux deux méthodes envisagées et montrer qu'elles sont toutes les deux optimales pour le coût de transport de Monge $c_1(x, y) = |x - y|$, $x, y \in \mathbb{R}$.*

1.4 Le problème de Monge-Kantorovich

a. Permission de casser les cailloux

Au vu des résultats précédents, le problème de Monge apparaît comme étant très mal posé et relativement pathologique.

Dans les années 40 Leonid Kantorovich, dans une étude qui lui a valu le prix Nobel d'économie, a introduit une généralisation du problème qui a de bien meilleures propriétés mathématiques.

Pour introduire cette version relâchée, revenons aux exemples de non-existence (d'application de transport ou d'application optimale). Intuitivement, dans les deux cas, on sent qu'on doit pouvoir s'autoriser à répartir la masse présente en x , sans l'envoyer uniquement au point $T(x)$. Par exemple, la technique optimale pour transporter μ sur ν dans l'exemple des lignes est de *couper en deux* chaque élément de masse $\mu(dx_1 dx_2)$ autour du point $(0, x_2)$ et d'en envoyer la moitié sur le point $(-1, x_2)$ et l'autre moitié sur le point $(1, x_2)$.

On veut donc remplacer l'application de transport T par quelque chose de plus général. Une bonne façon de faire est de dire que la masse présente en x va être elle-même distribuée suivant une certaine mesure de probabilité : intuitivement, on peut « casser » le caillou présent en x pour répartir la poussière selon une certaine distribution. Ceci est formalisé par la notion de noyau :

Définition 1.2. Un noyau sur \mathcal{X} est une application $p : x \mapsto p_x(\cdot)$, de \mathcal{X} dans l'ensemble $\mathcal{P}(\mathcal{X})$ des probabilités sur \mathcal{X} , telle que pour toute fonction mesurable bornée $f : \mathcal{X} \rightarrow \mathbb{R}$, la fonction $x \mapsto \int f(y) p_x(dy)$ est Borel mesurable.

Étant donné une mesure source μ et un noyau de transition p , on peut définir l'analogue d'une mesure image (souvent noté $\nu = \mu p$)

$$\nu(C) = \int \int \mathbf{1}_C(y) p_x(dy) d\mu(x).$$

Intuitivement, tirer selon la loi ν correspond à tirer x suivant la loi μ puis y suivant la loi $p_x(\cdot)$.

Dans les deux cas précédents, il existe bien un noyau qui envoie la mesure source sur la mesure cible. Revenons à l'Exemple 1.7. Pour $x = (0, x_2)$ on définit le noyau

$$p_x(dy) = \frac{1}{2} \delta_{(-1, x_2)}(dy) + \frac{1}{2} \delta_{(1, x_2)}(dy)$$

Le cout de transport associé à ce noyau est

$$\iint |(0, x_2) - y|^p p_x(dy) d\mu(x) = \int \left(\frac{1}{2} |(0, x_2) - (1, x_2)|^p + \frac{1}{2} |(0, x_2) - (-1, x_2)|^p \right) m(dx_2) = 1.$$

Autoriser les noyaux permet donc d'atteindre une solution qui ne pouvait pas s'écrire par une application de transport.

b. Noyaux et couplages

Un aspect important de cette généralisation est qu'elle peut se réécrire d'une manière *symétrique* en μ et ν . En effet, d'un point de vue probabiliste on a décrit une façon de construire un couple (X, Y) :

- on tire X suivant la loi μ ,
- conditionnellement au tirage de $X = x$, on tire Y suivant la loi p_x .

Mathématiquement,

$$\mathbb{E}[\phi(X, Y)] = \int \int \phi(x, y) p_x(dy) d\mu(x) = \iint \phi(x, y) \pi(dx dy),$$

où π est une mesure sur $\mathcal{X} \times \mathcal{X}$. Si la fonction ϕ ne dépend que de x , $\iint \phi(x, y) d\pi(x, y) = \int \phi(x) \mu(dx)$; si ϕ ne dépend que de y $\iint \phi(y) d\pi(y) = \int \phi(y) d(\mu p)(y)$.

Définition 1.3 (Couplage). Soient $\mu, \nu \in \mathcal{P}(\mathcal{X})$; un couplage entre μ et ν est une probabilité π sur $\mathcal{X} \times \mathcal{X}$ dont la première marginale est μ et la seconde est ν . On note $\Pi(\mu, \nu)$ l'ensemble de tous les couplages entre μ et ν .

On rappelle que la première marginale de π , notée π_1 est l'image de π par la première application de coordonnée : $\mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X} : (x, y) \mapsto x$.

D'un point de vue probabiliste, un couplage n'est autre que la loi d'un *couple* de variables aléatoires (X, Y) définies sur le même espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ telles que X est distribuée selon μ et Y selon ν .

À toute mesure μ et tout noyau p on peut donc associer une mesure π sur $\mathcal{X} \times \mathcal{X}$ qui couple μ et $\nu = \mu p$. Réciproquement on peut vérifier qu'il est essentiellement équivalent de s'intéresser aux noyaux ou aux couplages.

Théorème 1.4. Soit $\pi \in \Pi(\mu, \nu)$; il existe un noyau de transition $p : \mathcal{X} \rightarrow \mathcal{P}(\mathcal{X})$ telle que pour toute fonction Borel mesurable bornée $f : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ on a

$$\int f(x, y) \pi(dxdy) = \int \left(\int f(x, y) p_x(dy) \right) \mu(dx).$$

Le noyau de transition p est unique μ -presque sûrement. En particulier, si (X, Y) est un couple de variable aléatoire distribué selon π , alors pour toute fonction mesurable bornée $f : \mathcal{X} \rightarrow \mathbb{R}$, on a, presque sûrement

$$\mathbb{E}[f(Y)|X] = \int f(y) p_X(dy).$$

On dit que $x \rightarrow p_x$ est une version régulière de la loi conditionnelle de Y sachant X .

On écrit de façon abrégée

$$\pi(dxdy) = \mu(dx)p_x(dy).$$

Remarque 1.1. Dans certain cas apparaissant fréquemment, le noyau p_x peut être écrit explicitement :

- Si π est la loi d'un couple (X, Y) où X est à valeurs discrètes, alors pour tout $x \in X$ tel que $\mathbb{P}(X = x) > 0$,

$$p_x(A) = \mathbb{P}(Y \in A | X = x) = \frac{\mathbb{P}(X = x, Y \in A)}{\mathbb{P}(X = x)}, \quad \forall A.$$

- Dans le cas où π est une probabilité sur $\mathbb{R}^n \times \mathbb{R}^n$ admettant une densité par rapport à Lebesgue : $\pi(dxdy) = h(x, y) dx dy$, alors pour μ presque tout $x \in \mathbb{R}^n$,

$$p_x(dy) = \frac{h(x, y)}{\int h(x, z) dz} dy.$$

Exemples de couplage :

- Couplage indépendant : $\mu \otimes \nu \in \Pi(\mu, \nu)$. Ce couplage par mesure produit correspond à la loi d'un couple (X, Y) de variables aléatoires indépendantes. En termes de noyau, on a pour tout x $p_x(dy) = \nu(dy)$: tous les éléments de masse sont répartis suivant ν , indépendamment de l'endroit d'où ils partent.
- Couplage déterministe. Si T est une application de transport de μ sur ν , alors $\pi_T = \text{Loi}(X, T(X))$ avec X une variable aléatoire de loi μ est un couplage de μ et ν . Il s'écrit :

$$\int \phi(x, y) \pi_T(dxdy) = \mathbb{E}[\phi(X, T(X))] = \int \phi(x, T(x)) \mu(dx).$$

En termes de noyau, p_x est maintenant une mesure de Dirac en $T(x)$, et on a pour toute fonction ϕ borélienne bornée

$$\iint f(x, y) \pi_T(dxdy) = \int f(x, y) \delta_{T(x)}(dy) \mu(dx) = \int \phi(x, T(x)) \mu(dx).$$

c. Le problème de Monge-Kantorovich

Comme précédemment, on se donne une fonction de coût mesurable $c : \mathcal{X} \times \mathcal{X} \rightarrow [0, +\infty]$.

Définition 1.5 (Coût de transport associé à un couplage). Soient $\mu, \nu \in \mathcal{P}(\mathcal{X})$; le coût de transport associé à un couplage $\pi \in \Pi(\mu, \nu)$ est défini par

$$I_c(\pi) = \iint c(x, y) \pi(dx dy).$$

Problème de Monge-Kantorovich

Étant donné : Deux mesures de probabilité μ et ν , une fonction de coût c
 minimiser : le coût de transport $I_c(\pi) = \int c(x, y) d\pi(x, y)$
 s. l. c. : π , mesure sur $\mathcal{X} \times \mathcal{X}$, a pour marginales μ et ν .

Le coût de transport optimal entre μ et ν est la valeur de ce problème :

$$\mathcal{T}_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} I_c(\pi).$$

Si π^* réalise le minimum, $I_c(\pi^*) = \mathcal{T}_c(\mu, \nu)$, on dit que π est un *couplage optimal* ou encore un *plan de transport optimal* de μ sur ν .

Remarque 1.2 (Monge-Kantorovich comme problème relâché). Si T envoie μ sur ν , le couplage déterministe donné par le noyau $p_x = \delta_{T(x)}$ couple μ et ν . Par conséquent, $I_c(\pi_T) = \int c(x, T(x)) \mu(dx) = C(T)$, et on a donc toujours

$$C(\mu, \nu) \geq \mathcal{T}_c(\mu, \nu).$$

Si on montre que le problème de Kantorovich admet un couplage optimal π^* et que ce couplage est déterministe, i.e. de la forme π_T , alors l'application T sera une solution du problème de Monge. En effet, on aura alors

$$C(\mu, \nu) \geq \mathcal{T}_c(\mu, \nu) = I_c(\pi_T) = C(T) \geq C(\mu, \nu)$$

et donc $C(T) = C(\mu, \nu)$.

L'avantage principal de la formulation de Kantorovich, développé dans le chapitre suivant, est que sous des hypothèses très générales sur la fonction de coût c , il existe toujours au moins un couplage optimal (pas toujours unique ni déterministe en revanche). Remarquons déjà une première différence : si l'existence d'applications de transport n'est pas garantie (on ne peut pas transporter un Dirac sur le carré), l'existence de couplage est toujours vraie (la mesure produit $\mu \otimes \nu$ a toujours pour marginales μ et ν).

Exercice 1.2. Montrer que la suite des couplages π_{T_n} associés aux applications T_n définies dans l'exemple 1.7 converge étroitement vers π^* lorsque $n \rightarrow \infty$.

1.5 Plan du cours

Donnons maintenant un bref aperçu du contenu du reste du cours.

Dans le chapitre suivant nous établirons l'existence de solution au problème relâché de Monge-Kantorovich. Nous donnerons ensuite un aperçu des résultats que l'on peut obtenir dans le cas simple de la dimension 1.

Les deux chapitres suivants concernent le problème de Monge et sa solution dans un cas particulier par Brenier. Pour le coût quadratique dans l'espace euclidien, on montrera le résultat suivant :

Théorème 1.6 (Brenier). *Soient μ, ν deux mesures de probabilité sur \mathbb{R}^n équipé de sa norme euclidienne canonique. Si μ est absolument continue par rapport à la mesure de Lebesgue et si $\int |x|^2 \mu(dx) < \infty$ et $\int |y|^2 \nu(dy) < +\infty$, alors il existe une application de transport T^* de μ sur ν , telle que $C_2(T^*) = C_2(\mu, \nu)$. Cette application T^* s'écrit sous la forme $T^* = \nabla \phi$, où $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction convexe.*

De plus, cette application optimale est essentiellement unique : si T est une autre application de transport telle que $C_2(T) = C_2(\mu, \nu)$, alors $T(x) = T^(x)$ pour μ -presque tout x .*

Remarque 1.3.

- *On verra qu'une fonction convexe est différentiable Lebesgue-presque partout sur l'intérieur de son domaine, donc μ -presque partout, ce qui permet de donner un sens à la formule $T^* = \nabla \phi$.*
- *Le cas $p = 1$ correspond au problème de Monge historique et est beaucoup plus difficile à étudier que le cas quadratique ; la structure des applications de transport est alors nettement moins rigide.*

On s'intéressera ensuite dans une parenthèse algorithmique à la résolution effective du problème de transport optimal, ce qui permettra de donner un autre éclairage sur la notion fondamentale de problème dual.

La notion de transport optimal sera ensuite appliquée dans deux directions différentes : en probabilité, où elle permet de définir une famille de distances naturelle entre lois de variables aléatoires, puis en analyse fonctionnelle, où elle donne un outil puissant pour établir des inégalités fonctionnelles par changement de variables.

Chapitre 2

Existence de plans de transport optimaux

Dans cette partie, nous allons établir l'existence de plans de transport optimaux, sous des hypothèses très générales. Ce résultat d'existence va découler d'un argument de compacité du type « une fonction continue sur un compact atteint ses bornes ». Nous allons introduire dans ce qui suit les quelques outils de topologie dont nous aurons besoin. Le premier outil est la notion de *semi-continuité inférieure*, plus faible que la notion de continuité, et qui permet de garantir l'atteinte de la borne inférieure sur un compact. Le deuxième outil est la *topologie de la convergence étroite* sur l'espace des probabilités et un critère permettant de reconnaître les ensembles compacts au sens de cette topologie.

2.1 Quelques rappels de topologie

a. Semi-continuité inférieure

On se place dans un espace métrique (E, d) .

Définition 2.1. Une fonction $f : E \rightarrow \mathbb{R} \cup \{+\infty\}$ est dite semi-continue inférieurement en $x \in E$ si pour toute suite $(x_n)_{n \in \mathbb{N}}$ convergeant vers x , on a

$$\liminf_{n \rightarrow \infty} f(x_n) \geq f(x). \quad (2.1)$$

Elle est dite semi-continue inférieurement si elle est semi-continue inférieurement en tout point de E .

Dans la suite, on abrègera « semi-continue inférieurement » en *sci*. Remarquons qu'une fonction continue (en x) est *sci* (en x).

Exercice 2.1. Montrer que, si A est un borélien quelconque, la fonction $\mathbf{1}_A$ est *sci* si et seulement si A est un ouvert.

Proposition 2.2. Une fonction $f : E \rightarrow \mathbb{R} \cup \{+\infty\}$ est *sci* sur E si et seulement si pour tout $r \in \mathbb{R}$, l'ensemble $\{x \in E : f(x) \leq r\}$ est fermé.

Démonstration. Supposons que f soit *sci* et montrons que les ensembles $\{f \leq r\}$, $r \in \mathbb{R}$ sont fermés. Fixons $r \in \mathbb{R}$ et prenons une suite $(x_n)_{n \in \mathbb{N}}$ à valeurs dans $\{f \leq r\}$ et convergeant vers un point $x \in E$. Pour tout $n \in \mathbb{N}$, on a $f(x_n) \leq r$. Par conséquent, $\liminf_{n \rightarrow \infty} f(x_n) \leq r$ et donc, d'après

(2.1), $f(x) \leq r$, ce qui prouve que $x \in \{f \leq r\}$. Par la caractérisation séquentielle des ensembles fermés, l'ensemble $\{f \leq r\}$ est bien fermé.

Réciproquement, supposons que les ensembles $\{f \leq r\}$, $r \in \mathbb{R}$, sont fermés et montrons que f est sci. Prenons $(x_n)_{n \in \mathbb{N}}$ une suite convergeant vers un point $x \in E$. Supposons $f(x)$ fini, et fixons $\varepsilon > 0$. Comme l'ensemble $\{f > f(x) - \varepsilon\}$ est ouvert et contient x , $f(x_n) > f(x) - \varepsilon$ à partir d'un certain rang. Par conséquent $\liminf f(x_n) \geq f(x) - \varepsilon$. Comme ε est arbitraire, f est bien sci en x . Si $f(x) = \infty$ on conclut de manière similaire en considérant les ouverts $\{f > M\}$. \square

Proposition 2.3. *Une fonction sci sur un compact est minorée et atteint sa borne inférieure.*

Démonstration. Soit $f : K \rightarrow \mathbb{R} \cup \{\infty\}$ une fonction sci définie sur un compact $K \subset E$. Raisonnons par l'absurde et supposons que f ne soit pas minorée. On peut alors construire une suite $(x_n)_{n \in \mathbb{N}}$ telle que $f(x_n) \rightarrow -\infty$ lorsque $n \rightarrow \infty$. Comme la suite est à valeurs dans un compact, on peut, sans perte de généralité, supposer qu'elle converge vers $x \in K$. La fonction f étant sci en x , on a donc $f(x) \leq \liminf_{n \rightarrow \infty} f(x_n) = -\infty$ — contradiction. On conclut que f est bien minorée. Montrons à présent qu'elle atteint sa borne inférieure $m = \inf_{x \in K} f(x)$. On commence par construire une suite $(x_n)_{n \in \mathbb{N}}$ telle que $f(x_n) \rightarrow m$ lorsque $n \rightarrow \infty$. Par compacité, on peut supposer que cette suite converge vers un certain $x \in K$. Comme f est sci en x on a alors $f(x) \leq \liminf_{n \rightarrow \infty} f(x_n) = m$. Comme par ailleurs, $f(x) \geq m$, on conclut que $f(x) = m$. \square

Proposition 2.4 (Sup de fonctions sci, régularisation).

1. Si $(f_i)_{i \in I}$ est une famille quelconque de fonctions sci, alors $f = \sup_{i \in I} f_i$ est sci.
2. Si f est une application sci minorée, alors il existe une suite croissante $(f_n)_{n \in \mathbb{N}}$ de fonctions continues bornées telles que $f = \sup_{n \in \mathbb{N}} f_n$. On peut même supposer que pour tout $n \in \mathbb{N}$, la fonction f_n est n -Lipschitz sur \mathcal{X} .

Démonstration. 1. Pour tout $r \in \mathbb{R}$, $\{f \leq r\} = \bigcap_{i \in I} \{f_i \leq r\}$. Pour chaque $i \in I$, l'ensemble $\{f_i \leq r\}$ est fermé, puisque f_i est sci. Donc l'ensemble $\{f \leq r\}$ est également fermé (intersection de fermés). D'après la Proposition 2.2, on conclut que f est sci.

2. Si f est identiquement égale à $+\infty$, alors $f_n = n$ convient. On suppose désormais que f prend au moins une valeur finie en un point x_o et on pose

$$f_n(x) = \inf_{y \in E} \{f(y) + nd(x, y)\}, \quad \forall x \in E.$$

Pour mieux comprendre cette définition on peut considérer le cas où f est l'indicatrice de l'ouvert $]0, 1[$, et tracer pour plusieurs valeurs de y la fonction $g_y(x) = f(y) + nd(x, y)$.

Reprenons le cas général. On voit immédiatement que pour tout $x \in E$, $f_n(x) \leq f(x)$ (prendre $y = x$) et $f_n(x) < +\infty$ (prendre $y = x_o$). On voit aussi simplement que $f_n \leq f_{n+1}$. En particulier, $f_n \geq f_0 = \inf f$ et donc les f_n sont minorées. De plus, f_n est n -Lipschitz et donc continue. En effet, si $x_1, x_2 \in E$, alors, en appliquant l'inégalité triangulaire, on obtient

$$f_n(x_1) = \inf_{y \in E} \{f(y) + nd(x_1, y)\} \leq \inf_{y \in E} \{f(y) + nd(x_2, y) + nd(x_1, x_2)\} = f_n(x_2) + nd(x_1, x_2).$$

Finalement, montrons que $f_n(x)$ converge vers $f(x)$ lorsque $n \rightarrow \infty$. Il existe une suite $(y_n)_{n \in \mathbb{N}^*}$ telle que $f_n(x) + 1/n \geq f(y_n) + nd(x, y_n)$, pour tout $n \in \mathbb{N}^*$. Supposons d'abord que $f(x) < +\infty$. Soit $m \in \mathbb{R}$ un minorant de f ; on a alors $f(x) + 1/n \geq m + nd(y_n, x)$ et donc la suite $nd(y_n, x)$ est majorée, ce qui entraîne que y_n converge vers x . En passant à la limite dans l'inégalité $f(x) + 1/n \geq f_n(x) + 1/n \geq f(y_n)$, on obtient $f(x) \geq \liminf_{n \rightarrow \infty} f_n(x) \geq$

$\liminf_{n \rightarrow \infty} f(y_n) \geq f(x)$, la dernière inégalité venant de la semi-continuité inférieure de f . Étant croissante, la suite $f_n(x)$ converge donc vers $f(x)$ quand $n \rightarrow +\infty$.

Si $f(x) = +\infty$, raisonnons par l'absurde et supposons que la suite $f_n(x)$ est majorée par un certain M . En procédant comme précédemment on conclut que y_n converge vers x puis que $M \geq \liminf_{n \rightarrow \infty} f(y_n) = +\infty$ — contradiction. On en déduit que $f_n(x)$ n'est pas majorée et donc, comme cette suite est croissante, converge vers $+\infty$.

Pour finir, en posant $\tilde{f}_n = f_n \wedge n$, on obtient une suite de fonctions continues, bornées, n -Lipschitz convergeant en croissant vers f . □

b. Topologie faible sur l'espace des mesures de probabilité

Dans ce qui suit, (\mathcal{X}, d) est un espace polonais (métrique séparable et complet) et on va munir l'espace $\mathcal{P}(\mathcal{X})$ des mesures de probabilité de Borel sur \mathcal{X} d'une topologie appelée *topologie de la convergence étroite*. On notera $\mathcal{C}_b(\mathcal{X})$ l'ensemble des fonctions continues bornées sur \mathcal{X} .

Définition 2.5 (Topologie de la convergence étroite). *Un ensemble $O \subset \mathcal{P}(\mathcal{X})$ est dit ouvert si pour tout $\mu_o \in O$, il existe $\varphi_1, \dots, \varphi_p \in \mathcal{C}_b(\mathcal{X})$ et $r_1 > 0, \dots, r_p > 0$ tels que*

$$\bigcap_{i=1}^p \left\{ \mu \in \mathcal{P}(\mathcal{X}) : \left| \int \varphi_i d\mu - \int \varphi_i d\mu_o \right| < r_i \right\} \subset O$$

Exercice 2.2. *Montrer que la topologie faible est la topologie la moins fine rendant continues les applications $\mu \mapsto \int \varphi d\mu$, $\varphi \in \mathcal{C}_b(\mathcal{X})$.*

En particulier, une suite $(\mu_n)_{n \in \mathbb{N}}$ de mesures de probabilité converge vers μ pour la topologie de la convergence étroite (on dit qu'elle *converge étroitement*) si

$$\int \varphi d\mu_n \rightarrow \int \varphi d\mu, \quad \text{lorsque } n \rightarrow \infty,$$

pour toute $\varphi \in \mathcal{C}_b(\mathcal{X})$.

Remarque 2.1 (Terminologie). *La topologie de la convergence étroite est celle qui intervient en probabilités dans la notion de convergence en loi : dire qu'une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoire converge en loi signifie que la suite $(\mu_n)_{n \in \mathbb{N}}$ des lois des X_n converge étroitement.*

La topologie de la convergence étroite est *métrisable* par de nombreuses métriques. Citons, par exemple, la métrique de Fortet-Mourier notée $d_{F.M.}$ et définie par

$$d_{F.M.}(\mu, \nu) = \sup \left\{ \left| \int \varphi d\mu - \int \varphi d\nu \right| : \varphi : \mathcal{X} \rightarrow \mathbb{R} \text{ 1-Lipschitz et } \|\varphi\|_\infty \leq 1 \right\}$$

ou encore la distance de Levy-Prokhorov définie par

$$d_{L.P.}(\mu, \nu) = \inf \{ \varepsilon > 0 : \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ et } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon, \forall A \in \mathcal{B}(\mathcal{X}) \},$$

où $\mathcal{B}(\mathcal{X})$ désigne la tribu de Borel de (\mathcal{X}, d) et où, pour tout $A \subset \mathcal{X}$, A^ε désigne un *épaisi* de A défini par $A^\varepsilon = \{y \in \mathcal{X} : \exists x \in A, d(x, y) < \varepsilon\}$.

Proposition 2.6. *La distance de Fortet-Mourier et la distance de Levy-Prokhorov engendrent la topologie de la convergence étroite.*

Cette proposition est admise, mais nous montrerons en détail un résultat proche dans le chapitre 7 consacré à l'étude d'une distance liée au transport optimal, la distance de Wasserstein.

Étant métrisable, la topologie de la convergence étroite peut être « séquentialisée » au sens où les notions de fermeture et de compacité s'expriment en terme de *convergence de suites*. Ainsi, un ensemble (de probabilités) est fermé (pour la topologie étroite) si et seulement si toute suite convergente à valeurs dans cet ensemble a sa limite dans l'ensemble ; un ensemble est compact si et seulement si de toute suite à valeurs dans l'ensemble on peut extraire une sous-suite convergeant dans l'ensemble.

Nous aurons besoin du critère de compacité suivant dû à Prokhorov, qui décrit les compacts de $\mathcal{P}(\mathcal{X})$ à partir de ceux de l'espace \mathcal{X} lui-même.

Théorème 2.7. *Soit \mathcal{X} un espace polonais. Un ensemble $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ est d'adhérence compacte si et seulement si il est uniformément tendu : pour tout $\varepsilon > 0$, il existe un compact $K \subset \mathcal{X}$ tel que*

$$\sup_{\mu \in \mathcal{A}} \mu(\mathcal{X} \setminus K) \leq \varepsilon.$$

En particulier, quand on applique le résultat précédent au compact $K = \{\mu\}$, où $\mu \in \mathcal{P}(\mathcal{X})$, on voit que μ est *tendue*, c'est-à-dire que pour tout $\varepsilon > 0$, il existe un compact K_ε tel que $\mu(K_\varepsilon) \geq 1 - \varepsilon$.

2.2 Le théorème d'existence

On suppose que (\mathcal{X}, d) est un espace polonais.

Théorème 2.8. *Soit $c : \mathcal{X} \times \mathcal{X} \rightarrow [0, +\infty]$ une fonction de coût sci. Pour toutes mesures de probabilité $\mu, \nu \in \mathcal{P}(\mathcal{X})$, il existe $\pi^* \in \Pi(\mu, \nu)$ tel que*

$$\iint c(x, y) \pi^*(dxdy) = \mathcal{T}_c(\mu, \nu).$$

Démonstration. Pour tout $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, on rappelle que $I_c(\pi) = \iint c(x, y) \pi(dxdy) \in [0, +\infty]$. Procédons par étapes.

La fonction I_c est sci sur $\mathcal{P}(\mathcal{X} \times \mathcal{X})$. La fonction c étant positive et sci, on sait d'après le point 2. de la Proposition 2.4 qu'il existe une suite de fonctions $c_n \geq 0$ continues et bornées telle que $c = \sup_n c_n$. On notera I_{c_n} la fonction de coût associée à c_n . Par le théorème de convergence monotone, pour tout $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, $I_c(\pi) = \sup_{n \in \mathbb{N}} I_{c_n}(\pi)$. Comme c_n est continue bornée, par définition de la topologie de la convergence étroite sur $\mathcal{P}(\mathcal{X} \times \mathcal{X})$, la fonction I_{c_n} est continue. D'après le point 1. de la Proposition 2.4, la fonction I_c est donc elle-même sci comme sup de fonctions sci.

L'ensemble des couplages $\Pi(\mu, \nu)$ est compact. Rappelons que l'ensemble $\Pi(\mu, \nu)$ est un ensemble de couplages, autrement dit un ensemble de mesures de probabilités sur $\mathcal{X} \times \mathcal{X}$, et que $\mathcal{P}(\mathcal{X} \times \mathcal{X})$ est naturellement muni de la topologie étroite.

Montrons d'abord que $\Pi(\mu, \nu) \subset \mathcal{P}(\mathcal{X} \times \mathcal{X})$ est fermé. Soit π_n une suite d'éléments de $\Pi(\mu, \nu)$ convergeant étroitement vers un certain π , alors pour toute $f \in \mathcal{C}_b(\mathcal{X})$,

$$\int f(x) \mu(dx) = \iint f(x) \pi_n(dxdy) \rightarrow \iint f(x) \pi(dxdy),$$

lorsque $n \rightarrow \infty$. On en déduit que pour toute $f \in \mathcal{C}_b(\mathcal{X})$, on a $\int f(x) \pi(dxdy) = \int f(x) \mu(dx)$. De même, on montre que $\iint f(y) \pi(dxdy) = \int f(y) \nu(dy)$. Ces égalités ayant lieu pour toute $f \in \mathcal{C}_b(\mathcal{X})$, on conclut que π admet bien μ et ν pour marginales : une limite étroite de couplages reste un couplage, et $\Pi(\mu, \nu)$ est donc bien fermé.

Montrons maintenant que $\Pi(\mu, \nu)$ est tendu, ce qui entraînera la compacité de $\overline{\Pi(\mu, \nu)}$ ($= \Pi(\mu, \nu)$) par le critère de Prokhorov. D'après la remarque suivant le Théorème 2.7, les probabilités μ et ν sont tendues. Pour tout $\varepsilon > 0$, il existe donc $K_\varepsilon \subset \mathcal{X}$, tel que $\mu(\mathcal{X} \setminus K_\varepsilon) \leq \varepsilon$ et $\nu(\mathcal{X} \setminus K_\varepsilon) \leq \varepsilon$. Posons $L_\varepsilon = K_\varepsilon \times K_\varepsilon \subset \mathcal{X} \times \mathcal{X}$ qui est compact comme produit de compacts. Le complémentaire de L_ε est l'union (non-disjointe!) des produits $(\mathcal{X} \setminus K_\varepsilon) \times \mathcal{X}$ et $\mathcal{X} \times (\mathcal{X} \setminus K_\varepsilon)$. Si $\pi \in \Pi(\mu, \nu)$, on a

$$\pi(\mathcal{X} \times \mathcal{X} \setminus L_\varepsilon) \leq \pi((\mathcal{X} \setminus K_\varepsilon) \times \mathcal{X}) + \pi(\mathcal{X} \times (\mathcal{X} \setminus K_\varepsilon)) \leq 2\varepsilon.$$

Par conséquent la famille $\Pi(\mu, \nu)$ est uniformément tendue et donc d'adhérence compacte.

Conclusion. On peut maintenant conclure grâce à la Proposition 2.3 : la fonction sci I_c atteint sa borne inférieure sur le compact $\Pi(\mu, \nu)$, ce qui, en d'autres termes, établit l'existence d'au moins un plan de transport optimal. \square

Chapitre 3

Transport optimal en dimension un

3.1 Quelques généralités sur les fonctions de répartition

Dans tout le chapitre, l'espace \mathbb{R}^d est muni de sa tribu borélienne, *i.e.* la plus petite tribu contenant la classe des ouverts de \mathbb{R}^d . On désignera par $\mathcal{P}(\mathbb{R}^d)$ l'ensemble des mesures de probabilité boréliennes sur \mathbb{R}^d .

Définition 3.1 (Fonction de répartition). *Soit μ une mesure de probabilité sur \mathbb{R}^d ; on appelle fonction de répartition de μ la fonction F définie sur \mathbb{R}^d par*

$$F(x) = \mu([\!-\infty, x_1] \times [\!-\infty, x_2] \times \cdots \times [\!-\infty, x_d]), \quad \forall x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Si X est une variable aléatoire de loi μ définie sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$, on a

$$F(x) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_d \leq x_d), \quad \forall x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Dans la suite, on notera souvent F_μ la fonction de répartition de μ .

On rappelle que la fonction de répartition caractérise la mesure de probabilité.

Proposition 3.2. *La fonction de répartition caractérise la loi. En d'autres termes, deux mesures de probabilité μ et ν sur \mathbb{R}^d sont égales si et seulement si $F_\mu = F_\nu$.*

Démonstration. Cette propriété est classique et se montre par un argument de classe monotone. \square

Définition 3.3 (Inverse généralisé d'une fonction de répartition). *Soit $F : \mathbb{R} \rightarrow \mathbb{R}$ la fonction de répartition d'une probabilité $\mu \in \mathcal{P}(\mathbb{R})$. L'inverse généralisé de F est la fonction notée F^{-1} définie sur $[0, 1]$ par*

$$F^{-1}(t) = \inf\{x \in \mathbb{R}; F(x) \geq t\}, \quad \forall t \in [0, 1].$$

Pour $p \in (0, 1)$, on appelle le nombre $F^{-1}(p)$ le *quantile d'ordre p* . En particulier, pour $p = 1/2$, $F^{-1}(1/2)$ est appelé la *médiane* de μ .

Remarque 3.1 ((Non)-Unicité). *On peut définir un autre inverse généralisé en considérant $\sup\{x \in \mathbb{R}, F(x) \leq t\}$. Les deux inverses coïncident si F est injective. Quand elle ne l'est pas, il y a donc plusieurs notions différentes de quantiles.*

Exercice 3.1. *Vérifier que si F est une bijection de \mathbb{R} dans $[0, 1]$, alors l'inverse généralisé de F coïncide bien avec la fonction réciproque de F .*

Proposition 3.4. Soit μ une probabilité sur \mathbb{R} et F sa fonction de répartition.

1. F est continue à droite et F^{-1} continue à gauche.
2. F est continue en $x \in \mathbb{R}$ si et seulement si $\mu(\{x\}) = 0$.
3. F est constante sur $[a, b]$ ou sur $]a, b]$ si et seulement si $\mu(]a, b]) = 0$.
 F est constante sur $[a, b[$ ou sur $]a, b[$ si et seulement si $\mu(]a, b[) = 0$.
4. $F^{-1}(p) \leq x$ si et seulement si $p \leq F(x)$.
 $F^{-1}(p) > x$ si et seulement si $p > F(x)$.
5. F^{-1} est continue en $p \in (0, 1)$ si et seulement si $\{x \in \mathbb{R}; F(x) = p\}$ est réduit à un point.
6. $F \circ F^{-1}(p) \geq p$ avec égalité sauf si $F^{-1}(p)$ est un atome de μ .
7. $F^{-1} \circ F(x) \leq x$ avec égalité sauf si $\mu(]x - \varepsilon, x]) = 0$ pour un certain $\varepsilon > 0$. Autrement dit l'égalité a lieu sauf si x est à l'intérieur ou à l'extrémité droite d'un palier de F .

Démonstration.

1. Si $x_n \downarrow x$, alors $\bigcap_{n \in \mathbb{N}}]-\infty, x_n] =]-\infty, x]$ et donc

$$F(x_n) = \mu(]-\infty, x_n]) \rightarrow \mu(]-\infty, x]) = F(x)$$

lorsque $n \rightarrow \infty$, ce qui prouve la continuité à droite.

2. Si $x_n \uparrow x$, alors $\bigcup_{n \in \mathbb{N}}]-\infty, x_n] =]-\infty, x[$ et donc

$$F(x_n) = \mu(]-\infty, x_n]) \rightarrow \mu(]-\infty, x[) = F(x) - \mu(\{x\}),$$

lorsque $n \rightarrow \infty$. On en déduit que F est continue à gauche en x si et seulement si $\mu(\{x\}) = 0$.

3. La fonction F est constante sur $]a, b]$ si et seulement si $\mu(]-\infty, x]) = \mu(]-\infty, y])$ pour tout $a < x \leq y < b$, ce qui équivaut à $\mu(]x, y]) = 0$ pour tout $a < x \leq y < b$. Comme $]a, b[= \bigcup_{b > y \geq x > a}]x, y]$, cela équivaut à $\mu(]a, b[) = 0$. Les autres cas se montrent de la même manière.
4. Si $F^{-1}(p) \leq x$, alors pour tout $\varepsilon > 0$, $\inf\{y \in \mathbb{R} : F(y) \geq p\} < x + \varepsilon$ et donc il existe $y_\varepsilon < x + \varepsilon$ tel que $F(y_\varepsilon) \geq p$. Comme F est croissante, on a donc $F(x + \varepsilon) \geq p$ et comme F est continue à droite, $F(x) \geq p$. Réciproquement, si $F(x) \geq p$, alors par la définition même de $F^{-1}(p)$, on $F^{-1}(p) \leq x$.
5. Si $[a, b] \subset \{x \in \mathbb{R} : F(x) = p\}$, alors pour tout $q > p$, on a $F^{-1}(p) \leq a \leq b \leq F^{-1}(q)$. On en déduit facilement que F^{-1} est continue à droite en p si et seulement si $\{x \in \mathbb{R} : F(x) = p\}$ est réduit à un point.
6. En prenant $x = F^{-1}(p)$ dans la propriété 4., on obtient $F \circ F^{-1}(p) \geq p$. Or pour tout $s < F^{-1}(p)$, on a $F(s) < p$ et donc $\mu(]-\infty, F^{-1}(p)[) \leq p$. Par suite si $F \circ F^{-1}(p) > p$ on a $F \circ F^{-1}(p) > \mu(]-\infty, F^{-1}(p)[)$ et donc $\mu(\{F^{-1}(p)\}) > 0$. La réciproque est du même style.
7. En prenant $p = F(x)$ dans 4., on obtient $F^{-1} \circ F(x) \leq x$. De plus, $F^{-1}(F(x)) < x$ si et seulement s'il existe $\varepsilon > 0$ tel que $F^{-1}(F(x)) \leq x - \varepsilon$, ce qui se traduit d'après 4. en $F(x) \leq F(x - \varepsilon)$ et donc $F(x) = F(x - \varepsilon)$ ce qui équivaut à $\mu(]x - \varepsilon, x]) = 0$.

□

La proposition suivante donne un moyen concret pour simuler une variable aléatoire quand on connaît sa fonction de répartition.

Proposition 3.5. *Si F est la fonction de répartition d'une probabilité μ sur \mathbb{R} , et si U une variable aléatoire uniformément distribuée sur $[0, 1]$, alors la variable aléatoire $X = F^{-1}(U)$ est distribuée selon μ .*

Démonstration. Calculons la fonction de répartition de $X = F^{-1}(U)$. Pour tout $t \in \mathbb{R}$,

$$\mathbb{P}(X \leq t) = \mathbb{P}(F^{-1}(U) \leq t) = \mathbb{P}(U \leq F(t)) = F(t),$$

la deuxième inégalité venant de la propriété 4. de la proposition 3.4. □

La proposition suivante est une sorte de réciproque de la précédente.

Proposition 3.6. *Soit F la fonction de répartition d'une probabilité μ sur \mathbb{R} et X une variable aléatoire distribuée selon μ . Si μ est sans atome, alors $F(X)$ suit une loi uniforme sur $[0, 1]$.*

Démonstration. Déterminons la loi de $F(X)$. Sans perte de généralité, on peut supposer que $X = F^{-1}(U)$ avec U une loi uniforme sur $[0, 1]$. On sait que $F \circ F^{-1}(x) = x$ sauf si $F^{-1}(x)$ est un atome de μ . La probabilité μ n'ayant pas d'atome, on a donc $F(X) = U$. □

L'inégalité suivante va jouer un rôle important dans la suite.

Théorème 3.7 (Hoeffding-Fréchet). *Si $\pi \in \mathcal{P}(\mathbb{R}^2)$ est une mesure de probabilité sur \mathbb{R}^2 ayant pour marginales μ et ν , alors, en notant F la fonction de répartition de π , on a*

$$F_-(x, y) := [F_\mu(x) + F_\nu(y) - 1]_+ \leq F(x, y) \leq \min(F_\mu(x), F_\nu(y)) =: F_+(x, y), \quad \forall x, y \in \mathbb{R}.$$

Démonstration. Pour l'inégalité de droite,

$$F(x, y) = \mathbb{P}[X \leq x, Y \leq y] \leq \mathbb{P}[X \leq x] = F_\mu(x),$$

et de même $F(x, y) \leq F_\nu(y)$. Pour celle de gauche, $\mathbb{P}[X > x \text{ ou } Y > y] \leq \mathbb{P}[X > x] + \mathbb{P}[Y > y] = 2 - F_\mu(x) - F_\nu(y)$, donc

$$F(x, y) = \mathbb{P}[X \leq x \text{ et } Y \leq y] \geq F_\mu(x) + F_\nu(y) - 1. \quad \square$$

L'inégalité d'Hoeffding-Fréchet permet d'isoler deux éléments particuliers dans la classe $\Pi(\mu, \nu)$ de tous les couplages entre μ et ν .

Définition 3.8. *Soit μ, ν deux probabilités sur \mathbb{R} .*

1. *Le couplage croissant de μ et ν est défini par*

$$\pi_+ = \text{Loi}(F_\mu^{-1}(U), F_\nu^{-1}(U)),$$

2. *Le couplage décroissant de μ et ν est défini par*

$$\pi_- = \text{Loi}(F_\mu^{-1}(U), F_\nu^{-1}(1 - U)),$$

où U est une variable aléatoire distribuée uniformément sur $[0, 1]$.

Exercice 3.2. *Vérifier que π_+ (resp. π_-) admet pour fonction de répartition la fonction F_+ (resp. F_-) du théorème 3.7.*

Proposition 3.9. Soient μ, ν deux mesures de probabilité sur \mathbb{R} . On suppose que μ est sans atome et on note X une variable aléatoire de loi μ . Alors

$$\pi_+ = \text{Loi}(X, T(X)), \quad \text{et} \quad \pi_- = \text{Loi}(X, S(X))$$

avec $T(x) = F_\nu^{-1} \circ F_\mu(x)$ et $S(x) = F_\nu^{-1}(1 - F_\mu(x))$, $x \in \mathbb{R}$.

Autrement dit, π^+ et π^- sont déterministes :

$$\pi^+(dxdy) = \mu(dx)\delta_{T(x)}(dy) \quad \text{et} \quad \pi^-(dxdy) = \mu(dx)\delta_{S(x)}(dy).$$

On remarque que l'application T est croissante tandis que l'application S est décroissante, d'où le nom attribué aux deux couplages. L'application T est appelée parfois *réarrangement croissant* ou *transformation par quantiles* de μ sur ν . Cette dernière appellation vient du fait que

$$T(F_\mu^{-1}(p)) = F_\nu^{-1}(p), \quad \forall p \in (0, 1).$$

Démonstration. Comme μ est sans atome, $U := F_\mu(X)$ est une variable aléatoire uniformément distribuée sur $[0, 1]$. Par conséquent, $\pi_+ = \text{Loi}(F_\mu^{-1} \circ F_\mu(X), T(X))$. Mais on sait, d'après le point 4. de la proposition 3.4, que $F_\mu^{-1} \circ F_\mu(x) = x$ sauf s'il existe $\varepsilon > 0$ tel que $\mu(]x - \varepsilon, x]) = 0$. En notant A l'ensemble de ces points, on voit sans peine que

$$A \subset \cup_{a < b \in \mathbb{Q} \text{ t.q. } \mu(]a, b]) = 0}]a, b[$$

Ce dernier ensemble est une union dénombrable d'ensembles de probabilité nulle donc $\mu(A) = 0$. On conclut que $F_\mu^{-1} \circ F_\mu(X) = X$ avec probabilité 1. Par conséquent, $\pi_+ = \text{Loi}(X, T(X))$. On montre de même l'identité concernant π_- . \square

3.2 Transport optimal en dimension un

Dans tout ce qui suit, μ et ν sont deux probabilités sur \mathbb{R} ; on rappelle que l'ensemble de tous les couplages de μ et ν est noté $\Pi(\mu, \nu)$.

On considère une fonction de coût $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ mesurable et on supposera dans la suite que c vérifie l'une des deux conditions suivantes :

$$c(x_1, y_1) + c(x_2, y_2) \leq c(x_1, y_2) + c(x_2, y_1), \quad \forall x_1 \leq x_2 \text{ et } y_1 \leq y_2 \quad (3.1)$$

$$c(x_1, y_1) + c(x_2, y_2) \geq c(x_1, y_2) + c(x_2, y_1), \quad \forall x_1 \leq x_2 \text{ et } y_1 \leq y_2 \quad (3.2)$$

En multipliant ces inégalités par $1/2$, on obtient l'interprétation suivante : avec une fonction de coût c vérifiant la condition (3.1), si on veut transporter $\mu = \frac{1}{2}\delta_{x_1} + \frac{1}{2}\delta_{x_2}$ sur $\nu = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$, avec $x_1 \leq x_2$ et $y_1 \leq y_2$, il est moins coûteux d'envoyer x_1 sur y_1 et x_2 sur y_2 que de croiser les extrémités en envoyant x_1 sur y_2 et x_2 sur y_1 . En revanche, si c vérifie la condition (3.2), c'est l'autre transport qui est le moins coûteux. On peut se convaincre par un petit calcul que le premier transport est donné par le couplage π_+ tandis que le second est donné par π_- . Le résultat suivant montre essentiellement que si un coût vérifie (3.1) (resp. (3.2)), le couplage π_+ (resp. π_-) sera toujours optimal pour le transport entre deux probabilités μ, ν générales.

Dans ce qui suit, on dira que l'espérance d'une variable aléatoire Z existe si $\mathbb{E}[Z_+] < +\infty$ ou si $\mathbb{E}[Z_-] < +\infty$ et on notera dans ce cas $\mathbb{E}[Z] = \mathbb{E}[Z_+] - \mathbb{E}[Z_-] \in \mathbb{R} \cup \{\pm\infty\}$. Si $\mathbb{E}[Z] \in \mathbb{R}$, on dira que l'espérance est finie. On adoptera la même terminologie pour les intégrales.

Théorème 3.10. Soient $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de coût continue à gauche vérifiant (3.1) et μ, ν deux mesures de probabilité sur \mathbb{R} telles qu'il existe $x_o, y_o \in \mathbb{R}$ tels que $\int c(x_o, y) \nu(dy)$ et $\int c(x, y_o) \mu(dx)$ sont finies alors pour tout $\pi \in \Pi(\mu, \nu)$, on a

$$\int c(x, y) \pi_+(dxdy) \leq \int c(x, y) \pi(dxdy),$$

à condition que les intégrales existent. Autrement dit, pour tout couple de variables aléatoires (X, Y) de lois respectives μ et ν , on a

$$\mathbb{E}[c(X_+, Y_+)] \leq \mathbb{E}[c(X, Y)],$$

à condition que les espérances existent, en notant (X_+, Y_+) un couple distribué selon π_+ . En particulier, si c est minoré alors

$$\mathcal{T}_c(\mu, \nu) = \iint c(x, y) \pi$$

Théorème 3.11. Soient $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de coût continue à gauche vérifiant (3.2) et μ, ν deux mesures de probabilité sur \mathbb{R} telles qu'il existe $x_o, y_o \in \mathbb{R}$ tels que $\int c(x_o, y) \nu(dy)$ et $\int c(x, y_o) \mu(dx)$ sont finies alors pour tout $\pi \in \Pi(\mu, \nu)$, on a

$$\int c(x, y) \pi_-(dxdy) \leq \int c(x, y) \pi(dxdy),$$

à condition que les intégrales existent. Autrement dit, pour tout couple de variables aléatoires (X, Y) de lois respectives μ et ν , on a

$$\mathbb{E}[c(X_-, Y_-)] \leq \mathbb{E}[c(X, Y)],$$

à condition que les espérances existent, en notant (X_-, Y_-) un couple distribué selon π_- .

Avant de démontrer le théorème 3.10, nous allons prouver la proposition suivante qui donne une condition suffisante simple pour (3.1) et (3.2).

Proposition 3.12.

1. La condition (3.1) (resp. (3.2)) est réalisée si $c(x, y) = \theta(x - y)$, avec $\theta : \mathbb{R} \rightarrow \mathbb{R}$ une fonction convexe (resp. concave).
2. Plus généralement, une fonction $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 vérifie la condition (3.1) (resp. (3.2)), si et seulement si

$$\frac{\partial^2 c}{\partial x \partial y}(x, y) \leq 0, \quad \left(\text{resp.} \quad \frac{\partial^2 c}{\partial x \partial y}(x, y) \geq 0 \right) \quad (3.3)$$

3. Encore plus généralement, une fonction $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ continue à gauche vérifie (3.1) (resp. (3.2)) si et seulement si il existe une mesure positive (resp. négative) m sur \mathbb{R}^2 telle que

$$c(x_1, y_2) + c(x_2, y_1) - c(x_1, y_1) - c(x_2, y_2) = m([x_1, x_2] \times [y_1, y_2]), \quad \forall x_1 \leq x_2, \quad \forall y_1 \leq y_2.$$

Remarque 3.2.

1. La continuité à gauche de c signifie que si x_n, y_n sont deux suites croissantes convergent respectivement vers x et y , alors $c(x_n, y_n) \rightarrow c(x, y)$ lorsque $n \rightarrow +\infty$.

2. Si c est de classe \mathcal{C}^2 et vérifie (3.3), alors

$$m(dudv) = -\frac{\partial^2 c}{\partial x \partial y}(u, v) dudv.$$

En particulier, pour la fonction de coût $c_p(x, y) = |x - y|^p$, avec $1 < p$, la mesure m_p associée est donnée par

$$m_p(dudv) = p(p-1)|u-v|^{p-2} dudv.$$

On remarque que pour $p \in]1, 2[$, la fonction $(u, v) \mapsto |u-v|^{p-2}$ est bien localement intégrable.

3. Pour la fonction de coût $c(x, y) = |x - y|$, la mesure m associée est la mesure de Lebesgue sur la diagonale :

$$\iint f(u, v) m(dudv) = \int f(x, x) dx.$$

Démonstration. 1. Il s'agit de montrer que si θ est convexe, on a

$$\theta(x_1 - y_1) + \theta(x_2 - y_2) \leq \theta(x_1 - y_2) + \theta(x_2 - y_1)$$

pour tous $x_1 < x_2$ et $y_1 < y_2$. En notant $a = x_1 - y_2$, $b = x_1 - y_1$ et $c = x_2 - x_1$, cette inégalité s'écrit

$$\theta(b) + \theta(a + c) \leq \theta(a) + \theta(b + c),$$

ou encore

$$\frac{1}{c} [\theta(b + c) - \theta(b)] \geq \frac{1}{c} [\theta(a + c) - \theta(a)].$$

On reconnaît là l'inégalité des pentes qui caractérise les fonctions convexes.

2. Pour tout $x = (x_1, x_2)$ et $y = (y_1, y_2)$, notons

$$\Delta(x, y) = c(x_1, y_2) + c(x_2, y_1) - c(x_1, y_1) - c(x_2, y_2)$$

la quantité intervenant dans (3.1). Comme c est supposée régulière, on peut écrire

$$\Delta(x, y) = \int_{y_1}^{y_2} \frac{\partial c}{\partial y}(x_1, v) dv - \int_{y_1}^{y_2} \frac{\partial c}{\partial y}(x_2, v) dv = - \int_{y_1}^{y_2} \left(\int_{x_1}^{x_2} \frac{\partial^2 c}{\partial x \partial y}(u, v) du \right) dv \quad (3.4)$$

La condition (3.1) équivaut donc à $\iint_P \frac{\partial^2 c}{\partial x \partial y}(u, v) dudv \leq 0$ pour tout pavé $P = [x_1, x_2] \times [y_1, y_2]$ avec $x_1 \leq x_2$ et $y_1 \leq y_2$. Comme $\frac{\partial^2 c}{\partial x \partial y}$ est continue sur \mathbb{R}^2 , cela équivaut à $\frac{\partial^2 c}{\partial x \partial y} \leq 0$ sur \mathbb{R}^2 .

3. On admettra le point 3. Voir par exemple le livre *Measure Theory and Integration, Second Edition* de M.M. Rao (p. 106-107). \square

Démonstration du théorème 3.10. Pour simplifier, on se place dans le cas où c est une fonction de classe \mathcal{C}^2 vérifiant $\frac{\partial^2 c}{\partial x \partial y} \leq 0$. D'après l'équation (3.4), on voit que

$$\begin{aligned} \tilde{c}(x, y) &:= c(x, y) - c(x, y_o) - c(x_o, y) + c(x_o, y_o) \\ &= \int_{y_o}^y \left(\int_{x_o}^x \frac{\partial^2 c}{\partial x \partial y}(u, v) du \right) dv \\ &= \iint k(x, y, u, v) \left| \frac{\partial^2 c}{\partial x \partial y}(u, v) \right| dudv, \end{aligned}$$

avec k la fonction définie par

$$k(x, y, u, v) = \begin{cases} -\mathbf{1}_{[x_o, x] \times [y_o, y]}(u, v) & \text{si } x_o \leq x \text{ et } y_o \leq y, \\ +\mathbf{1}_{[x, x_o] \times [y_o, y]}(u, v) & \text{si } x < x_o \text{ et } y_o \leq y, \\ +\mathbf{1}_{[x_o, x] \times [y, y_o]}(u, v) & \text{si } x_o \leq x \text{ et } y_o > y, \\ -\mathbf{1}_{[x, x_o] \times [y, y_o]}(u, v) & \text{si } x < x_o \text{ et } y < y_o. \end{cases}$$

En particulier, on remarque que $\tilde{c}_\pm(x, y) = \iint k_\pm(x, y, u, v) \left| \frac{\partial^2 c}{\partial x \partial y}(u, v) \right| dudv$.

Par conséquent, si (X, Y) est un couple de variable aléatoire avec $X \sim \mu$ et $Y \sim \nu$, alors, d'après le théorème de Fubini,

$$\begin{aligned} \mathbb{E}[\tilde{c}_\pm(X, Y)] &= \mathbb{E} \left[\iint k_\pm(X, Y, u, v) \left| \frac{\partial^2 c}{\partial x \partial y}(u, v) \right| dudv \right] \\ &= \iint \mathbb{E}[k_\pm(X, Y, u, v)] \left| \frac{\partial^2 c}{\partial x \partial y}(u, v) \right| dudv. \end{aligned}$$

Or,

$$\begin{aligned} \mathbb{E}[k_-(X, Y, u, v)] &= \mathbb{P}(x_o \leq u < X, y_o \leq v < Y) + \mathbb{P}(X \leq u < x_o, Y \leq v < y_o) \\ &= \mathbb{P}(u < X, v < Y) \mathbf{1}_{\{u \geq x_o, v \geq y_o\}} + \mathbb{P}(X \leq u, Y \leq v) \mathbf{1}_{\{u < x_o, v < y_o\}}. \end{aligned}$$

et

$$\begin{aligned} \mathbb{E}[k_+(X, Y, u, v)] &= \mathbb{P}(X \leq u < x_o, y_o \leq v < Y) + \mathbb{P}(x_o \leq u < X, Y \leq v < y_o) \\ &= \mathbb{P}(X \leq u, v < Y) \mathbf{1}_{\{u < x_o, v \geq y_o\}} + \mathbb{P}(u < X, Y \leq v) \mathbf{1}_{\{u \geq x_o, v < y_o\}}. \end{aligned}$$

En utilisant les identités

$$\begin{aligned} \mathbb{P}(X \leq u, Y \leq v) &= F(u, v) \\ \mathbb{P}(u < X, v < Y) &= 1 - F_\mu(u) - F_\nu(v) + F(u, v), \\ \mathbb{P}(X \leq u, v < Y) &= F_\mu(u) - F(u, v), \\ \mathbb{P}(u < X, Y \leq v) &= F_\nu(v) - F(u, v). \end{aligned}$$

et l'inégalité de Hoeffding-Fréchet $F(u, v) \leq F_+(u, v)$, on voit que pour tout $(u, v) \in \mathbb{R}^2$, on a

$$\mathbb{E}[k_-(X, Y, u, v)] \leq \mathbb{E}[k_-(X_+, Y_+, u, v)],$$

et

$$\mathbb{E}[k_+(X, Y, u, v)] \geq \mathbb{E}[k_+(X_+, Y_+, u, v)],$$

où (X_+, Y_+) est distribué selon π_+ . On en déduit que

$$\mathbb{E}[\tilde{c}_+(X, Y)] \geq \mathbb{E}[\tilde{c}_+(X_+, Y_+)]$$

et

$$\mathbb{E}[\tilde{c}_-(X, Y)] \leq \mathbb{E}[\tilde{c}_-(X_+, Y_+)]$$

Par suite, $\mathbb{E}[\tilde{c}(X, Y)] \geq \mathbb{E}[\tilde{c}(X_+, Y_+)]$ dès que les espérances existent. Comme $\mathbb{E}[c(x_o, Y)] = \mathbb{E}[c(x_o, Y_+)]$ et $\mathbb{E}[c(X, y_o)] = \mathbb{E}[c(X_+, y_o)]$ sont supposées finies on conclut que $\mathbb{E}[c(X, Y)] \geq \mathbb{E}[c(X_+, Y_+)]$. \square

Chapitre 4

Monotonie cyclique

Dans ce chapitre, on établit une propriété, appelée monotonie cyclique, caractérisant les plans de transport optimaux. Cette propriété nous permettra en particulier de démontrer le théorème de Brenier au chapitre suivant.

4.1 Monotonie cyclique

Dans tout le chapitre $c : X \times X \rightarrow [0, +\infty[$ sera une fonction de coût *continue*. Les résultats que nous allons présenter restent vrais sous l'hypothèse un peu plus faible de semi-continuité de c , mais nous n'entrerons pas dans ces développements afin de ne pas alourdir inutilement les démonstrations.

Définition 4.1. *Un sous ensemble $\Gamma \subset X \times X$ est dit c -cycliquement monotone, si pour toute famille de points $(x_1, y_1), (x_2, y_2), \dots, (x_p, y_p) \in \Gamma$, on a*

$$\sum_{i=1}^p c(x_i, y_i) \leq \sum_{i=1}^p c(x_i, y_{i+1}),$$

en posant $y_{p+1} = y_1$.

La proposition suivante permet de mieux comprendre la notion et son lien avec le transport optimal.

Proposition 4.2. *Si $\Gamma \subset X \times X$ est c -cycliquement monotone, alors pour toute famille finie de couples $(x_i, y_i)_{i \in I}$ éléments de Γ , et toute permutation σ de l'ensemble d'indices I , on a*

$$\sum_{i \in I} c(x_i, y_i) \leq \sum_{i \in I} c(x_i, y_{\sigma(i)}).$$

En particulier, si les x_i sont deux à deux distincts, l'application T définie sur $\{x_i : i \in I\}$ par $T(x_i) = y_i$ est optimale dans le problème de transport de Monge pour la mesure source $\mu = \frac{1}{\#I} \sum_{i \in I} \delta_{x_i}$ et la mesure cible $\nu = \frac{1}{\#I} \sum_{i \in I} \delta_{y_i}$.

Démonstration. On se donne $(x_i, y_i)_{i \in I}$ une famille de points de Γ et σ une permutation de I ; on note $m = \#I$. Supposons d'abord que σ est un cycle : I s'écrit donc $I = \{i_0, \sigma(i_0), \dots, \sigma^{m-1}(i_0)\}$. Posons $x'_j = x_{\sigma^j(i_0)}$ et $y'_j = y_{\sigma^j(i_0)}$ — on a simplement numéroté la famille des (x_i) . Alors

$$\sum_{i \in I} c(x_i, y_i) = \sum_{j=1}^m c(x'_j, y'_j) \leq \sum_{j=1}^m c(x'_j, y'_{j+1})$$

par définition de la monotonie cyclique appliquée aux (x'_j, y'_j) . Comme $y'_{j+1} = y_{\sigma(\sigma^j(i_0))}$,

$$\sum_{i \in I} c(x_i, y_i) \leq \sum_{j=0}^{m-1} c(x_{\sigma^j(i_0)}, y_{\sigma(\sigma^j(i_0))}) = \sum_{i \in I} c(x_i, y_{\sigma(i)}).$$

Une permutation σ quelconque de I peut toujours se décomposer en un produit de cycles à supports disjoints : on peut partitionner I en $\bigcup_k I_k$ tel que σ_{I_k} soit un cycle. Alors

$$\begin{aligned} \sum_{i \in I} c(x_i, y_i) &= \sum_k \sum_{i \in I_k} c(x_i, y_i) \\ &\leq \sum_k \sum_{i \in I_k} c(x_i, y_{\sigma(i)}) = \sum_{i \in I} c(x_i, y_{\sigma(i)}). \end{aligned} \quad \square$$

La monotonie cyclique permet de caractériser les couplages optimaux. Pour cela rappelons d'abord la définition suivante.

Définition 4.3 (Support). *Le support d'une mesure de probabilité γ sur un espace métrique polonais (E, d) est le plus petit fermé F tel que $\gamma(F) = 1$. De façon équivalente,*

$$\text{supp } \gamma = \{x \in E, \forall \delta > 0, \gamma(B(x, \delta)) > 0\}.$$

Théorème 4.4 (Caractérisation des couplages optimaux). *Soient $c : X \times X \rightarrow [0, +\infty[$ une fonction de coût continue et $\mu, \nu \in \mathcal{P}(X)$ telles que $\mathcal{T}_c(\mu, \nu) < \infty$. Un couplage $\pi \in \Pi(\mu, \nu)$ est optimal si et seulement si son support est c -cycliquement monotone.*

Dans un premier temps, nous ne ferons que la preuve de l'implication « Optimalité \Rightarrow c -cyclique monotonie du support ». L'autre implication nécessite des outils qui seront introduits dans la section suivante.

Preuve de la nécessité de la c -monotonie cyclique. Soit $\pi \in \Pi(\mu, \nu)$ un couplage tel que $I_c(\pi) < \infty$. On va établir que si le support de π n'est pas c -cycliquement monotone, alors π n'est pas optimal. Par hypothèse, il existe des points $(x_1, y_1), \dots, (x_p, y_p)$ dans le support de π tels que

$$\sum_{i=1}^p c(x_i, y_i) - \sum_{i=1}^p c(x_i, y_{i+1}) > 0,$$

en notant $y_{p+1} = y_1$. Comme la fonction c est continue sur $X \times X$, on en déduit qu'il existe I_1, \dots, I_p et J_1, \dots, J_p des ouverts de X tels que $x_i \in I_i$ et $y_i \in J_i$ et tels que, pour un certain $\varepsilon > 0$, on ait

$$\sum_{i=1}^p c(u_i, v_i) - \sum_{i=1}^p c(u_i, v_{i+1}) > \varepsilon > 0, \quad (4.1)$$

pour tous $u_i \in I_i, v_i \in J_i$. Comme $(x_i, y_i) \in I_i \times J_i$, on a, par définition du support de π , $\pi(I_i \times J_i) > 0$. On note $m_i = \pi(I_i \times J_i)$ et $\lambda = \min_i m_i$. On pose ensuite $\gamma_i = \frac{1}{m_i} \pi(\cdot \cap (I_i \times J_i))$. Les γ_i correspondent à des transferts de masse de I_i vers J_i . D'après (4.1), les transferts de masse de I_i vers J_{i+1} sont plus avantageux. On définit donc $\tilde{\gamma}_i$ comme la loi de (X_i, Y_{i+1}) où $(X_1, Y_1), \dots, (X_p, Y_p)$ est une suite de variables aléatoires indépendantes, telles que (X_i, Y_i) suit la loi γ_i . Finalement on considère π' défini par

$$\pi' = \pi + \frac{\lambda}{p} \sum_{i=1}^p (\tilde{\gamma}_i - \gamma_i).$$

Si $f \geq 0$, alors

$$\frac{\lambda}{p} \sum_{i=1}^p \iint f(x, y) \gamma_i(dx dy) \leq \iint f(x, y) \left(\frac{1}{p} \sum_{i=1}^p \mathbf{1}_{I_i \times J_i} \right) \pi(dx dy) \leq \iint f(x, y) \pi(dx dy).$$

Par conséquent, π' est bien une mesure positive. De plus, pour toute fonction borélienne bornée $f : X \times X \rightarrow \mathbb{R}$,

$$\iint f(x, y) \pi'(dx dy) = \iint f(x, y) \pi(dx dy) + \frac{\lambda}{p} \mathbb{E} \left[\sum_{i=1}^p (f(X_i, Y_{i+1}) - f(X_i, Y_i)) \right]$$

donc, en prenant successivement $f(x, y) = 1$, $f(x, y) = \varphi(x)$ puis $f(x, y) = \psi(y)$, on vérifie que $\pi' \in \Pi(\mu, \nu)$. Enfin, en prenant $f(x, y) = c(x, y)$, on obtient

$$I_c(\pi') = I_c(\pi) + \frac{\lambda}{p} \mathbb{E} \left[\sum_{i=1}^p (c(X_i, Y_{i+1}) - c(X_i, Y_i)) \right] \leq I_c(\pi) - \frac{\lambda \varepsilon}{p} < I_c(\pi),$$

ce qui montre bien que π n'est pas optimal. □

4.2 Fonctions c -convexes et ensembles c -cycliquement monotones

Commençons par définir deux opérateurs. On note $\overline{\mathbb{R}}$ l'ensemble $[-\infty, +\infty]$.

Définition 4.5 (Inf et sup-convolution). *Pour toute fonction $f : X \rightarrow \overline{\mathbb{R}}$, on définit l'inf-convolution de f par :*

$$Q_c f : y \mapsto \inf_{x \in X} \{f(x) + c(x, y)\}, \quad y \in X.$$

De même, la sup-convolution d'une fonction $g : X \rightarrow \overline{\mathbb{R}}$ est la fonction

$$P_c g : x \mapsto \sup_{y \in X} \{g(y) - c(x, y)\}, \quad x \in X.$$

Définition 4.6 (c -convexité). *Une fonction $f : X \rightarrow \overline{\mathbb{R}}$ est dite c -convexe s'il existe une fonction $g : X \rightarrow \overline{\mathbb{R}}$ telle que $f = P_c g$.*

Remarque 4.1 (Interprétation géométrique). *Déterminons l'épigraphe d'une fonction c -convexe $f = P_c g$;*

$$\text{epi}(f) = \{(x, t) \in X \times \mathbb{R} : t \geq f(x)\} = \bigcap_{y \in X} \{(x, t) \in X \times \mathbb{R} : t \geq g(y) - c(x, y)\}.$$

On dispose de formes géométriques de base : les épigraphes des fonctions de la forme

$$x \mapsto a - c(x, y), \quad a \in \mathbb{R}, y \in X.$$

Les fonctions c -convexes sont exactement les fonctions dont l'épigraphe est l'intersection d'une collection de ces formes géométriques. Par exemple dans le cas $c(x, y) = |x - y|^2$, l'épigraphe d'une fonction c -convexe est une intersection de dessus de paraboles concaves $x \mapsto -x^2 + bx + c$.

Remarque 4.2 (Liens avec les fonctions convexes). *Si $c(x, y) = -x \cdot y$, $x, y \in \mathbb{R}^n$, on verra plus loin qu'une fonction est c -convexe si et seulement si elle est convexe sci. En effet, l'épigraphe d'une telle fonction est un ensemble convexe fermé, intersection de demi-espaces fermés.*

Proposition 4.7. *Pour toute fonction $f : X \rightarrow \overline{\mathbb{R}}$, on a toujours $P_c Q_c f \leq f$. Une fonction f est c -convexe si et seulement si elle vérifie $P_c Q_c f = f$.*

Démonstration. Par définition,

$$P_c Q_c f(x) = \sup_{y \in X} \inf_{z \in X} \{f(z) + c(z, y) - c(x, y)\}.$$

On voit donc en prenant $z = x$ dans l'infimum que $P_c Q_c f(x) \leq \sup_{y \in X} \{f(x)\} = f(x)$. Si $P_c Q_c f = f$, alors $f = P_c g$, avec $g = Q_c f$ et donc, par définition, f est c -convexe. Réciproquement, supposons que $f = P_c g$, et vérifions que $P_c Q_c f = f$.

$$\begin{aligned} P_c Q_c f(x) &= \sup_{y \in X} \inf_{z \in X} \{f(z) + c(z, y) - c(x, y)\} \\ &= \sup_{y \in X} \inf_{z \in X} \sup_{\tilde{y} \in X} \{g(\tilde{y}) - c(z, \tilde{y}) + c(z, y) - c(x, y)\}. \end{aligned}$$

En prenant $\tilde{y} = y$ dans le deuxième supremum, on voit que

$$P_c Q_c f(x) \geq \sup_y \inf_z \{g(y) - c(x, y)\} = P_c g(x) = f(x).$$

Comme on a toujours $P_c Q_c f \leq f$, on conclut que $P_c Q_c f = f$. \square

Définition 4.8. *Soit $f : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ et $x \in X$. Le c -sous différentiel de f au point x est l'ensemble (éventuellement vide) noté $\partial_c f(x)$ des points y tels que*

$$f(x) = Q_c f(y) - c(x, y).$$

De manière équivalente, c'est l'ensemble des points y tels que

$$f(z) \geq f(x) + c(x, y) - c(z, y), \quad \forall z \in X.$$

On note $\partial_c f = \{(x, y) : x \in X, y \in \partial_c f(x)\}$ le graphe du c -sous-différentiel de f .

La notion de c -sous-différentiel permet de décrire les ensembles c -cycliquement monotones. Commençons par étudier le sens facile :

Proposition 4.9. *Si un ensemble Γ est inclus dans le c -sous différentiel d'une fonction c -convexe, alors il est c -cycliquement monotone.*

Démonstration. Comme un sous-ensemble d'un ensemble cycliquement monotone reste cycliquement monotone, on peut supposer que $\Gamma = \partial_c f$ pour une fonction c -convexe f . Prenons une famille de points $(x_1, y_1), \dots, (x_p, y_p)$ dans $\partial_c f$. Pour tout $i \in \{1, \dots, p\}$, on a

$$f(z) \geq f(x_i) + c(x_i, y_i) - c(z, y_i), \quad \forall z \in X.$$

Appliquons cette inégalité avec $z = x_{i-1}$, en notant $x_0 = x_p$, et sommons toutes les inégalités correspondantes :

$$0 = \sum_{i=1}^p (f(x_{i-1}) - f(x_i)) \geq \sum_{i=1}^p c(x_i, y_i) - \sum_{i=1}^p c(x_i, y_{i+1}),$$

ce qui prouve que $\sum_{i=1}^p c(x_i, y_i) \leq \sum_{i=1}^p c(x_i, y_{i+1})$. \square

La réciproque est vraie ! Pour comprendre la preuve qui suit, expliquons comment on peut en partie retrouver f à partir de son c -sous-différentiel. Supposons que $f : X \rightarrow \mathbb{R}$ soit c -convexe et telle que $\partial_c f(x) \neq \emptyset$ pour tout $x \in X$. Prenons une suite de points $(x_1, y_1), (x_2, y_2), \dots, (x_p, y_p)$ dans $\partial_c f$. On a

$$\begin{aligned} f(x) &\geq f(x_1) + c(x_1, y_1) - c(x, y_1) \\ &\geq f(x_2) + c(x_2, y_2) - c(x_1, y_2) + c(x_1, y_1) - c(x, y_1) \\ &\vdots \\ &\geq f(x_p) + \sum_{i=1}^p (c(x_i, y_i) - c(x_{i-1}, y_i)), \end{aligned}$$

en notant $x_0 = x$. Donc

$$f(x) \geq \sup_{p \in \mathbb{N}^*} \sup_{(x_1, y_1), \dots, (x_p, y_p) \in \partial_c f} \left\{ f(x_p) + \sum_{i=1}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) \right\}.$$

Par ailleurs, en prenant $p = 1$, et $x_1 = x$ et $y_1 \in \partial_c f(x)$, on voit que

$$\sup_{p \in \mathbb{N}^*} \sup_{(x_1, y_1), \dots, (x_p, y_p) \in \partial_c f} \left\{ f(x_p) + \sum_{i=1}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) \right\} \geq f(x).$$

Ainsi,

$$f(x) = \sup_{p \in \mathbb{N}^*} \sup_{(x_1, y_1), \dots, (x_p, y_p) \in \partial_c f} \left\{ f(x_p) + \sum_{i=1}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) \right\}.$$

A présent, fixons un point $a \in X$ et considérons

$$\varphi_a(x) = \sup_{p \in \mathbb{N}^*} \sup_{(x_i, y_i) \in \partial_c f, x_p = a} \left\{ \sum_{i=1}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) \right\}.$$

On a clairement, $f(x) = \sup_{a \in X} \{f(a) + \varphi_a(x)\}$.

Les fonctions φ_a vont nous servir pour démontrer le résultat suivant.

Théorème 4.10. *Soit $c : X \times X \rightarrow [0, +\infty[$ une fonction de coût continue et $\Gamma \subset X \times X$ un ensemble c -cycliquement monotone, alors il existe une fonction $\varphi : X \rightarrow \mathbb{R} \cup \{+\infty\}$ c -convexe telle que $\varphi < \infty$ sur $\Gamma_1 := \{x : \exists y, (x, y) \in \Gamma\}$ et telle que $\Gamma \subset \partial_c \varphi$.*

Démonstration. Bien sûr, on suppose que Γ est non vide... On fixe un point a arbitrairement choisi dans Γ_1 et on pose

$$\varphi_a(x) = \sup_{p \in \mathbb{N}^*} \sup_{(x_1, y_1), \dots, (x_p, y_p) \in \Gamma, x_p = a} \left\{ \sum_{i=1}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) \right\}, \quad x \in X,$$

en notant $x_0 = x$. Tout d'abord, $\varphi_a(a) \geq c(a, y) - c(a, y) = 0$ (y tel que $(a, y) \in \Gamma$) et donc $\varphi_a(a) \geq 0$. Par ailleurs, en utilisant la c -cyclique monotonie de Γ , on voit que $\varphi_a(a) \leq 0$ et donc $\varphi_a(a) = 0$ (c'est le seul endroit dans la preuve où on utilise l'hypothèse que Γ est c -cycliquement monotone). Montrons que φ_a est c -convexe.

$$\varphi_a(x) = \sup_{(x_1, y_1) \in \Gamma} \{ \psi(x_1, y_1) - c(x, y_1) \},$$

où

$$\psi(x_1, y_1) = c(x_1, y_1) + \sup \left\{ \sum_{i=2}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) : (x_2, y_2), \dots, (x_p, y_p) \in \Gamma, x_p = a \right\}.$$

On voit en particulier sur l'écriture de φ_a si dessus que si $x \rightarrow c(x, y_1)$ est une application K -Lipschitz pour tout y_1 alors φ_a est elle aussi K -Lipschitz. En posant, pour tout $y \in X$, $\bar{\psi}(y) = \sup\{\psi(x, y) : (x, y) \in \Gamma\}$ (avec la convention $\sup \emptyset = -\infty$), on voit donc que $\varphi_a(x) = \sup_{y \in X} \{\bar{\psi}(y) - c(x, y)\}$, ce qui prouve que φ_a est c -convexe. De plus, pour tout $(u, v) \in \Gamma$,

$$\begin{aligned} \varphi_a(x) &\geq \sup_{p \geq 2} \sup_{(x_i, y_i) \in \Gamma, (x_1, y_1) = (u, v), x_p = a} \left\{ \sum_{i=1}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) \right\} \\ &= c(u, v) - c(x, v) + \sup_{p \geq 2} \sup_{(x_i, y_i) \in \Gamma, x_1 = u, x_p = a} \left\{ \sum_{i=2}^p (c(x_i, y_i) - c(x_{i-1}, y_i)) \right\} \\ &\geq c(u, v) - c(x, v) + \varphi_a(u). \end{aligned}$$

Par conséquent, $\Gamma \subset \partial_c \varphi_a$. De plus, en prenant $u = a$ ou $x = a$, on voit que φ_a est à valeurs dans $\mathbb{R} \cup \{+\infty\}$ et que $\varphi_a < \infty$ sur Γ_1 . \square

4.3 Caractérisation des plans de transport optimaux

On est maintenant en mesure de terminer la preuve du Théorème 4.4. On va établir la forme plus précise suivante :

Théorème 4.11. *Soit $c : X \times X \rightarrow [0, \infty[$ une fonction de coût continue et μ, ν deux mesures de probabilité sur X telles que $\mathcal{T}_c(\mu, \nu) < \infty$. Pour un couplage $\pi \in \Pi(\mu, \nu)$, les propriétés suivantes sont équivalentes :*

1. π est optimal,
2. le support Γ de π est c -cycliquement monotone,
3. il existe une fonction c -convexe mesurable $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ finie sur $\Gamma_1 = \{x : \exists y \in X, (x, y) \in \Gamma\}$ telle que $Q_c f(y) = f(x) + c(x, y)$ pour π -presque tout (x, y) .

En utilisant la notion de noyau de transition (voir la Définition 1.2 et le Théorème 1.4), on peut donc écrire le couplage optimal π sous la forme suivante :

$$\pi(dxdy) = \mu(dx)p_x(dy),$$

où pour μ presque tout x , la probabilité p_x a son support dans l'ensemble $\partial_c f(x)$.

Démonstration. On a déjà montré que 1. \Rightarrow 2. et que 2. \Rightarrow 3. Montrons que 3. \Rightarrow 1. Comme c est supposée continue, on voit (par définition de la c -convexité) que f est un sup de fonctions continues et est donc, d'après la Proposition 2.4, semi-continue inférieurement. De même $Q_c f$ est un inf de fonctions continues et est donc semi-continue supérieurement. Les deux fonctions f et $Q_c f$ sont donc mesurables.

Traisons d'abord le cas particulier où c est bornée. Dans ce cas, $Q_c f(y) = \inf_{x \in X} \{f(x) + c(x, y)\} \leq \inf_x f(x) + M$, et donc $P_c Q_c f \leq \inf_x f(x) + 2M$, soit $f \leq \inf_x f(x) + 2M$. Cela entraîne en particulier que $\inf f > -\infty$ et que $\sup f < \infty$. On en déduit ensuite facilement que $Q_c f$ est également bornée. Les fonctions f et $Q_c f$ sont donc intégrables par rapport à π et on a

$$\iint c(x, y) \pi(dxdy) = \iint Q_c f(y) - f(x) \pi(dxdy) = \int Q_c f(y) \nu(dy) - \int f(x) \mu(dx).$$

Mais si $\pi' \in \Pi(\mu, \nu)$ est un autre couplage, on a

$$\int Q_c f(y) \nu(dy) - \int f(x) \mu(dx) = \iint Q_c f(y) - f(x) \pi'(dxdy) \leq \iint c(x, y) \pi'(dxdy),$$

où la dernière inégalité vient du fait que $Q_c f(y) \leq f(x) + c(x, y)$ pour tout $x, y \in X$. Par conséquent, $\iint c(x, y) \pi(dxdy) \leq \iint c(x, y) \pi'(dxdy)$ et donc π est optimal.

Traitons maintenant le cas général. Pour se ramener à des fonctions bornées, on utilisera le procédé de troncature suivant (assez subtil) :

$$T_n(u) = \begin{cases} u & \text{si } |u| \leq n \\ n & \text{si } u > n \\ -n & \text{si } u < -n \end{cases}$$

On pose ensuite $F(x, y) = Q_c f(y) - f(x)$ et pour tout n , $F_n(x, y) = T_n(Q_c f(y)) - T_n(f(x))$. On peut vérifier que sur l'ensemble $F(x, y) \geq 0$, $0 \leq F_n(x, y) \uparrow F(x, y)$, et que sur l'ensemble $F(x, y) < 0$ et $0 \geq F_n(x, y) \downarrow F(x, y)$. On a donc pour tout $(x, y) \in X \times X$,

$$F_n(x, y) \leq [F_n(x, y)]_+ = F_n(x, y) \mathbf{1}_{F \geq 0} \leq F(x, y) \mathbf{1}_{F \geq 0} \leq c(x, y)$$

Par conséquent, pour tout $\pi' \in \Pi(\mu, \nu)$,

$$\begin{aligned} \iint F_n(x, y) \pi(dxdy) &= \int T_n(Q_c f(y)) \nu(dy) - \int T_n(f(x)) \mu(dx) = \iint F_n(x, y) \pi'(dxdy) \\ &\leq \iint c(x, y) \pi'(dxdy). \end{aligned}$$

Par convergence monotone,

$$\iint F_n(x, y) \pi(dxdy) = \iint F_n(x, y) \mathbf{1}_{F \geq 0} \pi(dxdy) \rightarrow \iint F(x, y) \mathbf{1}_{F \geq 0} \pi(dxdy) = \iint c(x, y) \pi(dxdy),$$

et on en déduit que π est optimal. \square

4.4 Le théorème de dualité de Kantorovich

Théorème 4.12. *Soit $c : X \times X \rightarrow [0, \infty[$ une fonction semi-continue inférieurement, et $\mu, \nu \in \mathcal{P}(X)$ deux probabilités telles que $\mathcal{T}_c(\mu, \nu) < \infty$. On a alors*

$$\mathcal{T}_c(\mu, \nu) = \sup_{(\psi, \varphi) \in \Phi_c} \left\{ \int \psi(x) \mu(dx) + \int \varphi(y) \nu(dy) \right\}, \quad (4.2)$$

où Φ_c est l'ensemble des fonctions continues bornées telles que

$$\psi(x) + \varphi(y) \leq c(x, y), \quad \forall x, y \in X. \quad (4.3)$$

La formule (4.2) reste vraie si l'on élargit la classe Φ_c à la classe notée $\Phi_c(\mu, \nu)$ des couples $(\psi, \varphi) \in \mathbf{L}_1(\mu) \times \mathbf{L}_1(\nu)$ vérifiant (4.3).

Démonstration. Si $(\psi, \varphi) \in \Phi_c(\mu, \nu)$ et $\pi \in \Pi(\mu, \nu)$, alors on a toujours

$$\int \psi(x) \mu(dx) + \int \varphi(y) \nu(dy) = \iint \psi(x) + \varphi(y) \pi(dxdy) \leq \iint c(x, y) \pi(dxdy).$$

Donc en optimisant à la fois par rapport au couple $(\psi, \varphi) \in \Phi_c(\mu, \nu)$ et par rapport à $\pi \in \Pi(\mu, \nu)$, on obtient

$$\sup_{(\psi, \varphi) \in \Phi_c(\mu, \nu)} \left\{ \int \psi(x) \mu(dx) + \int \varphi(y) \nu(dy) \right\} \leq \mathcal{T}_c(\mu, \nu).$$

Reste à montrer l'autre inégalité. Supposons pour commencer que la fonction de coût c est continue et bornée. Considérons π^* un plan de transport optimal. D'après le Théorème 4.11, il existe une fonction c convexe $f : X \rightarrow \mathbb{R}$ telle que $Q_c f(y) = f(x) + c(x, y)$ pour π^* presque tout couple (x, y) . On a vu dans la preuve du théorème 4.11 que les fonctions f et $Q_c f$ étaient mesurables et bornées. De plus, pour tout (x, y) , on a $Q_c f(y) - f(x) \leq c(x, y)$. Donc,

$$\begin{aligned} \int c(x, y) \pi^*(dxdy) &= \iint Q_c f(y) - f(x) \pi^*(dxdy) = \int (-f)(x) \mu(dx) + \int Q_c f(y) \nu(dy) \\ &\leq \sup_{(\psi, \varphi) \in \Phi_c(\mu, \nu)} \left\{ \int \psi(x) \mu(dx) + \int \varphi(y) \nu(dy) \right\} \leq \mathcal{T}_c(\mu, \nu). \end{aligned}$$

Comme π^* est optimal, on conclut que toutes les inégalités ci-dessus sont en fait des égalités, ce qui prouve la formule de dualité de Kantorovich. On voit en particulier que le couple $\psi(x) = -f(x)$, $\varphi(y) = Q_c f(y)$ est solution du problème dual.

Traisons à présent le cas où la fonction de coût $c : X \times X \rightarrow [0, \infty[$ est sci. D'après la Proposition 2.4, il existe une suite croissante de fonctions $c_n : X \times X \rightarrow [0, \infty[$ bornées et n -Lipschitz telle que $c = \sup_n c_n$ (par exemple pour la distance $D((x, y), (x', y')) = \max(d(x, x'), d(y, y'))$). D'après le cas traité précédemment, pour tout $n \in \mathbb{N}$, il existe $f_n : X \rightarrow \mathbb{R}$ c_n -convexe et bornée et $\pi_n^* \in \Pi(\mu, \nu)$ optimal pour le coût c_n tels que

$$\int Q_{c_n} f_n(y) \nu(dy) - \int f_n(x) \mu(dx) = \iint c_n(x, y) \pi_n^*(dxdy).$$

On remarque que $(-f_n, Q_{c_n} f_n) \in \Phi_{c_n} \subset \Phi_c$ (on voit facilement que $Q_{c_n} f_n$ et $f_n = P_{c_n} Q_{c_n} f_n$ sont n -Lipschitz) donc, pour tout $n \in \mathbb{N}$ et tout $m \in \mathbb{N}$, $m \leq n$,

$$\sup_{(\psi, \phi) \in \Phi_c} \left\{ \int \psi(x) \mu(dx) + \int \phi(y) \nu(dy) \right\} \geq \iint c_m(x, y) \pi_n^*(dxdy),$$

et par conséquent puisque c_n est une suite croissante, pour tout $m \in \mathbb{N}$, $m \leq n$,

$$\sup_{(\psi, \phi) \in \Phi_c} \left\{ \int \psi(x) \mu(dx) + \int \phi(y) \nu(dy) \right\} \geq \iint c_m(x, y) \pi_n^*(dxdy), \quad (4.4)$$

La suite π_n^* est à valeurs dans $\Pi(\mu, \nu)$ et on a vu dans la preuve du Théorème 2.8 que cet ensemble était compact. On peut donc supposer sans perte de généralité que la suite π_n^* converge étroitement vers un certain $\pi^* \in \Pi(\mu, \nu)$. Puisque c_m est continue, $\lim_{n \rightarrow \infty} \iint c_m(x, y) \pi_n^*(dxdy) = \iint c_m(x, y) \pi^*(dxdy)$. Ainsi, par passage à la limite en n puis en m dans (4.4), on obtient d'après le théorème de convergence monotone

$$\sup_{(\psi, \phi) \in \Phi_c} \left\{ \int \psi(x) \mu(dx) + \int \phi(y) \nu(dy) \right\} \geq \iint c(x, y) \pi^*(dxdy) \geq \mathcal{T}_c(\mu, \nu).$$

Comme on a toujours $\mathcal{T}_c(\mu, \nu) \geq \sup_{(\psi, \varphi) \in \Phi_c(\mu, \nu)} \left\{ \int \psi(x) \mu(dx) + \int \varphi(y) \nu(dy) \right\}$, on conclut que la formule de dualité avec Φ_c ou $\Phi_c(\mu, \nu)$ est vraie. \square

Chapitre 5

Existence d'applications de transport optimal

Le but de ce chapitre est de montrer, pour certaines fonctions de coût, l'existence d'applications de transport optimales pour le problème de Monge-Kantorovich, en particulier le théorème de Brenier (voir Théorème 5.3).

5.1 Préliminaires sur les fonctions convexes

On rappelle qu'une fonction $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est convexe si pour tous $x, y \in \mathbb{R}^n$ et $t \in [0, 1]$

$$\phi((1-t)x + ty) \leq (1-t)\phi(x) + t\phi(y),$$

On rappelle que si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est de classe \mathcal{C}^2 , la convexité de f est équivalente à la positivité de la matrice Hess $f(x)$ en tout point $x \in \mathbb{R}^n$: pour tout $u \in \mathbb{R}^n$, $u^T \cdot \text{Hess } f(x) \cdot u \geq 0$.

Exercice 5.1. Montrer que si $(f_i)_{i \in I}$ est une famille de fonctions convexes sur \mathbb{R}^n alors $f = \sup_i f_i$ est convexe sur \mathbb{R}^n .

Définition 5.1 (Domaine d'une fonction convexe). L'ensemble $\text{dom}(\phi) = \{x \in \mathbb{R}^n : \phi(x) < +\infty\}$ est appelé le domaine de ϕ . On notera Ω l'intérieur du domaine de ϕ .

Dans la preuve du théorème de Brenier, on fera appel au résultat suivant :

Théorème 5.2. Soit $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe. On pose $\Omega_{\text{diff}} = \{x \in \Omega : \phi \text{ est différentiable en } x\}$. Notons λ la mesure de Lebesgue sur \mathbb{R}^n .

- Le bord du domaine de ϕ est négligeable pour la mesure de Lebesgue.
- La fonction ϕ est presque partout différentiable sur son domaine : $\lambda(\text{dom}(\phi) \setminus \Omega_{\text{diff}}) = 0$.
- La fonction ϕ est au-dessus de ses hyperplans tangents : pour tout $x \in \Omega_{\text{diff}}$ et tout $y \in \text{dom}(\phi)$,

$$\phi(y) \geq \phi(x) + \nabla \phi(x) \cdot (y - x).$$

Démonstration. Notons $\partial K = \overline{K} \setminus \text{int}(K)$ le bord de K . Pour le premier point, puisque $\partial K \subset \partial \overline{K}$ on peut se ramener à montrer que le bord d'un convexe fermé est Lebesgue-négligeable. Si K est d'intérieur vide alors il est nécessairement dans un hyperplan donc de mesure nulle. Sinon, puisque la mesure de Lebesgue est invariante par translation, on peut supposer que 0 est dans l'intérieur de K . Pour tout entier $n \geq 1$, l'ensemble $K_n = K \cap B_f(0, n)$ où $B_f(0, n) = \{x \in \mathbb{R}^n, |x| \leq 1\}$, est

un compact convexe tel que $0 \in \text{int}(K)$. Par ailleurs, on vérifie simplement que $\partial K \subset \cup_{n \in \mathbb{N}} \partial K_n$. Puisque $\lambda(\partial K) \leq \sum_{n \in \mathbb{N}} \lambda(\partial K_n)$, il suffit de montrer que le bord d'un compact convexe C tel que $0 \in \text{int}(C)$ est Lebesgue-négligeable.

Pour tout ε , on vérifie facilement que $(1 - \varepsilon)C \subset \text{int}(C)$. En effet, puisque $0 \in \text{int}(C)$, il existe $\eta > 0$ tel que $B(0, \eta) \subset C$ et par convexité pour tout $x \in C$, $B((1 - \varepsilon)x, \varepsilon\eta) = ((1 - \varepsilon)x + \varepsilon B(0, \eta)) \subset C$. On en déduit que $(1 - \varepsilon)^n \lambda(C) \leq \lambda(\text{int}(C)) < \infty$, d'où la conclusion en faisant tendre ε vers 0.

Pour le troisième point, on peut se ramener au cas unidimensionnel en considérant la fonction $\psi : [0, 1] \rightarrow \mathbb{R}$ définie par $\psi(s) = \phi((1 - s)x + sy)$. Cette fonction est convexe, dérivable à droite en 0 de dérivée $\nabla \phi(x) \cdot (y - x)$. Comme ψ est au-dessus de sa tangente,

$$\phi(y) = \psi(1) \geq \psi(0) + \psi'(0)(1 - 0) = \phi(x) + \nabla \phi(x) \cdot (y - x).$$

Il reste donc à montrer le deuxième point. Commençons par traiter le cas de la dimension 1. Notons $I = \text{dom}(\phi)$ (intervalle de \mathbb{R}). On sait que la convexité d'une fonction est équivalente à la croissance des pentes : pour tout $x \in I$, la fonction $p_x(y) = \frac{\phi(x) - \phi(y)}{x - y}$ est croissante sur son ensemble de définition. En tout point à l'intérieur de I , la fonction p_x admet donc une limite à gauche et à droite finie. En effet, si x est dans l'intérieur de I , alors pour tous $y < x < z$, on a $p_x(y) \leq p_x(z)$, et donc $\{p_x(y) : y < x\}$ est majoré, donc $\lim_{y \rightarrow x^-} p_x(y) = \sup\{p_x(y) : y < x\} \leq p_x(x) < +\infty$. On voit de même que $\lim_{y \rightarrow x^+} p_x(y) = \inf\{p_x(z) : z > x\} > -\infty$. En conclusion en tout point x à l'intérieur de I , $\phi'_-(x)$ et $\phi'_+(x)$ sont finies et de plus $\phi'_-(x) \leq \phi'_+(x)$. Par ailleurs, les fonction ϕ'_\pm sont croissantes sur l'intérieur de I . En effet, si $x_1 < x_2$ et $z \in]x_1, x_2[$

$$\phi'_+(x_1) \leq \frac{\phi(z) - \phi(x_1)}{z - x_1} = p_z(x_1) \leq p_z(x_2) = p_{x_2}(z).$$

Donc en prenant la limite quand $z \rightarrow x_2$ par valeurs inférieures, on trouve que $\phi'_+(x_1) \leq \phi'_-(x_2)$.

Nous allons montrer maintenant que ϕ est dérivable sur l'intérieur de I sauf peut être en une infinité dénombrable de points. Pour cela, nous allons montrer que l'ensemble J des points x à l'intérieur de I tels que $\phi'_-(x) < \phi'_+(x)$ est au plus dénombrable. Pour tout $x \in J$, on choisit un rationnel r_x dans l'intervalle $]\phi'_-(x), \phi'_+(x)[$. Si $x_1 < x_2$ sont dans J , alors $\phi'_+(x_1) \leq \phi'_-(x_2)$ et donc les intervalles $]\phi'_-(x_1), \phi'_+(x_1)[$ et $]\phi'_-(x_2), \phi'_+(x_2)[$ sont disjoints. On en déduit que $r_{x_1} < r_{x_2}$. On a ainsi construit une application injective de J dans \mathbb{Q} , ce qui prouve que J est au plus dénombrable.

A présent traitons le cas de la dimension n . Commençons par montrer que si $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction convexe, alors elle admet des dérivées partielles $\partial_i \phi(x)$ $i = 1, \dots, n$, en Lebesgue presque tout point de $\Omega = \text{int}(\text{dom}(\phi))$. On pose $V = \{v \in \mathbb{R}^{n-1}; \exists w \in \mathbb{R}, (v, w) \in \Omega\}$ et, pour tout $v \in V$, on note $I_v = \{w \in \mathbb{R} : (v, w) \in \Omega\}$ et on pose $\phi_v(w) = \phi(v_1, \dots, v_{n-1}, w)$ définie sur l'intervalle I_v . D'après le cas précédent, pour chaque $v \in V$, il existe un ensemble au plus dénombrable $N_v \subset I_v$ tel que ϕ_v soit dérivable en tout point de $I_v \setminus N_v$. Notons $N^n = \{(v, w) : v \in V, w \in N_v\}$. D'après le théorème de Fubini,

$$|N^n| = \int_V |N_v| dv = 0,$$

puisque étant dénombrable, N_v est de mesure de Lebesgue nulle pour tout $v \in V$. Ainsi, en tout point x de $\Omega \setminus N^n$, la fonction ϕ admet une dérivée partielle $\partial_n \phi(x)$. En échangeant le rôle des coordonnées, on montre de même que pour tout $i \in \{1, \dots, n\}$, il existe un ensemble négligeable, N^i tel que ϕ admet une dérivée partielle $\partial_i \phi(x)$ pour tout $x \in \Omega \setminus N^i$. En posant $N = N^1 \cup \dots \cup N^n$ on obtient un ensemble négligeable en dehors duquel ϕ admet des dérivées partielles par rapport à chacune des coordonnées.

Montrons pour finir que ϕ est différentiable en tout point de $\Omega_{\text{diff}} := \Omega \setminus N$. Prenons $x \in \Omega_{\text{diff}}$, et considérons

$$\varepsilon_x(y) = \phi(y) - \phi(x) - \sum_{i=1}^n \partial_i f(x)(y_i - x_i), \quad \forall y \in \Omega.$$

Le but est de montrer que $\frac{\varepsilon_x(y)}{\|y-x\|_1} \rightarrow 0$ lorsque $y \rightarrow x$ (on utilise la norme $\|u\|_1 = \sum_{i=1}^n |u_i|$). Soit $r = \|y-x\|_1$. On note B_1 la boule unité fermée pour la norme $\|\cdot\|_1$. La fonction ε_x étant convexe, elle atteint son maximum sur le convexe $x + rB_1$ en l'un de ses $2n$ points extrémaux. On a donc

$$\frac{1}{r} \max_{\|y-x\|_1=r} \varepsilon_x(y) = \frac{1}{r} \max_i \varepsilon_x(x \pm r e_i) := \alpha(r),$$

en notant e_1, \dots, e_n la base canonique de \mathbb{R}^n . Pour tout i , $\varepsilon_x(x \pm r e_i)/r \rightarrow 0$, par définition des dérivées partielles $\partial_i f(x)$. Donc $\alpha(r) \rightarrow 0$ lorsque $r \rightarrow 0$. Ainsi, $\varepsilon_x(y) \leq \|y-x\|_1 \alpha(\|y-x\|_1)$, avec α ayant une limite nulle en 0. Pour conclure, il suffit de montrer que ε_x est positive. Remarquons pour cela que pour tout $h \in \mathbb{R}^n$, et $t \in \mathbb{R}$, on a $\varepsilon_x(x+th) \leq |t| \|h\|_1 \alpha(|t| \|h\|_1)$. Donc $\limsup_{t \rightarrow 0^+} \varepsilon_x(x+th)/t \leq 0$ et de même $\liminf_{t \rightarrow 0^-} \varepsilon_x(x+th)/t \geq 0$. Comme la fonction $t \mapsto \varepsilon_x(x+th)$ est convexe et nulle en 0, la fonction $t \mapsto \varepsilon_x(x+th)/t$ est croissante. On en déduit que $\lim_{t \rightarrow 0} \varepsilon_x(x+th)/t = 0$. On a donc pour tout $t > 0$, $\varepsilon_x(x+th) \geq 0$, ce qui achève la preuve. \square

5.2 Le théorème de Brenier

On se place dans $\mathcal{X} = \mathbb{R}^n$ muni de sa norme euclidienne standard notée $|\cdot|$ et on considère la fonction de coût $c_2(x, y) = \frac{1}{2}|x-y|^2$, $x, y \in \mathbb{R}^n$. Pour toutes $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$, on note $\mathcal{T}_2(\mu, \nu)$ le coût de transport optimal associé à ce coût quadratique c_2 .

Théorème 5.3 (Brenier). *Soient μ, ν deux mesures de probabilité sur \mathbb{R}^n équipé de sa norme euclidienne canonique; si μ est absolument continue par rapport à la mesure de Lebesgue et si $\mathcal{T}_2(\mu, \nu) < \infty$ alors il existe un unique plan de transport optimal $\pi^* \in \Pi(\mu, \nu)$. Ce plan de transport optimal est déterministe : il existe une application de transport T^* de μ sur ν telle que $\pi^*(dx dy) = \mu(dx) \delta_{T^*(x)}(dy)$.*

De plus, cette application T^ s'écrit sous la forme $T^* = \nabla \phi$, où $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction convexe telle que $\mu(\phi < \infty) = 1$, et elle est essentiellement unique : si T est une autre solution optimale du problème de Monge alors $T = T^*$ μ -presque sûrement.*

Commençons par remarquer que si T_1 et T_2 sont deux applications telles que $T_1(x) = T_2(x)$ pour μ presque tout x , alors $T_{1\#}\mu = T_{2\#}\mu$ car pour toute fonction test f ,

$$\int f(T_1(x)) \mu(dx) = \int_{T_1=T_2} f(T_1(x)) \mu(dx) = \int_{T_1=T_2} f(T_2(x)) \mu(dx) = \int f(T_2(x)) \mu(dx).$$

On utilise cette remarque pour donner un sens à $T_{\#}\mu$ lorsque T n'est définie que μ -presque partout : on étend T de manière arbitraire à l'espace tout entier et la remarque faite juste au-dessus montre que la mesure image ne dépend pas du prolongement choisi.

Démonstration. Le point le plus délicat est la preuve d'existence.

Existence d'un plan gradient. D'après le Théorème 4.11, il existe une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ c_2 -convexe et un couplage $\pi^* \in \Pi(\mu, \nu)$ optimal telle que le support Γ de π^* est contenu

dans $\partial_{c_2} f$. De plus la fonction f est finie sur $\Gamma_1 = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n (x, y) \in \Gamma\}$ et $\mu(\Gamma_1) = 1$. Par définition des fonctions c_2 -convexes, il existe une fonction $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ telle que

$$f(x) = \sup_{y \in \mathbb{R}^n} \left\{ g(y) - \frac{1}{2}|x - y|^2 \right\} = \sup_{y \in \mathbb{R}^n} \left\{ g(y) - \frac{1}{2}|y|^2 + x \cdot y \right\} - \frac{1}{2}|x|^2.$$

La fonction $\phi(x) := f(x) + \frac{1}{2}|x|^2$ est donc un sup de fonctions affines et est donc convexe (voir Exercice 5.1).

Par définition du c_2 sous-différentiel de f en $x \in \mathbb{R}^n$, on a $y \in \partial_{c_2} f(x)$ si et seulement si

$$f(z) \geq f(x) + c_2(x, y) - c_2(z, y), \quad \forall z \in \mathbb{R}^n.$$

Autrement dit, en développant les carrés,

$$f(z) \geq f(x) + \frac{1}{2}|x|^2 - \frac{1}{2}|z|^2 + (z - x) \cdot y, \quad \forall z \in \mathbb{R}^n,$$

ce qui se traduit par

$$\phi(z) \geq \phi(x) + (z - x) \cdot y, \quad \forall z \in \mathbb{R}^n.$$

D'après le Théorème 5.2, la fonction ϕ est différentiable pour tout x dans un ensemble Ω_{diff} de complémentaire Lebesgue-négligeable. Si $x \in \Omega_{\text{diff}}$, alors en prenant $z = x + th$ avec $h \in \mathbb{R}^n$, on obtient

$$\nabla \phi(x) \cdot h = \lim_{t \rightarrow 0} \frac{\phi(x + th) - \phi(x)}{t} \geq h \cdot y.$$

On a donc $(\nabla \phi(x) - y) \cdot h \geq 0$ pour tout $h \in \mathbb{R}^n$, et donc $\nabla \phi(x) = y$. Ainsi pour tout $x \in \Omega_{\text{diff}}$, $\partial_{c_2} f(x) \subset \{\nabla \phi(x)\}$. L'inclusion réciproque est assurée puisque ϕ est au-dessus de son hyperplan tangent (Théorème 5.2).

Comme μ est absolument continue par rapport à Lebesgue, on a $\mu(\text{dom}(\phi)) = \mu(\Omega_{\text{diff}})$. Comme $1 = \mu(\Gamma_1) \leq \mu(\text{dom} \phi)$ on en déduit $\mu(\Omega_{\text{diff}}) = 1$. Pour π^* presque tout (x, y) , on a donc simultanément $x \in \Omega_{\text{diff}}$ et $(x, y) \in \partial_{c_2} f$, donc $y = \nabla \phi(x)$. Autrement dit

$$\pi^* (\{(x, \nabla \phi(x)) : x \in \Omega_{\text{diff}}\}) = 1.$$

Si on prend une fonction test $h : \mathbb{R}^n \rightarrow \mathbb{R}$ continue bornée, on a donc

$$\int h(x, y) \pi^*(dxdy) = \int h(x, \nabla \phi(x)) \pi^*(dxdy) = \int h(x, \nabla \phi(x)) \mu(dx),$$

ce qui prouve que $\pi^*(dxdy) = \mu(dx) \delta_{\nabla \phi(x)}(dy)$.

Unicité du couplage optimal. Supposons que π_1 et π_2 sont deux couplages optimaux et considérons $\pi_3 = \frac{1}{2}\pi_1 + \frac{1}{2}\pi_2$. Il est facile de vérifier que le couplage π_3 est encore optimal. Mais, d'après ce qui précède, il existe des applications de transport telles que $\pi_1(dxdy) = \mu(dx) \delta_{T_1(x)}(dy)$, $\pi_2(dxdy) = \mu(dx) \delta_{T_2(x)}(dy)$ et $\pi_3(dxdy) = \mu(dx) \delta_{T_3(x)}(dy)$. Par conséquent,

$$\pi_3(dxdy) = \mu(dx) \left(\frac{1}{2} \delta_{T_1(x)}(dy) + \frac{1}{2} \delta_{T_2(x)}(dy) \right) = \mu(dx) \delta_{T_3(x)}(dy).$$

En intégrant la fonction indicatrice de l'ensemble $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n, y \neq T_3(x)\}$, cela impose $T_1(x) = T_2(x) = T_3(x)$ pour μ presque tout x , et *a fortiori* $\pi_1 = \pi_2$. \square

5.3 Le cas des coûts strictement convexes

On considère dans cette partie un coût de transport $c : \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, +\infty[$ de la forme

$$c(x, y) = \theta(y - x), \quad x, y \in \mathbb{R}^n,$$

où dans tout ce qui suit $\theta : \mathbb{R}^n \rightarrow [0, +\infty[$ est une fonction convexe de classe \mathcal{C}^1 .

On supposera dans toute la suite que θ est même *strictement convexe*, c'est-à-dire que si $x \neq y$, alors

$$\theta((1-t)x + ty) < (1-t)\theta(x) + t\theta(y), \quad \forall t \in]0, 1[.$$

Autrement dit la fonction θ n'est affine sur aucun segment de \mathbb{R}^n .

Cette hypothèse entraîne que l'application $\nabla\theta : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est injective. On peut donc définir $(\nabla\theta)^{-1}$ sur $\nabla\theta(\mathbb{R}^n)$.

Proposition 5.4. *Si $\theta : \mathbb{R}^n \rightarrow \mathbb{R}$ est strictement convexe, alors $\nabla\theta$ est injective.*

Démonstration. Supposons que $\nabla\theta(x_1) = \nabla\theta(x_2)$ pour $x_1 \neq x_2$. La fonction

$$\varphi(t) = \theta((1-t)x_1 + tx_2), \quad t \in [0, 1]$$

est convexe sur $[0, 1]$. Sa dérivée $\varphi'(t) = \nabla\theta((1-t)x_1 + tx_2) \cdot (x_2 - x_1)$ est donc croissante. Comme $\varphi'(0) = \varphi'(1)$ la fonction φ' est en fait constante sur $[0, 1]$. Par conséquent, $\varphi(t) = \varphi(0) + t\varphi'(0)$ et donc f est affine sur le segment $[x_1, x_2]$. \square

Théorème 5.5. *Soient μ, ν deux mesures de probabilité sur \mathbb{R}^n . On note O l'intérieur du support de μ et on suppose que*

— μ est absolument continue par rapport à Lebesgue et $\mu(\partial O) = 0$,

et

— ν est à support compact.

Si $\mathcal{T}_c(\mu, \nu) < \infty$, alors il existe un unique plan de transport optimal $\pi^* \in \Pi(\mu, \nu)$. Ce plan de transport optimal est déterministe : il existe une application de transport T^* de μ sur ν , unique μ presque sûrement, telle que $\pi^*(dx dy) = \mu(dx) \delta_{T^*(x)}(dy)$. De plus, cette application T^* s'écrit sous la forme

$$T^*(x) = x + (\nabla\theta)^{-1}(\nabla f(x)),$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction c -convexe telle que le support de π^* soit inclus dans $\partial_c f$.

Remarque 5.1. *Pour certaines fonctions de coût, les hypothèses du théorème précédent peuvent être affaiblies. Par exemple, pour les fonctions de coûts $c_p(x, y) = \frac{1}{p}|x - y|^p$, $p > 1$ (où $|\cdot|$ est la norme euclidienne) la conclusion du théorème précédent reste valable sous la seule hypothèse que $\mathcal{T}_{c_p}(\mu, \nu) < +\infty$. On pourra consulter par exemple l'article de Gangbo et Mc Cann (The geometry of optimal transportation, *Acta Mathematica* 177 (1996), 113-161).*

Démonstration. D'après le Théorème 4.11, il existe une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ c -convexe et un couplage $\pi^* \in \Pi(\mu, \nu)$ optimal tels que le support Γ de π^* soit contenu dans $\partial_c f$. De plus f est finie sur $\Gamma_1 = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n : (x, y) \in \Gamma\}$. On a toujours $\mu(\Gamma_1) = 1$. De plus comme $\Gamma \subset \text{Supp}(\mu) \times \text{Supp}(\nu)$ et que $\text{Supp}(\nu)$ est compact on voit facilement que Γ_1 est fermé (prendre une suite à valeurs dans $\Gamma_1 \dots$) et par conséquent $\Gamma_1 = \text{Supp}(\mu)$.

Pour tout $x \in \text{Supp}(\mu)$, il existe donc $y \in \text{Supp}(\nu)$ tel que

$$Q_c f(y) = f(x) + \theta(y - x).$$

Comme $Q_c f(y) = \inf_{z \in \mathbb{R}^n} \{f(z) + \theta(y - z)\}$, on en déduit que l'application $z \mapsto f(z) + \theta(y - z)$ atteint son minimum en x . Admettons un instant que f est différentiable en μ -presque tout point de O et notons O_{dif} l'ensemble de ces points de différentiabilité. On conclut alors que pour tout $x \in O_{dif}$, il existe $y \in \text{Supp}(\nu)$ tel que

$$\nabla f(x) = \nabla \theta(y - x),$$

ce qui équivaut à

$$y = x + (\nabla \theta)^{-1}(\nabla f(x)).$$

Autrement dit $\partial_c f(x) = \{T^*(x)\}$ pour tout $x \in O_{dif}$. Comme par hypothèse $\mu(\partial O) = 0$, on conclut que $\mu(O_{dif}) = 1$. En raisonnant comme dans la preuve du théorème de Brenier, on conclut que $\pi^*(dx dy) = \mu(dx) \delta_{T^*(x)}(dy)$ et T^* transporte μ sur ν . L'unicité de T^* s'obtient de la même manière.

Il reste à montrer que f est bien différentiable en μ -presque tout point de O . Etant c -convexe, f vérifie

$$f(x) = \sup_{y \in \mathbb{R}^n} \{Q_c f(y) - c(x, y)\}, \quad \forall x \in \mathbb{R}^n.$$

Pour tout $x \in \text{Supp}(\mu)$, il existe $y \in \text{Supp}(\nu) := K$ tel que $f(x) = Q_c f(y) - c(x, y)$. Donc, si l'on pose

$$g(x) := \sup_{y \in K} \{Q_c f(y) - c(x, y)\}, \quad x \in \mathbb{R}^n,$$

on obtient une fonction (c -convexe) vérifiant $g \leq f$ sur \mathbb{R}^n et $g = f$ sur $\text{Supp}(\mu)$. Admettons un instant que g est différentiable en Lebesgue presque tout point de l'intérieur de $\text{dom}(g) = \{x \in \mathbb{R}^n : g(x) < \infty\}$. Comme μ est absolument continue par rapport à Lebesgue et $\text{Supp}(\mu) \subset \text{dom}(f)$ on a $O \subset \text{int dom}(g)$ et comme $f = g$ sur O , on conclut que f est différentiable en Lebesgue presque tout point de O .

Pour finir, montrons que g est différentiable en Lebesgue presque tout point de l'intérieur de son domaine. Nous donnerons de ce résultat deux preuves. La première impose une condition de régularité un peu plus forte à la fonction θ et utilise le théorème 5.2 de différentiabilité presque sûr des fonctions convexes. La deuxième preuve utilise moins de régularité, mais fait appel à un théorème de Rademacher (que nous admettrons) qui montre que toute fonction localement Lipschitz est différentiable presque sûrement.

Première preuve : Nous supposons dans ce qui suit que θ est de classe \mathcal{C}^2 . Soit a un point de l'intérieur du domaine de g et $r_a > 0$ tel que la boule euclidienne fermée $\overline{B}(a, r_a)$ soit contenue dans $\text{dom}(g)$. Posons

$$\lambda_a = \sup \{u^T \cdot \text{Hess} \theta(y - x) \cdot u : |u| = 1, y \in K, |x - a| \leq r_a\}.$$

Autrement dit λ_a est un majorant de la plus grande valeur propre de la matrice hessienne de θ sur le compact $K - \overline{B}(a, r_a)$. Posons $\varphi(y) = Q_c f(y)$ et écrivons

$$g(x) + \lambda_a \frac{|x|^2}{2} = \sup_{y \in K} \left\{ \varphi(y) + \lambda_a \frac{|x|^2}{2} - \theta(y - x) \right\}$$

Par construction, pour tout $y \in K$, la fonction $x \mapsto \psi(x, y) = \varphi(y) + \lambda_a \frac{|x|^2}{2} - \theta(y - x)$ est convexe sur $\overline{B}(a, r_a)$. On en déduit facilement que $x \mapsto \sup_{y \in K} \psi(x, y)$ est encore convexe sur $\overline{B}(a, r_a)$ (voir Exercice 5.1). Ainsi la fonction $x \mapsto g(x) + \lambda_a \frac{|x|^2}{2}$ est convexe sur $\overline{B}(a, r_a)$, ce qui d'après le théorème 5.2 entraîne que cette fonction, et par conséquent la fonction g elle-même, est différentiable en Lebesgue presque tout point de $B(a, r_a)$. En se restreignant aux boules centrées en des points à

coordonnées rationnelles et de rayons rationnels, on conclut sans peine que g est différentiable en Lebesgue presque tout point de l'intérieur de son domaine.

Deuxième preuve : Prenons de nouveau a tel que $\overline{B}(a, r_a) \subset \text{dom}(g)$ et posons cette fois

$$L = \sup\{|\nabla\theta(x - y)| : y \in K, |x - a| \leq r_a\}.$$

Pour tout $y \in K$, la fonction $x \mapsto \psi(x, y) = \varphi(y) - \theta(y - x)$ est cette fois L -Lipschitz sur $\overline{B}(a, r_a)$. Comme un supremum de fonctions L -Lipschitz reste L -Lipschitz (le vérifier), on conclut que g est L -Lipschitz sur $\overline{B}(a, r_a)$. On conclut grâce au théorème de Rademacher énoncé ci-dessous. \square

Théorème 5.6 (Rademacher). *Si $f : U \rightarrow \mathbb{R}$ est une fonction Lipschitz sur un ouvert U de \mathbb{R}^n alors f est dérivable en Lebesgue presque tout point de U .*

Démonstration. La preuve est du même style que celle donnée pour le Théorème 5.2 mais est beaucoup plus délicate. On la trouve par exemple dans le livre de C. Villani (Optimal-Transport : Old and New - Theorem 10.8, p. 222). \square

Chapitre 6

Algorithmique du transport optimal

6.1 Diverses solutions

Le transport optimal, en plus de ses utilisations théoriques, dont certaines seront vues plus bas, a également des « applications appliquées » dans de très nombreux domaines : né du problème d'ingénierie de Monge, il retourne volontiers au monde « concret »... Il est en particulier utilisé pour diverses tâches de traitement d'images ; citons :

- le transfert de couleurs (*color transfer*), où l'on veut modifier une image source pour que ses couleurs ressemblent à celles d'une image cible (pour uniformiser des portraits, rendre des couchers de soleil plus rouges, faire ressembler n'importe quelle photo à un Van Gogh...)
- la génération aléatoire de textures à partir d'échantillons.

L'utilisation pratique du transport suppose de pouvoir calculer, ou au moins approcher, les coûts de transport et les couplages optimaux.

Problème informel

Étant donné : Deux mesures μ et ν , le coût c

minimiser : $\iint c(x, y) d\pi(x, y)$

sous la contrainte (s.l.c.) : π a pour marginales μ et ν

On peut dénombrer (au moins !) trois pistes possibles pour résoudre le problème.

1. Discrétiser les données : si μ et ν sont des sommes finies de masses de Dirac, on peut espérer traiter explicitement le problème.
2. Utiliser les propriétés de l'application de transport optimal : on verra plus loin que dans le cas quadratique, T^* vérifie une ÉDP — l'équation de Monge-Ampère — que l'on peut résoudre numériquement.
3. Toujours dans le cas quadratique, voyons $\mu = \rho_0$ comme une densité de particules dans un fluide, au temps $t = 0$, et $\nu = \rho_1$ comme une densité finale au temps 1. Supposons que le fluide bouge avec un champ de vitesses v_t : la densité ρ évolue suivant l'équation de transport

$$\partial_t \rho_t + \nabla_x(\rho_t v_t) = 0.$$

On peut alors montrer que le transport optimal entre μ et ν peut être construit en cherchant les champs v_t qui amènent μ sur ν en minimisant l'énergie cinétique totale

$$\int_0^1 \int_{\mathbb{R}^n} \rho_t(x) |v_t(x)|^2 dx dt.$$

Ce point de vue « Eulérien » de mécanique des fluides, introduit par Benamou et Brenier [BB99] pour fournir une méthode numérique de résolution du problème, a ouvert de larges avenues théoriques. En pratique, un bon choix de variables rend le problème convexe et permet sa résolution numérique efficace.

Dans la suite de ce chapitre nous nous concentrerons sur la première piste ; nous donnons en fin de chapitre des pointeurs bibliographiques pour en savoir plus sur les autres méthodes et les applications.

6.2 Différents problèmes

Nous abandonnons donc le problème général pour nous restreindre à la question suivante.

Transport discret

Étant donné : Deux mesures discrètes $\mu = \sum_{i=1}^m p_i \delta_{a_i}$ et $\nu = \sum_{j=1}^n \delta_{b_j}$; le coût c
 minimiser : $\iint c(x, y) d\pi(x, y)$
 s. l. c. : π a pour marginales μ et ν

Remarque 6.1. *Nous avons en partie repoussé le problème : si l'on sait résoudre le transport discret, peut-on résoudre approximativement le problème général ? Nous donnerons quelques éléments de réponse plus loin.*

Nous allons réécrire ce problème sous une forme légèrement différente. Notons $A = \{a_1, \dots, a_m\}$ et $B = \{b_1, \dots, b_n\}$. Comme π a pour marginales μ et ν ,

$$\pi(A \times X) = \mu(A) = 1, \quad \pi(X \times B) = \nu(B) = 1,$$

donc $\pi(A \times B) = \pi((A \times X) \cap (X \times B)) = 1$. Par conséquent $\pi \in \Pi(\mu, \nu)$ est discrète, et est entièrement caractérisée par la famille $\mathbf{x} = (x_{ij})_{ij} = (\pi(\{a_i\} \times \{b_j\}))_{i,j}$.

Dans ce cas discret, un plan de transport correspond donc à une matrice \mathbf{x} de taille $m \times n$. Les conditions de marginales se lisent sur les... marges de la matrice, en faisant les sommes en ligne et en colonne :

$$\forall i, \quad \sum_{j=1}^n x_{ij} = p_i, \quad \forall j, \quad \sum_{i=1}^m x_{ij} = q_j.$$

Donnons deux exemples très simples, pour $\mu = \frac{1}{4}\delta_1 + \frac{1}{2}\delta_2 + \frac{1}{4}\delta_3 = \frac{1}{8}(2\delta_1 + 4\delta_2 + 2\delta_3)$ et $\nu = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2 = \frac{1}{8}(4\delta_1 + 4\delta_2)$. En prenant $1/8$ comme unité, les deux matrices suivantes correspondent à des couplages de μ et ν , que l'on retrouve respectivement comme sommes en ligne et en colonne.

2	4	4	2	4	4
4	1	1	4	2	0
2	2	2	2	2	2
2	1	1	2	0	2

Le premier couplage correspond à des variables indépendantes ($\pi = \mu \otimes \nu, x_{ij} = p_i q_j$). Le second n'est autre que le couplage croissant.

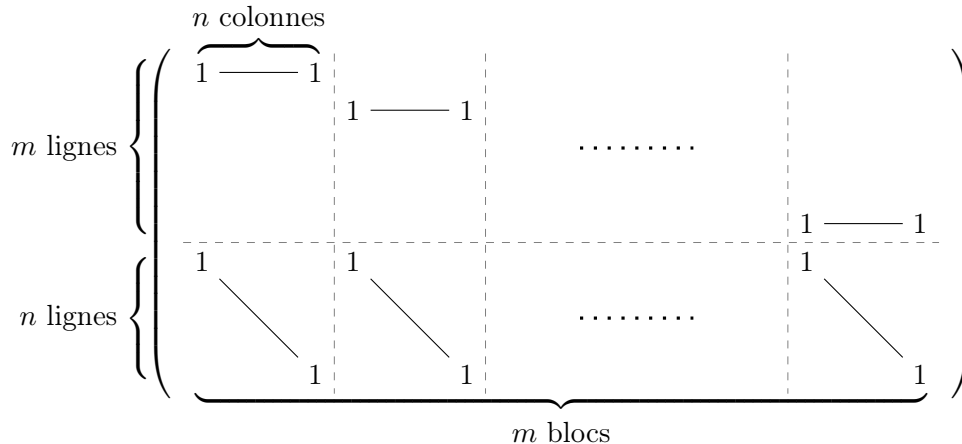
Pour tout (i, j) , notons $c_{ij} = c(a_i, b_j)$ et notons \mathbf{c} la matrice des c_{ij} . Le problème se réécrit de la manière suivante :

Transport discret v.2.0

Étant donné : Deux vecteurs de probabilités $(p_1, \dots, p_m), (q_1, \dots, q_n)$; la matrice de coûts \mathbf{c}

$$\begin{aligned} \text{minimiser : } & \sum_{i,j} x_{ij} c_{ij} \\ \text{s. l. c. : } & \begin{cases} \forall i, \sum_j x_{ij} = p_i \\ \forall j, \sum_i x_{ij} = q_j \\ \forall i, j, x_{ij} \geq 0. \end{cases} \end{aligned}$$

Récrivons ce problème sous une dernière forme. On identifie \mathbf{c} et le vecteur colonne à mn éléments $(c_{11}, c_{12}, \dots, c_{1n}, c_{21}, \dots, c_{m1}, \dots, c_{mn})^\top$, et on fait de même pour \mathbf{x} . On définit le vecteur \mathbf{b} de taille $m+n$: $\mathbf{b} = (p_1, \dots, p_m, q_1, \dots, q_n)^\top$, et la matrice A de taille $(m+n) \times (mn)$:



On peut alors écrire une dernière reformulation du problème, entièrement sous forme matricielle.

Transport discret v.3.0

Étant donné : Le vecteur $\mathbf{b} = (p_1, \dots, p_m, q_1, \dots, q_n)^\top$; la matrice de coûts \mathbf{c}

$$\begin{aligned} \text{minimiser : } & \mathbf{c}^\top \mathbf{x} \\ \text{s. l. c. : } & \begin{cases} \mathbf{Ax} = \mathbf{b}, \\ \forall i, j, x_{ij} \geq 0. \end{cases} \end{aligned}$$

On a donc affaire à un problème de *minimisation* d'une fonction « objectif » *linéaire* (la forme linéaire sur \mathbb{R}^{mn} : $\mathbf{x} \mapsto \mathbf{c}^\top \mathbf{x}$), sous des contraintes *linéaires*. Ce type de problème est connu sous le nom d'*optimisation linéaire* ou *programmation linéaire* et est très bien étudié; la littérature qui lui est consacrée est gigantesque et de nombreux algorithmes efficaces ont été développés depuis les

années 1940 pour le résoudre (avec des applications aussi variées qu'optimiser la production en temps de guerre ou chercher le régime alimentaire le moins coûteux respectant des apports nutritionnels recommandés...).

Définition 6.1 (Terminologie). *Si un vecteur \mathbf{x} respecte les contraintes du problème il sera dit faisable. On dira qu'il est optimal si il réalise le minimum de la fonction $\mathbf{c}^\top \mathbf{x}$. Ce minimum est appelé valeur du problème.*

Enfin, nous considérerons deux problèmes légèrement différents. Le premier est le cas particulier du problème de transport quand les mesures de départ et d'arrivée sont de la forme $\mu = \frac{1}{N} \sum_{i=1}^N \delta_{a_i}$ et $\nu = \frac{1}{N} \sum_{j=1}^N \delta_{b_j}$. Pour alléger les notations, on remplace le \mathbf{x} précédent par $N\mathbf{x}$ et l'on obtient le problème suivant, dont le nom sera expliqué plus loin.

LSAP relâché

Étant donné : La matrice de coûts \mathbf{c}

minimiser : $\mathbf{c}^\top \mathbf{x} = \sum_{ij} c_{ij} x_{ij}$

$$\text{s. l. c. : } \begin{cases} \forall i, \sum_j x_{ij} = 1 \\ \forall j, \sum_i x_{ij} = 1 \\ \forall i, j, x_{ij} \geq 0. \end{cases}$$

Notre dernier problème naît là encore des applications : si les mesures μ et ν à transporter représentent des quantités d'objets (disons des palettes de yaourts à transporter entre les usines situées en a_i et les points de vente situés en b_j , les a_i pouvant se répéter quand une usine produit plus d'une palette, idem pour les b_j), on veut éviter de couper les (palettes de) yaourts en morceaux. On cherche donc à *assigner* à chaque site de production i un (et un seul) site de livraison j , ce qui revient à poser $x_{ij} = 1$ et $x_{il} = 0$ pour $l \neq j$. On arrive au problème suivant.

Linear Sum Assignment Problem (LSAP)

Étant donné : La matrice de coûts \mathbf{c}

minimiser : $\mathbf{c}^\top \mathbf{x} = \sum_{ij} c_{ij} x_{ij}$

$$\text{s. l. c. : } \begin{cases} \forall i, \sum_j x_{ij} = 1 \\ \forall j, \sum_i x_{ij} = 1 \\ \forall i, j, x_{ij} \in \{0, 1\}. \end{cases}$$

Remarque 6.2 (Liens entre les problèmes).

- *LSAP relâché est un cas particulier du transport discret, qui est lui-même un problème de programmation linéaire.*
- *Si on sait résoudre LSAP relâché, on peut résoudre approximativement le transport discret (exactement si les poids sont rationnels).*
- *Nous expliciterons plus loin le lien entre LSAP et LSAP relâché.*

6.3 Dualité de Kantorovich et « lâcheté complémentaire »

Considérons le problème LSAP relâché dans son interprétation économique. On dispose de N palettes de yaourts que l'on doit envoyer dans N endroits; l'envoi de la palette i à l'endroit j coûte c_{ij} .

Un transporteur arrive et propose le marché suivant : il s'occupe de tout le transport, et fait payer $\phi(i)$ pour enlever le yaourt i et $\psi(j)$ pour déposer un yaourt au site j . Pour rendre son offre intéressante il garantit que, pour tout couple (i, j) ,

$$\phi(i) + \psi(j) \leq c_{ij}.$$

Par conséquent on ne perd pas d'argent en laissant le transporteur s'occuper du yaourt i : quel que soit l'endroit où l'on voulait l'envoyer, le transporteur le fait pour moins cher (on détaillera ce point plus bas).

On dira que le couple (ϕ, ψ) est *faisable* s'il vérifie ces inégalités. L'objectif du transporteur est de maximiser son profit, ce qui correspond à répondre au problème suivant :

LSAP *

Étant donné : la matrice des coûts \mathbf{c}

maximiser : $\sum_i \phi(i) + \sum_j \psi(j)$

s. l. c. : $\forall i, j, \quad \phi(i) + \psi(j) \leq c_{ij}$

Ce problème de maximisation est appelé « problème dual » du « problème primal » LSAP (relâché); sa *valeur* est la borne supérieure de $\sum_i \phi(i) + \sum_j \psi(j)$ sur tous les couples (ϕ, ψ) faisables.

Pour comparer les valeurs et les solutions optimales du problème primal et du problème dual, définissons d'abord une notion. À tout couple (i, j) correspond une contrainte de positivité pour \mathbf{x} ($x_{ij} \geq 0$) et une contrainte $\phi(i) + \psi(j) \leq c_{ij}$ pour (ϕ, ψ) . Ces contraintes peuvent être « lâches » (si elles sont vérifiées avec une inégalité stricte; il y a « du mou ») ou « saturées » (si elles sont vérifiées avec égalité, on dit aussi que la contrainte est *tendue*). La « lâcheté complémentaire » dit que ces deux contraintes ne peuvent pas être « lâches » en même temps :

Définition 6.2 (Lâcheté complémentaire). *Si \mathbf{x} est faisable pour le problème primal, et (ϕ, ψ) faisable pour le dual, on dit que \mathbf{x} et (ϕ, ψ) vérifient la condition de lâcheté complémentaire (ou d'écart complémentaire, en anglais complementary slackness) si, pour tout i, j ,*

$$OU \left\{ \begin{array}{l} x_{ij} = 0 \\ \phi(i) + \psi(j) = c_{ij}. \end{array} \right.$$

Théorème 6.3 (Dualité). *Soit \mathbf{x} et (ϕ, ψ) des vecteurs faisables respectivement pour le problème primal et le problème dual. Alors*

$$\mathbf{c}^\top \mathbf{x} \geq \sum_i \phi(i) + \sum_j \psi(j). \quad (6.1)$$

Les conditions suivantes sont équivalentes :

- i) il y a égalité dans (6.1);
- ii) \mathbf{x} et (ϕ, ψ) sont de « lâcheté complémentaire ».

Si elles sont vérifiées alors \mathbf{x} est une solution optimale du problème primal et (ϕ, ψ) est une solution optimale du problème dual.

Enfin, pour toute solution optimale \mathbf{x} , il existe une solution duale (ϕ, ψ) vérifiant ces conditions ; les deux problèmes ont donc même valeur.

Le premier point confirme l'intuition plus haut : il est toujours avantageux de faire appel au transporteur. Le second donne une condition suffisante d'optimalité, le dernier montre qu'elle est d'une certaine manière nécessaire.

Remarque 6.3. Ce théorème est en fait un cas particulier de la dualité de Kantorovich vue plus haut (Théorèmes 4.11 et 4.12). Pour le voir, rappelons d'abord que si (ϕ, ψ) est optimale pour le problème dual, alors $(Q_c(-\psi), \psi)$ l'est aussi, ainsi que $(Q_c(-\psi), Q_c(-Q_c(-\psi)))$, où Q_c est l'inf-convolution

$$Q_c f(i) = \min_j c_{ij} + f(j).$$

En posant f la fonction c -convexe $f = P_c \phi = -Q_c(-\psi)$, on a donc un couple optimal $(-f, Q_c f)$. Pour un tel couple, la notion d'arête saturée n'est autre que notre vieil ami le sous différentiel :

$$\begin{aligned} (i, j) \text{ saturée} &\iff (-f)(i) + Q_c f(j) = c_{ij} \\ &\iff f(i) = Q_c f(j) - c_{ij} \\ &\iff (i, j) \in \partial_c f. \end{aligned}$$

La « lâcheté complémentaire » dit alors que le support du plan \mathbf{x} est inclus dans le sous-différentiel $\partial_c f$.

Remarque 6.4. On peut aussi voir le théorème comme un cas particulier du théorème de dualité en programmation linéaire. On peut, pour un problème d'optimisation linéaire quelconque, définir un problème dual. Si le primal et le dual sont faisables, alors les valeurs optimales des deux problèmes sont les mêmes, et on dispose également d'une caractérisation par « lâcheté complémentaire ».

Le cadre est ici suffisamment simple pour que l'on puisse proposer une preuve directe.

Démonstration. Tout d'abord, il est clair que l'on ne peut pas perdre de l'argent en faisant appel au transporteur : pour tout plan \mathbf{x} et tout (ϕ, ψ) faisables,

$$\begin{aligned} \sum_{i,j} c_{ij} x_{ij} &\geq \sum_{i,j} \phi(i) x_{ij} + \sum_{i,j} \psi(j) x_{ij} \\ &= \sum_i \phi(i) + \sum_j \psi(j). \end{aligned}$$

où la première ligne vient de la faisabilité de (ϕ, ψ) (et la positivité de x_{ij}), et la seconde de la faisabilité de \mathbf{x} ($\sum_i x_{ij} = 1$, $\sum_j x_{ij} = 1$). Il y a égalité entre le premier et le dernier terme si et seulement si, pour tout (i, j) ,

$$x_{ij} c_{ij} = x_{ij} (\phi(i) + \psi(j)),$$

ce qui donne exactement la lâcheté complémentaire : (i) et (ii) sont bien équivalents.

Supposons l'égalité vérifiée pour \mathbf{x}_0 et (ϕ_0, ψ_0) . Alors pour tout point \mathbf{x} faisable, comme (ϕ_0, ψ_0) est faisable on a, en appliquant ce que l'on vient d'établir :

$$\mathbf{c}^\top \mathbf{x} \geq \sum_i \phi_0(i) + \sum_j \psi_0(j) = \mathbf{c}^\top \mathbf{x}_0.$$

Donc \mathbf{x}_0 est bien optimal. On montre de façon similaire que (ϕ_0, ψ_0) est optimal.

Le dernier point sera établi plus bas constructivement : on va donner un algorithme permettant de déterminer à la fois un optimum primal \mathbf{x}_0 et un optimum dual (ϕ_0, ψ_0) vérifiant les conditions de lâcheté complémentaire. Pour une solution optimale \mathbf{x} quelconque, on a alors

$$\mathbf{c}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{x}_0 = \sum_i \phi_0(i) + \sum_j \psi_0(j)$$

donc (i) est vérifié pour \mathbf{x} et (ϕ_0, ψ_0) . □

L'algorithme donne même mieux puisqu'on verra qu'il construit une solution qui vérifie $x_{ij} \in \{0, 1\}$, ce qui implique le résultat suivant.

Théorème 6.4. *Parmi les solutions \mathbf{x} optimales pour le problème LSAP relâché, il en existe toujours une qui vérifie $x_{ij} \in \{0, 1\}$; les deux problèmes ont donc même valeur.*

Remarque 6.5. *Ce résultat peut aussi être vu comme conséquence d'un théorème de Birkhoff, qui assure que les matrices bistochastiques (c'est à dire les \mathbf{x} faisables pour LSAP relâché) sont l'enveloppe convexe des matrices de permutation (les \mathbf{x} faisables pour LSAP).*

De nombreuses conditions ont été étudiées pour savoir sous quelle condition sur la matrice A cette équivalence entre problème « à valeurs entières » et problème relâché était vraie. C'est le cas par exemple si A est *totalelement unimodulaire*¹, auquel cas le problème

$$\begin{aligned} \text{Étant donné :} & \quad \text{le vecteur } \mathbf{c}, \text{ la matrice } A, \text{ un vecteur } \mathbf{b} \text{ à coordonnées entières} \\ \text{maximiser :} & \quad \mathbf{x}^\top \mathbf{c} \\ \text{s. l. c. :} & \quad \mathbf{x}_i \geq 0, Ax = b \end{aligned}$$

s'il a une solution optimale, en a une à coordonnées entières. Cette propriété rend le problème « facile » du point de vue du calcul : le problème à contraintes entière se ramène au problème de programmation linéaire.

Faisons le point sur différentes approches possibles pour la résolution algorithmique des problèmes LSAP/LSAP relâché.

- Pour LSAP, l'ensemble des \mathbf{x} faisables est fini. Une méthode naïve est donc de calculer $\mathbf{c}^\top \mathbf{x}$ pour toutes les $N!$ assignations possibles. Cette méthode est naturellement beaucoup trop lente dès que N grandit.
- si les coûts viennent d'un problème de transport en dimension 1 (pour le coût $d(x, y)^p$) on connaît l'optimum (le couplage croissant). Il suffit donc de trier les points pour résoudre le problème, ce qui se fait en un temps $O(N \log(N))$.
- le théorème précédent montre que si l'on sait résoudre LSAP relâché, on connaît la valeur de LSAP. On peut facilement passer d'une solution optimale de LSAP relâché à une solution de LSAP. Ceci permet d'utiliser les (très nombreuses) méthodes générales de programmation linéaire comme la méthode du simplexe.
- Enfin, on peut construire des algorithmes spécifiques au problème, ce que nous allons voir ci-dessous. Le développement et l'étude de ces algorithmes a fortement inspiré de nombreuses méthodes plus générales de programmation linéaire.

1. C'est-à-dire que tous ses mineurs valent 0, 1 ou -1 .

6.4 L'algorithme hongrois

Cette méthode construit simultanément une assignation \mathbf{x} ($x_{ij} = 1$ si le yaourt i est envoyé en j , 0 sinon) et une solution du problème dual (ϕ, ψ) . Tout le long de l'algorithme on disposera :

- d'un couple (ϕ, ψ) faisable pour le problème dual, c'est-à-dire vérifiant $\phi(i) + \psi(j) \leq c_{ij}$. On dira que l'arête (i, j) est *saturée* si $\phi(i) + \psi(j) = c_{ij}$.
- d'une assignation partielle $x_{ij} \in \{0, 1\}$ vérifiant $\sum_i x_{ij} \leq 1$, $\sum_j x_{ij} \leq 1$, et la lâcheté complémentaire : si x_{ij} vaut 1 alors l'arête (i, j) est saturée. Autrement dit : seules les arêtes saturées peuvent être assignées.

Le nombre de sites assignés sera croissant pendant l'algorithme ; celui-ci finira quand tous les sites seront assignés.

Présentons d'abord une vue d'ensemble du fonctionnement de l'algorithme.

Initialisation On part de l'assignation vide $\mathbf{x} = 0$ et de (ϕ, ψ) identiquement nulles. On « augmente » d'abord ϕ et ψ en deux étapes.

1. Pour i de 1 à N on pose $\phi(i) = \min_j c_{ij}$. C'est le ϕ le plus grand tel que le couple (ϕ, ψ) reste faisable, qui n'est autre que l'inf-convolution : $\phi = Q_c(-0)$.
2. Puis, pour j de 1 à N on pose $\psi(j) = \min_i (c_{ij} - \phi(i))$. À ϕ fixé, c'est le ψ le plus grand tel que (ϕ, ψ) reste faisable, et cela correspond exactement à poser $\psi = Q_c(-\phi) = Q_c(-Q_c(-0))$.

Enfin, pour i entre 1 et N :

- si il existe j non-assigné tel que l'arête (i, j) est saturée, on assigne i au premier j trouvé (on change x_{ij} en 1) ;
- sinon i reste non-assigné.

Boucle principale Pour chaque sommet non assigné :

1. Chercher dans le graphe des arêtes saturées un *chemin augmentant*.
2. S'il y en a un, l'*inverser*.
3. S'il n'y en a pas, *créer des arêtes saturées* en modifiant (ϕ, ψ) , et revenir au 1.

Nous allons expliquer ces différentes étapes sur deux exemples, en justifiant au fur et à mesure qu'elles conservent bien les deux propriétés de faisabilité et de lâcheté complémentaire.

a. Un premier exemple

Pour le premier exemple on considère la matrice de coûts

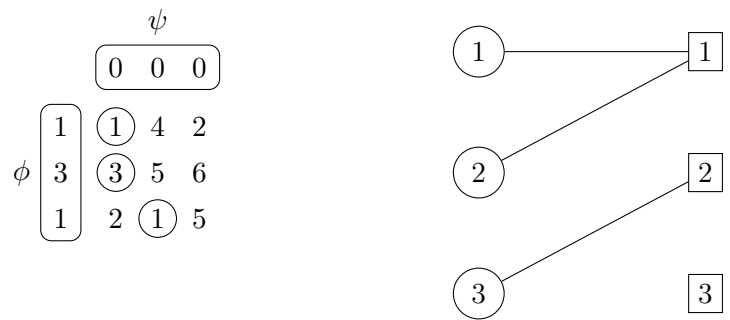
$$\mathbf{c} = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 5 & 6 \\ 2 & 1 & 5 \end{pmatrix}$$

On notera les valeurs courantes de ϕ et ψ à gauche et en haut du tableau, et on entourera les termes correspondants à des couples (i, j) *saturés*, c'est-à-dire pour lesquels $\phi(i) + \psi(j) = c_{ij}$. Par ailleurs on représentera l'état actuel de l'assignation sur un graphe : les sommets de gauche correspondent aux yaourts $i = 1, \dots, N$, les sommets de droite aux points de vente $j = 1, \dots, N$. On dessine une arête noire entre i et j si le couple (i, j) est saturé, et une arête rouge si l'assignation courante envoie i en j (c'est-à-dire si $x_{ij} = 1$).

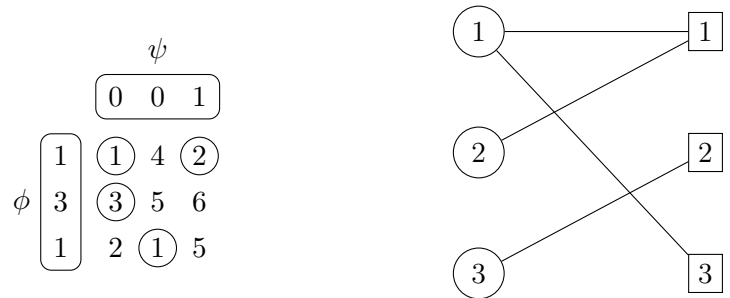
Initialisation. Au départ de l'algorithme on a donc :



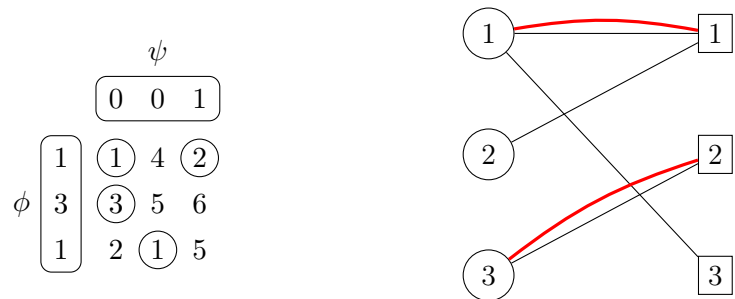
On met à jour ϕ , ce qui sature trois arêtes :



On met à jour ψ . Seul $\psi(3)$ peut être augmenté jusqu'à 1, ce qui donne :



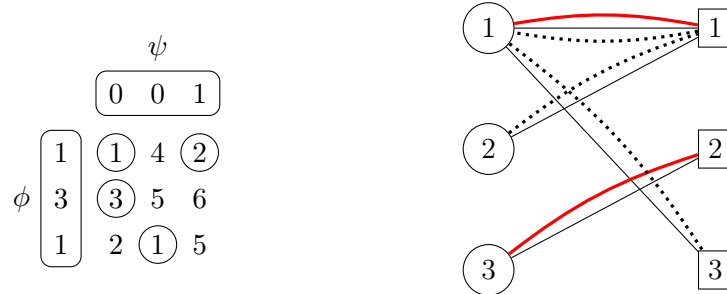
Enfin on essaie d'assigner les sommets de gauche : (1) est assigné à [1], (2) ne peut pas être assigné puisque [1] est déjà pris, et (3) est assigné à [2].



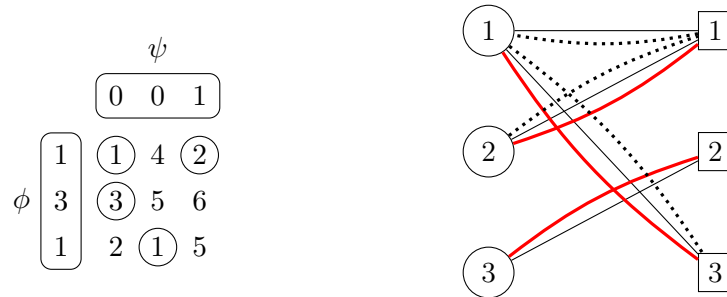
Recherche d'un chemin augmentant. Partant d'un sommet de gauche u_1 non-assigné, on cherche les chemins $(u_1, v_1, u_2, v_2, \dots)$ dans le graphe des arêtes saturées tels que :

1. les arêtes « aller » $(u_1, v_1), (u_2, v_2), \dots$ sont non-assignées ;
2. les arêtes « retour » $(v_1, u_2), (v_2, u_3), \dots$ sont assignées.

On construit ainsi un *arbre* dont les branches sont des chemins *alternés* entre arêtes assignées et non-assignées. Si un des chemins finit en un sommet de droite v_k non-assigné, on dit que le chemin est *augmentant*. Un chemin augmentant contient donc k arêtes non-assignées et $k - 1$ arêtes assignées. Dans notre exemple le chemin $\textcircled{2} - \boxed{1} - \textcircled{1} - \boxed{3}$ est augmentant.



Inversion On modifie l'assignation des arêtes le long du chemin augmentant : les arêtes non-assignées deviennent assignées et réciproquement. Comme le chemin augmentant n'utilise que des arêtes saturées, la condition de lâcheté reste vérifiée, mais le nombre d'arêtes assignées augmente de 1 (on désassigne $k - 1$ arêtes et on en assigne k). On arrive finalement à



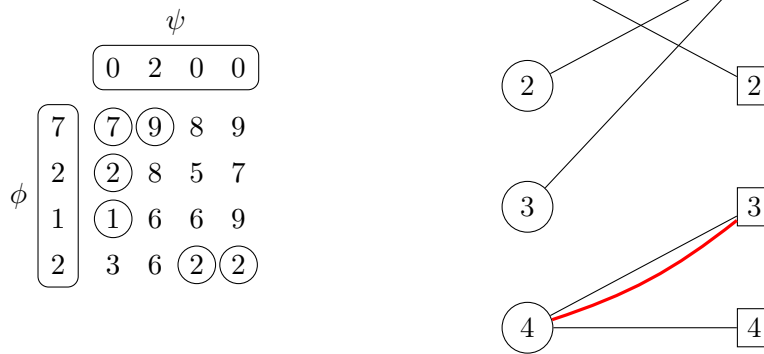
On remarque que l'assignation est complète à ce stade, et vérifie la propriété de lâcheté complémentaire : on a donc trouvé un plan de transport optimal. Le coût de transport vaut $6 = \sum \phi(i) + \sum \psi(j) = c_{13} + c_{21} + c_{32}$.

b. Un exemple complet

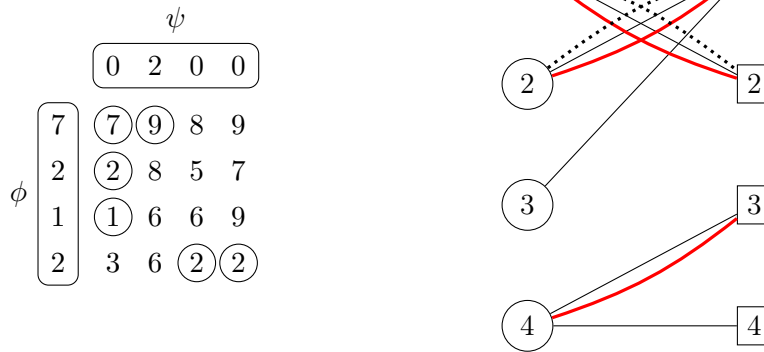
Considérons l'exemple un peu plus complexe donné par la matrice de coûts :

$$\mathbf{c} = \begin{pmatrix} 7 & 9 & 8 & 9 \\ 2 & 8 & 5 & 7 \\ 1 & 6 & 6 & 9 \\ 3 & 6 & 2 & 2 \end{pmatrix}$$

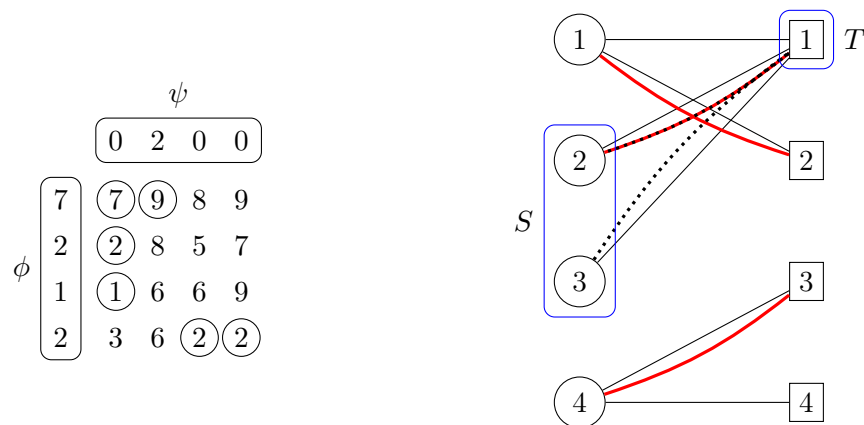
Après initialisation de ϕ , ψ , et de l'assignation on obtient



Le premier sommet de gauche non-assigné est le $\textcircled{2}$. Un seul chemin augmentant en part. On l'inverse :



On passe au sommet $\textcircled{3}$. Le seul chemin alternant qui en part est $\textcircled{3}-\boxed{1}-\textcircled{2}$, et il n'est pas augmentant puisqu'il finit en un sommet de gauche et assigné.



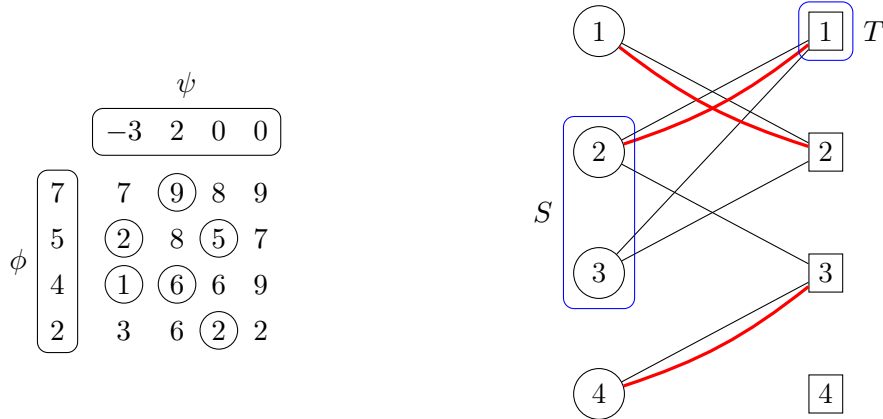
On met alors en œuvre une procédure d'ouverture d'arêtes saturées. Notons S l'ensemble des sommets de gauche sur l'arbre de chemins alternés, et T l'ensemble des sommets de droite : ici $S = \{\textcircled{2}, \textcircled{3}\}$ et $T = \{\boxed{1}\}$. Notons enfin

$$\delta = \min_{i \in S, j \notin T} \{c_{ij} - \phi(i) - \psi(j)\},$$

intuitivement c'est « ce qu'il faut rajouter aux $\phi(i)$ pour $i \in S$ pour tendre une arête vers T^c ». Dans notre exemple, $\delta = 3$: le minimum est atteint pour les deux arêtes $(i, j) = (2, 3)$ et $(3, 2)$. On met à jour ϕ et ψ en posant :

$$\tilde{\phi}(i) = \begin{cases} \phi(i) + \delta & \text{si } i \in S, \\ \phi(i) & \text{sinon,} \end{cases} \quad \tilde{\psi}(j) = \begin{cases} \psi(j) - \delta & \text{si } j \in T, \\ \psi(j) & \text{sinon.} \end{cases}$$

On obtient dans l'exemple :



Cette modification peut fermer des arêtes (ici $\textcircled{1}-\boxed{1}$ n'est plus saturée) et en ouvrir d'autres (ici, $\textcircled{3}-\boxed{2}$ et $\textcircled{2}-\boxed{3}$). Quatre points sont à remarquer :

1. les arêtes assignées restent saturées,
2. les arêtes saturées joignant S à T restent saturées,
3. la procédure ouvre au moins une arête nouvelle partant de S ,
4. le couple (ϕ, ψ) reste faisable.

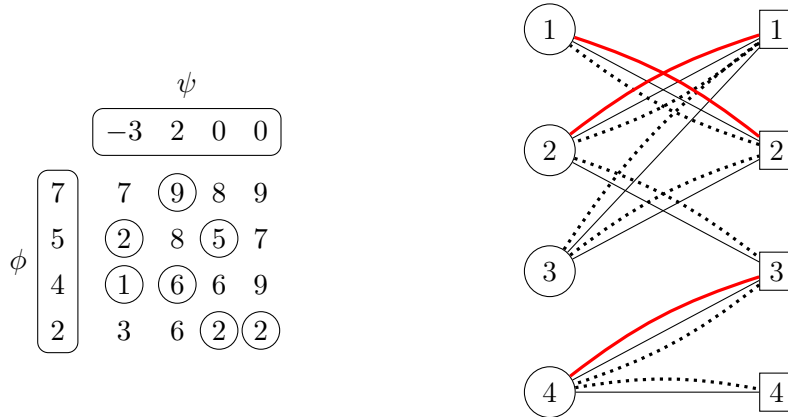
Démonstration. Soit (i, j) un couple de sommets. Quatre cas sont possibles.

1. Si $i \in S$ et $j \in T$, la somme $\tilde{\phi}(i) + \tilde{\psi}(j) = \phi(i) + \psi(j)$ n'est pas modifiée par la procédure. Le statut de l'arête (saturée ou non) ne change pas.
2. Il n'y a pas non plus de changement si ni l'un ni l'autre sommet ne sont dans l'arbre.
3. Si $i \in S$ et $j \notin T$, $\tilde{\phi}(i) + \tilde{\psi}(j) = \phi(i) + \psi(j) + \delta$. Par définition de δ ceci sature les arêtes pour lesquelles le min est atteint, mais conserve la contrainte $\tilde{\phi}(i) + \tilde{\psi}(j) \leq c_{ij}$ (le couple (ϕ, ψ) reste faisable).
4. Enfin, si $i \notin S$ et $j \in T$, $\tilde{\phi}(i) + \tilde{\psi}(j) = \phi(i) + \psi(j) - \delta$, ce qui peut désaturer l'arête (i, j) mais conserve la contrainte de faisabilité.

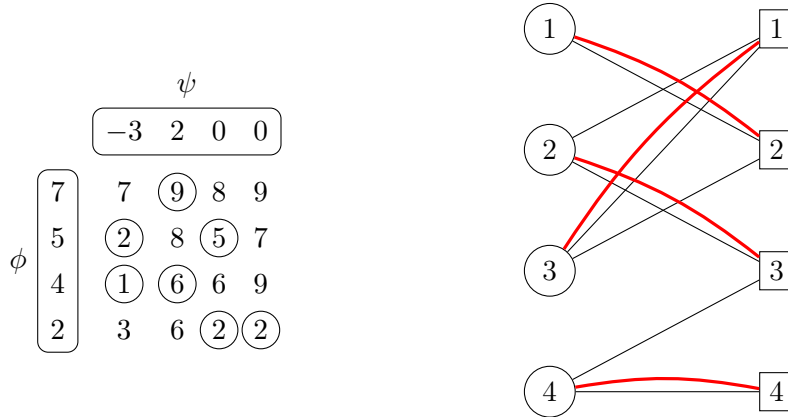
Notons enfin que par construction de l'arbre, si $v \in T$ est assigné à un sommet \textcircled{u} , on rajoute nécessairement l'arête (v, u) à l'arbre, donc $u \in S$. Les arêtes assignées sont donc soit entièrement dans l'arbre soit entièrement en dehors, et ne sont donc pas désaturées dans la procédure. \square

Remarque 6.6. On peut vérifier (exercice!) que si le couple de départ est de la forme $(Q_c f, -f)$ avec f c -convexe, c'est encore le cas (pour une fonction $\tilde{f} \dots$) après modification.

Cette modification étant faite, on revient à la recherche de chemins augmentants : l'arbre alternant aura au moins une arête de plus que précédemment. Dans notre exemple :



L'arbre contient deux chemins, $(3) \text{---} (2) \text{---} (1)$ qui n'est pas augmentant, et $(3) \text{---} (1) \text{---} (2) \text{---} (3) \text{---} (4) \text{---} (4)$ qui l'est. En inversant les assignations le long de ce dernier, on arrive finalement à :



On a une assignation complète et un couple dual (ϕ, ψ) faisable et lâchement complémentaire : on a trouvé une solution au problème de transport, ici de valeur 17.

c. Synthèse

L'algorithme présenté dans cette section :

- termine : après chaque inversion de chemin, le nombre d'arêtes assignées augmente de 1, et la procédure d'ouverture d'arêtes finit toujours par trouver un chemin augmentant.
- trouve une solution du problème : la condition de lâcheté complémentaire reste toujours vérifiée, donc dès que l'assignation est complète on a bien une solution du problème.

Rappelons que l'algorithme naïf consistant à examiner toutes les assignations possibles prend au moins $N!$ étapes; en codant intelligemment l'algorithme hongrois le nombre d'étapes peut être borné par $CN^{5/2} \log(N)$.

6.5 L'algorithme des enchères

Présentons brièvement un algorithme différent pour résoudre le problème LSAP; là encore il s'explique par une analogie économique. Supposons donc que l'on confie à chaque producteur de

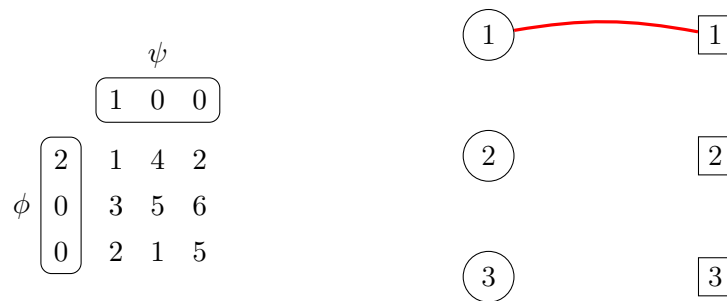
yaourts le soin d'envoyer son yaourt au site le plus rentable pour lui : pour cela il accepte de payer une certaine somme. On notera $\psi(j)$ le prix du site \boxed{j} , et $\phi(i)$ la dépense de l'usine \textcircled{i} . Au départ les prix sont nuls et aucune assignation n'est décidée.

Reprenons le premier exemple précédent, donné par la matrice de coûts :

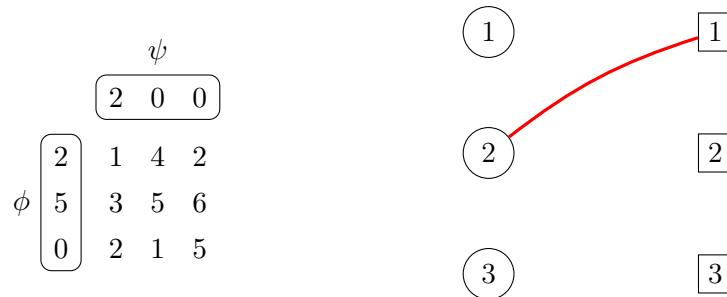
$$\mathbf{c} = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 5 & 6 \\ 2 & 1 & 5 \end{pmatrix}$$

La première usine $\textcircled{1}$ voudrait envoyer son yaourt au site $\boxed{1}$; elle est prête à payer jusqu'à 1€ pour avoir ce privilège (au-delà, le site $\boxed{3}$ deviendrait comparativement moins cher).

Le prix du site $\boxed{1}$, que nous noterons $\psi(1)$, passe donc à 1 et on assigne $\textcircled{1}$ à $\boxed{1}$. La dépense du site $\boxed{1}$ est donc $(c_{11} + \psi(1)) = 2$.



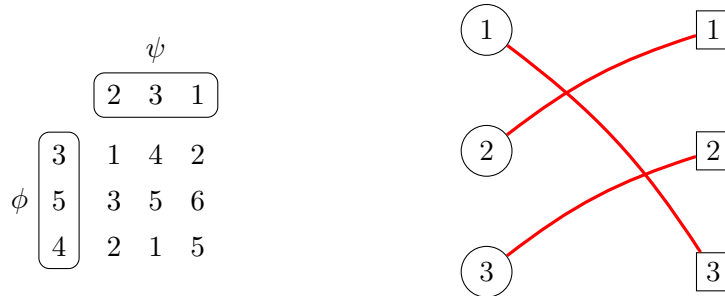
On passe à la deuxième usine : envoyer au site $\boxed{1}$ lui coûte maintenant $c_{21} + \psi(1) = 4$, envoyer en $\boxed{2}$ coûte 5 et envoyer en $\boxed{3}$ coûte 6. Le site le moins cher est donc $\boxed{1}$, et l'usine est prête à payer $5 - 4 = 1\text{€}$ supplémentaire pour y livrer son yaourt. Le prix $\psi(1)$ passe à 2, et on assigne $\textcircled{2}$ à $\boxed{1}$.



La première usine est désassignée. Le site le moins cher pour elle est maintenant le $\boxed{3}$, avec un écart de 1 : $\psi(3)$ passe à 1, et on assigne $\textcircled{1}$ à $\boxed{3}$.

Les coûts de livraison pour la troisième usine sont respectivement 4, 1 et 6 : on assigne $\textcircled{3}$ à $\boxed{2}$ en fixant le prix $\psi(2) = 3$.

À ce moment de l'algorithme tous les sommets de gauche sont assignés. Les valeurs finales de ϕ et ψ sont données par :



On remarque pour cet exemple que $(\phi, -\psi)$ est un couple faisable pour le dual, et que $(\mathbf{x}, (\phi, -\psi))$ vérifie la condition de lâcheté complémentaire : l'assignation est donc bien optimale.

Plus généralement, remarquons qu'à chaque nouvelle enchère, si \textcircled{i} enchérit pour \boxed{j} , le couple (i, j) devient tendu pour $(\phi, -\psi)$ puisqu'après avoir fixé $\psi(j)$ on pose $\phi(i) = c_{ij} + \psi(j)$. Les autres couples (k, l) :

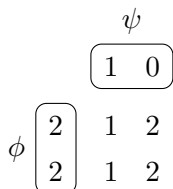
1. restent faisables : $\phi(k) - \psi(l)$ ne peut que diminuer si $k \neq i$ (puisque les prix ψ augmentent toujours et $\phi(k)$ ne change pas), et après la mise à jour des prix, $\psi(l) + c_{il} \geq \psi(j) + c_{ij}$, donc

$$\begin{aligned} \phi(i) - \psi(l) &\leq \phi(i) - \psi(j) + c_{il} - c_{ij} \\ &\leq c_{il}. \end{aligned}$$

2. restent tendus s'ils étaient assigné (puisque dans ce cas $\phi(k)$ et $\psi(l)$ ne changent pas).

Notons enfin que chaque fois qu'un nouveau site reçoit une enchère, il reste assigné dans la suite.

Si l'algorithme termine, il finit donc sur une solution lâchement complémentaire et donc optimale. Malheureusement, ce n'est pas toujours le cas : il se peut que deux sites de production alternent infiniment des enchères sans faire augmenter les prix. Ce cas arrive par exemple dans l'exemple élémentaire suivant :



Les deux sites continuent à enchérir l'un après l'autre pour le site $\boxed{1}$ sans augmenter son prix et l'algorithme boucle.

Pour contourner ce problème, on peut forcer les usines à enchérir plus. On fixe un $\varepsilon > 0$, et quand \textcircled{i} choisit \boxed{j} , il fixe le prix $\psi(j)$ à

$$\max_{l \neq j} (\psi(l) + c_{il}) + \varepsilon.$$

Ceci force les enchères à terminer, mais le résultat n'est plus nécessairement optimal : dans l'état final $(\mathbf{x}, (\phi, -\psi))$ le couple (ϕ, ψ) n'est *a priori* pas faisable, et la lâcheté complémentaire n'est vérifiée qu'à ε près (c'est-à-dire que pour tout (i, j) , $x_{ij} |c_{ij} - \phi(i) - \psi(j)| \leq \varepsilon$).

On peut montrer que si ε est assez petit, l'assignation \mathbf{x} obtenue est en fait optimale ; si les coûts c_{ij} sont entiers, il suffit par exemple de fixer $\varepsilon < 1/N$ (où N est le nombre d'usines). Il est alors souvent utile de faire une approche multi-échelles : résoudre un premier problème avec un ε_1 grand, et utiliser la solution comme point de départ d'un problème avec $\varepsilon_2 < \varepsilon_1$, etc.

6.6 Remarques bibliographiques

Pour un bref panorama des différentes approches algorithmes du problème de transport optimal, on pourra consulter les introductions de [RPDB11, PPO14, MO14], qui donnent aussi des exemples d'application en particulier en traitement d'images. Parmi les très nombreux livres sur la programmation linéaire nous avons utilisé [MG07]. L'explication de l'algorithme hongrois est adaptée de [BDM09] qui détaille de nombreuses variantes des problèmes d'assignation et de nombreux algorithmes pour les traiter, y compris l'algorithme des enchères. Pour ce dernier on peut aussi consulter le survol [Ber01] et ses références.

Chapitre 7

Distance de Wasserstein

7.1 Définition, premières propriétés

On se replace dans un cadre général en étudiant le problème de transport sur un espace métrique polonais (X, d) avec une fonction de coût c . Le coût de transport $\mathcal{T}_c(\mu, \nu)$ mesure d'une certaine manière la différence entre les mesures μ et ν : il est naturel de se demander si l'on peut, à partir du transport, définir une vraie *distance* sur l'espace des mesures de probabilités.

Pour obtenir une quantité symétrique, on considère le problème de Monge-Kantorovich. On se restreint au cas particulier des coûts définis en fonction de la distance sur X . Pour $p \in [1, \infty[$, notons

$$\mathcal{T}_p(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \iint d(x, y)^p d\pi(x, y).$$

Fixons un point x_0 quelconque dans X . Le coût de transport d'une mesure μ quelconque sur la masse de Dirac δ_{x_0} vaut

$$\mathcal{T}_p(\delta_{x_0}, \mu) = \inf_{\pi \in \Pi(\delta_{x_0}, \mu)} \iint d(x_0, y) d\pi(x, y) = \int d(x_0, y)^p d\mu(y).$$

Il est donc naturel de se restreindre aux mesures pour lesquelles ce coût est fini.

Définition 7.1. *Pour $p \in [1, \infty[$, on note $\mathcal{P}_p(X)$ l'ensemble des mesures boréliennes sur X qui admettent un moment d'ordre p :*

$$\mathcal{P}_p(X) = \left\{ \mu \in \mathcal{P}(X), \int d(x_0, x)^p d\mu(x) < \infty \right\}.$$

Il est facile de voir que cet ensemble ne dépend pas du point d'origine x_0 choisi.

Si μ et ν sont des mesures de Dirac en x et y , il n'y a qu'un seul couplage possible, et $\mathcal{T}_p(\delta_x, \delta_y) = d(x, y)^p$; pour obtenir une quantité qui se comporte comme une distance, on élève donc à la puissance $1/p$.

Définition 7.2 (Distance de Wasserstein). *La quantité*

$$W_p(\mu, \nu) = \mathcal{T}_p(\mu, \nu)^{1/p},$$

définie pour tous $(\mu, \nu) \in \mathcal{P}_p(X)$, est une distance sur $\mathcal{P}_p(X)$.

Remarque 7.1. *Cette distance est connue sous de nombreux noms différents suivant les valeurs de p , l'espace X , les auteurs et les communautés : distance de Mallows, distance du terrassier (earth mover's distance), distance de Kantorovich...*

Lors de l'étude de $W_p(\mu, \nu)$, on écrira souvent les calculs sous forme probabiliste : un couplage $\pi \in \Pi(\mu, \nu)$ sera vu comme la loi d'un couple de variables (X, Y) , dont les lois marginales sont μ et ν . On fera parfois l'abus de notation $W_p(X, Y) = W_p(\mu, \nu)$.

Vérifions que l'on a bien défini une distance.

Démonstration. La symétrie et la positivité sont immédiates. Si $W_p(\mu, \nu) = 0$, alors il existe un couplage π tel que $X = Y$ π -p.s, donc X et Y ont même loi : $\mathbb{E}[\phi(X)] = \mathbb{E}_\pi[\phi(X)] = \mathbb{E}_\pi[\phi(Y)] = \mathbb{E}[\phi(Y)]$.

Il reste donc à vérifier l'inégalité triangulaire. Soit μ_1, μ_2 et ν trois mesures dans $\mathcal{P}_p(X)$. On peut construire un vecteur (X_1, Y, X_2) , dont on notera la loi π , et qui vérifie :

- X_1, X_2 et Y ont respectivement pour loi μ_1, μ_2 et ν ;
- (X_1, Y) et (X_2, Y) sont des couples optimaux pour les problèmes de transport associés, c'est-à-dire :

$$\mathbb{E}_\pi [d(X_1, Y)^p] = \mathcal{T}_p(\mu_1, \nu), \quad \mathbb{E}_\pi [d(X_2, Y)^p] = \mathcal{T}_p(\mu_2, \nu).$$

Pour ce faire on utilise l'existence de couplages optimaux pour les deux couples (μ_1, ν) et (μ_2, ν) , qui peuvent tous deux être « désintégrés » grâce au théorème 1.4 en $d\pi_1(x, y) = p_1(y, dx)d\nu(y)$ et $d\pi_2(x_2, y) = p_2(y, dx)d\nu(y)$. En langage probabiliste, tirer (X_1, Y, X_2) suivant π correspond à

- tirer Y suivant ν , puis
- tirer X_1 suivant la loi conditionnelle $p_1(Y, \cdot)$ et X_2 suivant la loi conditionnelle $p_2(Y, \cdot)$.

En utilisant l'inégalité triangulaire dans (\mathcal{X}, d) , on a

$$0 \leq d(X_1, X_2)^p \leq (d(X_1, Y) + d(X_2, Y))^p.$$

donc, en notant $U_i = d(X_i, Y)$,

$$\begin{aligned} \mathbb{E}_\pi [d(X_1, X_2)^p]^{1/p} &\leq \mathbb{E}_\pi [(U_1 + U_2)^p]^{1/p} \\ &= \|U_1 + U_2\|_{L^p(\pi)} \\ &\leq \|U_1\|_{L^p(\pi)} + \|U_2\|_{L^p(\pi)} \\ &= W_p(\mu_1, \nu) + W_p(\mu_2, \nu) \end{aligned}$$

par l'inégalité triangulaire dans l'espace normé $L^p(\pi)$. Comme (X_1, X_2) est bien un couplage de μ_1 et μ_2 , on en déduit

$$W_p(\mu_1, \mu_2) \leq W_p(\mu_1, \nu) + W_p(\mu_2, \nu). \quad \square$$

Théorème 7.3. *Pour $1 \leq p \leq q$, $\mathcal{P}_q \subset \mathcal{P}_p$ et pour toutes μ, ν dans \mathcal{P}_q ,*

$$W_p(\mu, \nu) \leq W_q(\mu, \nu).$$

Démonstration. L'inclusion est claire, puisque pour toute mesure de probabilité, $L^q(\mu) \subset L^p(\mu)$. Si π est un couplage qui réalise $W_q(\mu, \nu)$,

$$\|d(X, Y)\|_{L^p(\pi)} \leq \|d(X, Y)\|_{L^q(\pi)} = W_q(\mu, \nu),$$

et le terme de gauche est supérieur à $W_p(\mu, \nu)$ puisque π couple μ et ν . □

7.2 Caractérisation duale de W_1

La dualité de Kantorovich donne, dans le cas $p = 1$, une expression très agréable de la distance de Wasserstein.

Théorème 7.4 (Dualité pour W_1). *Pour tout μ, ν dans \mathcal{P}_1 , la distance $W_1(\mu, \nu)$ est donnée par*

$$W_1(\mu, \nu) = \sup \left\{ \left| \int f d\mu - \int f d\nu \right| : f \text{ 1-Lipschitzienne} \right\}.$$

Remarque 7.2. *Rappelons l'expression similaire pour la distance de Fortet-Mourier :*

$$d_{F.M.}(\mu, \nu) = \sup \left\{ \left| \int \varphi d\mu - \int \varphi d\nu \right| : \varphi : X \rightarrow \mathbb{R} \text{ 1-Lipschitz et } \|\varphi\|_\infty \leq 1 \right\}$$

d'où l'on déduit en particulier que $d_{F.M.}(\mu, \nu) \leq W_1(\mu, \nu)$.

Démonstration. Notons S le sup ; il s'agit de montrer que $S = W_1(\mu, \nu)$.

Soit f une fonction 1-Lipschitzienne, π un couplage dans $\Pi(\mu, \nu)$ et (X, Y) un couple de loi π . Alors

$$\begin{aligned} \left| \int f d\mu - \int f d\nu \right| &= |\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]| \\ &= |\mathbb{E}[f(X) - f(Y)]| \\ &\leq \mathbb{E}[|f(X) - f(Y)|] \\ &\leq \mathbb{E}[|d(X, Y)|]. \end{aligned}$$

On en déduit $S \leq \mathbb{E}_\pi[d(X, Y)]$; ceci étant vrai pour tout couplage on a $S \leq W_1(\mu, \nu)$.

Montrons l'inégalité dans l'autre sens. Fixons un $\varepsilon > 0$. Par dualité de Kantorovich, il existe $\phi \in L^1(\mu)$, $\psi \in L^1(\nu)$ telles que :

$$\begin{cases} \forall x, y, & \phi(x) + \psi(y) \leq d(x, y) \\ \int \phi d\mu + \int \psi d\nu & \geq W_1(\mu, \nu) - \varepsilon. \end{cases}$$

Posons $f = P_c\psi$ la sup-convolution de f pour le coût $c(x, y) = d(x, y)$:

$$f(x) = \sup_y (\psi(y) - d(x, y)).$$

Remarquons d'abord que f est 1-Lipschitzienne : pour tout (x, x') , $d(x', y) \leq d(x, x') + d(x, y)$ par inégalité triangulaire donc

$$\begin{aligned} f(x) &= \sup_y (\psi(y) - d(x, y)) \\ &\leq \sup_y (\psi(y) - d(x', y) + d(x, x')) \\ &\leq f(x') + d(x, x'). \end{aligned}$$

Comme μ et ν sont dans \mathcal{P}_1 , on en déduit en particulier que f est dans $L^1(\mu) \cap L^1(\nu)$.

D'autre part, en prenant $y = x$ dans le sup définissant f on a $f(x) \geq \psi(x)$, et comme $\phi(x) + \psi(y) \leq d(x, y)$, on a également

$$f(x) = \sup_y (\psi(y) - d(x, y)) \leq -\phi(x).$$

Ces deux inégalités entraînent

$$S \geq - \int f d\mu + \int f d\nu \geq \int \phi d\mu + \int \psi d\nu \geq W_1 - \varepsilon,$$

ce qui achève la preuve puisque ε est arbitraire. \square

Remarque 7.3. *On a vu dans la preuve qu'une fonction c -convexe pour le coût $c(x, y) = d(x, y)$ était 1-Lipschitzienne : on pourra vérifier (exercice) que la réciproque est vraie, et relier ce fait à l'interprétation géométrique de la c -convexité.*

7.3 Lien avec la topologie faible

L'objectif de cette section est de montrer le résultat suivant.

Théorème 7.5 (Lien avec la convergence faible). *Soit μ et (μ_n) des mesures dans \mathcal{P}_p . Alors $W_p(\mu_n, \mu)$ tend vers 0 si et seulement si :*

1. μ_n converge faiblement vers μ ,
2. les moments d'ordre p convergent :

$$\int d(x_0, x)^p d\mu_n(x) \rightarrow \int d(x_0, x)^p d\mu(x).$$

Pour cela nous allons d'abord faire quelques rappels.

a. Convergence faible : rappels

Revenons d'abord sur la convergence faible pour des suites de mesures de probabilités (sur un espace métrique polonais).

Théorème 7.6 (Théorème du porte-manteau). *La convergence faible de μ_n vers μ équivaut à :*

- (i) $\int f d\mu_n \rightarrow \int f d\mu$ pour toute f continue bornée,
- (ii) $\int f d\mu_n \rightarrow \int f d\mu$ pour toute f lipschitzienne bornée,
- (iii) $\liminf \int f d\mu_n \geq \int f d\mu$ pour toute f s.c.i. minorée,
- (iv) $\liminf \mu_n(O) \geq \mu(O)$ pour tout ouvert O .

Démonstration. (i) \implies (ii) est clair.

(ii) \implies (iii). Soit f une fonction s.c.i. minorée. On a vu plus haut (Proposition 2.4) que f était limite croissante d'une suite de fonctions (f_k) Lipschitziennes bornées. Comme la suite est croissante, on a pour tout k :

$$\int f_k d\mu_n \leq \int f d\mu_n.$$

Comme f_k est Lipschitzienne bornée, on en déduit

$$\int f_k d\mu \leq \liminf_n \int f d\mu_n.$$

On fait tendre k vers l'infini ; par convergence monotone on conclut :

$$\int f d\mu \leq \liminf_n \int f d\mu_n.$$

(iii) \implies (iv) est direct puisque l'indicatrice d'un ouvert est minorée et s.c.i.

(iv) \implies (i). Soit f continue bornée. Quitte à la translater, on peut la supposer positive. On utilise l'astuce classique de décomposition en lignes de niveau :

$$\begin{aligned}\int f(x)d\mu_n(x) &= \int \left(\int_0^\infty \mathbf{1}_{t < f(x)} dt \right) d\mu_n(x) \\ &= \int_0^\infty \mu_n(f^{-1}(]t, \infty[)) dt\end{aligned}$$

par Fubini-Tonelli. Comme f est continue, $O_t = f^{-1}(]t, \infty[)$ est un ouvert, donc

$$\liminf_n \mu_n(O_t) \geq \mu(O_t)$$

par hypothèse. Posons $g_n(t) = \mu_n(O_t)$,

$$\begin{aligned}\int f d\mu &= \int_0^\infty \mu(O_t) dt \leq \int_0^\infty \liminf_n g_n(t) dt \\ &\leq \liminf_n \int_0^\infty g_n(t) dt && \text{(lemme de Fatou)} \\ &= \liminf_n \int f d\mu_n.\end{aligned}$$

On applique le même raisonnement à $g = (\|f\|_\infty - f)$ pour conclure. \square

Théorème 7.7 (Préservation par fonction continue). *Soit (X_n) une suite de v.a à valeurs dans un espace polonais (\mathcal{X}, d) et ϕ une application continue de (\mathcal{X}, d) vers un autre espace polonais (\mathcal{X}', d') . Si X_n converge en loi vers X alors $\phi(X_n)$ converge en loi vers $\phi(X)$.*

Démonstration. Il suffit de revenir à la définition en remarquant que pour toute fonction test f continue bornée, $f \circ \phi$ est continue bornée. \square

Théorème 7.8 (Convergence en loi vers une constante). *Si la suite de variables $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une constante c , alors la convergence a lieu en probabilité :*

$$\forall \varepsilon, \mathbb{P}[|X_n - c| \geq \varepsilon] \xrightarrow{n \rightarrow \infty} 0.$$

Rappelons que la réciproque est toujours vraie (la convergence en probabilité entraîne la convergence en loi).

Démonstration. Il suffit de remarquer que $\mathbb{P}[|X_n - c| < \varepsilon] = \mu_n(]c - \varepsilon, c + \varepsilon[)$, ce qui entraîne

$$1 \geq \liminf_n \mathbb{P}[|X_n - c| < \varepsilon] \geq \mathbb{P}[0 < \varepsilon] = 1$$

par le point (iv) du théorème du porte-manteau. \square

b. Uniforme intégrabilité

Nous avons vu plus haut la notion de *tension* pour un ensemble de variables aléatoires : \mathcal{A} est tendu si pour tout $\varepsilon > 0$, il existe un compact K tel que pour toute variable X de l'ensemble \mathcal{A} ,

$$\mathbb{P}[X \in K] \geq 1 - \varepsilon.$$

Remarquons que la tension ne fait intervenir que la topologie de l'espace ambiant. Nous allons maintenant considérer une notion similaire mais utilisant la métrique de \mathbb{R} .

Définition 7.9 (Uniforme intégrabilité). *Soit \mathcal{A} un ensemble de variables aléatoires réelles. On dit que \mathcal{A} est uniformément intégrable (abrégé en u.i.) si :*

$$\sup_{X \in \mathcal{A}} \mathbb{E} [|X| \mathbf{1}_{|X| > R}] \xrightarrow{R \rightarrow \infty} 0.$$

Une suite de variables $(X_n)_{n \in \mathbb{N}}$ est dite u.i. si l'ensemble $\{X_n; n \in \mathbb{N}\}$ l'est.

Exercice 7.1. *Soit X_n, Y_n deux suites de v.a. telle que $\mathbb{P}[X_n = 0] = \mathbb{P}[Y_n = 0] = 1 - 1/n$, $\mathbb{P}[X_n = \sqrt{n}] = \mathbb{P}[Y_n = n] = 1/n$. Montrer que les suites sont tendues. Montrer que X_n est uniformément intégrable, mais que Y_n ne l'est pas.*

Théorème 7.10 (Conditions suffisantes).

Union, familles finies *Si \mathcal{A} est u.i., toute variable $X \in \mathcal{A}$ est intégrable. Si X est intégrable, $\{X\}$ est u.i. Si \mathcal{A} et \mathcal{B} sont u.i., alors $\mathcal{A} \cup \mathcal{B}$ l'est aussi. Toute ensemble fini de variables intégrables est u.i.*

Borne dans L^p *Si l'ensemble \mathcal{A} est borné dans L^p pour un $p > 1$, c'est-à-dire si*

$$\sup_{X \in \mathcal{A}} \mathbb{E} [|X|^p] < \infty,$$

alors il est u.i.

Comparaison *Si pour tout X de \mathcal{A} il existe Y dans \mathcal{B} tel que $|X| \leq Y$, et si \mathcal{B} est u.i., alors \mathcal{A} l'est aussi. En particulier, si les variables de \mathcal{A} ont un « chapeau intégrable » Y , elles sont u.i.*

Stabilité *Si \mathcal{A} et \mathcal{B} sont u.i., alors $\{X + Y; X \in \mathcal{A}, Y \in \mathcal{B}\}$ l'est aussi. En particulier la somme de deux suites u.i. est u.i.*

Suites *Si les $(X_n)_{n \in \mathbb{N}}$ sont intégrables et si*

$$\lim_{R \rightarrow \infty} \limsup_n \mathbb{E} [|X_n| \mathbf{1}_{|X_n| > R}] = 0,$$

alors les $(X_n)_{n \in \mathbb{N}}$ sont u.i.

Démonstration. Union. Si \mathcal{A} est u.i. et $X \in \mathcal{A}$, il existe un R tel que $\mathbb{E} [|X| \mathbf{1}_{|X| > R}] \leq 1$, et pour lequel on a donc $\mathbb{E} [|X|] \leq R + 1 < \infty$. L'uniforme intégrabilité de $\{X\}$ est une application directe de la convergence dominée. Pour l'union proprement dite, on écrit

$$\sup_{Z \in \mathcal{A} \cup \mathcal{B}} \mathbb{E} [Z \mathbf{1}_{|Z| > R}] = \max \left(\sup_{X \in \mathcal{A}} \mathbb{E} [X \mathbf{1}_{|X| > R}], \sup_{Y \in \mathcal{B}} \mathbb{E} [Y \mathbf{1}_{|Y| > R}] \right).$$

On prend la limite en R : par monotonie c'est également la \limsup , or $\limsup_R \max(f(R), g(R)) = \max(\limsup_R f, \limsup_R g)$ pour toutes fonctions réelles f et g . On en déduit la propriété sur l'union. Le dernier point s'obtient par une récurrence immédiate.

Borne dans L^p . La preuve est similaire à celle de l'inégalité de Markov : pour tout X dans \mathcal{A} ,

$$\begin{aligned}\mathbb{E} [|X| \mathbf{1}_{|X|>R}] &\leq \mathbb{E} \left[|X| \left(\frac{|X|}{R} \right)^{p-1} \mathbf{1}_{|X|>R} \right] \\ &\leq \frac{1}{R^{p-1}} \sup_{Y \in \mathcal{A}} \mathbb{E} [|Y|^p],\end{aligned}$$

qui tend bien vers 0 quand R tend vers l'infini, uniformément par rapport à X .

Comparaison. Ce point est direct puisque $x \mapsto x \mathbf{1}_{x>R}$ est croissante.

Stabilité. Soit U, V deux variables positives et $R > 0$.

$$\mathbb{E} [(U + V) \mathbf{1}_{U+V>2R}] = \mathbb{E} [(U + V) \mathbf{1}_{U+V>2R} \mathbf{1}_{U \leq V}] + \mathbb{E} [(U + V) \mathbf{1}_{U+V>2R} \mathbf{1}_{V < U}].$$

Dans le premier terme, comme $U \leq V$, V est nécessairement plus grand que R ; de plus $(U + V)$ est majoré par $2V$. En raisonnant de même pour le second terme :

$$\mathbb{E} [(U + V) \mathbf{1}_{U+V>2R}] \leq 2\mathbb{E} [V \mathbf{1}_{V>R}] + 2\mathbb{E} [U \mathbf{1}_{U>R}].$$

En appliquant ces inégalités à $U = |X|$ et $V = |Y|$ on obtient l'uniforme intégrabilité de $\{|X| + |Y|\}$, qui entraîne celle de $\{X + Y\}$ par comparaison.

Suites La quantité $\limsup_n \mathbb{E} [|X_n| \mathbf{1}_{|X_n|>R}]$ étant décroissante en R , l'hypothèse se réécrit :

$$\inf_{R>0, n_0 \in \mathbb{N}} \sup_{n \geq n_0} \mathbb{E} [|X_n| \mathbf{1}_{|X_n|>R}] = 0.$$

Pour $\varepsilon > 0$, il existe donc n_0 tel que $\limsup_n \mathbb{E} [|X_n| \mathbf{1}_{|X_n|>R}] = \inf_{R>0} \sup_{n \geq n_0} \mathbb{E} [|X_n| \mathbf{1}_{|X_n|>R}] \leq \varepsilon$. Comme l'ensemble $\{X_k, k < n_0\}$ est u.i, on a par le même raisonnement que dans la preuve de la propriété d'union,

$$\limsup_{R>0} \sup_n \mathbb{E} [|X_n| \mathbf{1}_{|X_n|>R}] \leq \max(0, \varepsilon).$$

Ceci conclut la preuve puisque ε est arbitraire. □

L'intégrabilité uniforme est une notion importante en raison du résultat suivant de « passage à la limite dans l'intégrale », généralisation du théorème de convergence dominée.

Théorème 7.11. *Si la suite de variables aléatoires X_n :*

- converge en probabilité vers une variable X intégrable,
- est uniformément intégrable,

alors $\mathbb{E} [X_n]$ tend vers $\mathbb{E} [X]$.

Démonstration. Par le lemme de stabilité, $Y_n = |X_n - X|$ est uniformément intégrable, de plus Y_n converge en probabilité vers 0. Pour $0 < \varepsilon < R$ arbitraires, on découpe $\mathbb{E} [Y_n]$ en trois morceaux :

$$\begin{aligned}\mathbb{E} [Y_n] &= \mathbb{E} [Y_n \mathbf{1}_{Y_n \leq \varepsilon}] + \mathbb{E} [Y_n \mathbf{1}_{\varepsilon < Y_n \leq R}] + \mathbb{E} [Y_n \mathbf{1}_{Y_n > R}] \\ &\leq \varepsilon + R \mathbb{P} [Y_n > \varepsilon] + \sup_m \mathbb{E} [Y_m \mathbf{1}_{Y_m > R}].\end{aligned}$$

Comme Y_n converge en probabilité vers 0,

$$\limsup_n \mathbb{E} [Y_n] \leq \varepsilon + \sup_m \mathbb{E} [Y_m \mathbf{1}_{Y_m > R}].$$

On fait tendre ε vers 0 et R vers l'infini, le dernier terme tend vers 0 par uniforme intégrabilité donc $\mathbb{E} [Y_n]$ tend vers 0. Par conséquent $|\mathbb{E} [X_n] - \mathbb{E} [X]| \leq \mathbb{E} [Y_n]$ tend également vers 0, ce qui achève la preuve. □

Remarque 7.4. On peut affaiblir les hypothèses en ne demandant que la convergence en loi de X_n vers X et l'uniforme intégrabilité, voir par exemple [Bil99, Théorème 3.5].

On aura également besoin de la « réciproque » suivante.

Théorème 7.12. Si la suite de variables aléatoires X_n vérifie :

- X_n est positive et intégrable,
- X_n converge en loi vers X ,
- $\mathbb{E}[X_n] \rightarrow \mathbb{E}[X]$,

alors les (X_n) sont uniformément intégrables.

Démonstration. La fonction $f : x \mapsto x\mathbf{1}_{x < R}$ est s.c.i. Par le théorème du porte-manteau, on a donc

$$\liminf_n \mathbb{E}[X_n \mathbf{1}_{X_n < R}] \geq \mathbb{E}[X \mathbf{1}_{X < R}].$$

Comme $\mathbb{E}[X_n]$ tend vers $\mathbb{E}[X]$, on en déduit

$$\begin{aligned} \limsup_n \mathbb{E}[X_n \mathbf{1}_{X_n \geq R}] &= \limsup_n (\mathbb{E}[X_n] - \mathbb{E}[X_n \mathbf{1}_{X_n < R}]) \\ &= \mathbb{E}[X] - \liminf_n \mathbb{E}[X_n \mathbf{1}_{X_n < R}] \\ &\leq \mathbb{E}[X] - \mathbb{E}[X \mathbf{1}_{X < R}] \\ &= \mathbb{E}[X \mathbf{1}_{X \geq R}]. \end{aligned}$$

Comme X est intégrable, on en déduit

$$\lim_{R \rightarrow \infty} \limsup_n \mathbb{E}[|X_n| \mathbf{1}_{|X_n| > R}] = 0,$$

ce qui implique l'uniforme intégrabilité par le dernier point du théorème 7.10. □

c. Deux faits élémentaires

Lemme 7.13. Pour tout ε , tout $p \geq 1$, il existe $C(p, \varepsilon)$ telle que :

$$\forall (a, b) \in \mathbb{R}_+, \quad |a^p - b^p| \leq \varepsilon a^p + C(p, \varepsilon) |b - a|^p.$$

Démonstration. Remarquons d'abord que pour tout $p \geq 1$, tout $\varepsilon' > 0$,

$$\forall (a', b') \in \mathbb{R}_+, \quad (a' + b')^p \leq (1 + \varepsilon')^p (a')^p + (1 + \frac{1}{\varepsilon'})^p (b')^p.$$

En effet, le membre de gauche est majoré par le premier terme du membre de droite si $b' \leq \varepsilon' a'$, et par le second si $b' > \varepsilon' a'$. On applique cette inégalité à $a' = \min(a, b)$, $b' = \max(a, b) - \min(a, b) = |b - a|$ et $\varepsilon' = (1 + \varepsilon)^{1/p} - 1$ pour obtenir, en posant $C(p, \varepsilon) = (1 + \frac{1}{\varepsilon'})^p$,

$$\max(a, b)^p - \min(a, b)^p \leq \varepsilon \min(a, b)^p + C(p, \varepsilon) |b - a|^p.$$

qui entraîne l'inégalité annoncée. □

Lemme 7.14 (Un résultat classique sur les suites). Soit u_n une suite dans un espace topologique. Si il existe un l tel que, de toute sous-suite v_n de u_n , on peut extraire une sous-suite w_n qui converge vers l , alors u_n converge vers l .

Démonstration. Montrons la contraposée, en supposant que u_n ne converge pas vers l . Il existe un ouvert O contenant l et une sous-suite (v_n) telle que $v_n \notin O$. Une sous suite de (v_n) ne peut donc pas converger vers l . □

d. Preuve du théorème

Sens direct — convergence faible. On suppose que $W_p(\mu_n, \mu)$ tend vers 0. Par comparaison des distances de Wasserstein, $W_1(\mu_n, \mu)$ tend vers 0, donc pour toute fonction 1-Lipschitzienne,

$$\int f d\mu_n \rightarrow \int f d\mu.$$

Ceci reste vrai par homogénéité pour toute fonction Lipschitzienne, et *a fortiori* pour toute fonction Lipschitzienne bornée. Par le théorème du porte-manteau, on a donc bien convergence faible de μ_n vers μ .

Sens direct — convergence des moments. Pour montrer la convergence des moments, on réalise les mesures μ_n et μ comme lois de variables aléatoires sur un même espace de probabilités. On peut les construire de manière à vérifier :

$$\mathbb{E} [d(X_n, X)^p] = W_p(\mu_n, \mu)^p,$$

c'est-à-dire que (X_n, X) est un couplage optimal. On majore la différence des moments :

$$|\mathbb{E} [d(X_n, x_0)^p] - \mathbb{E} [d(X, x_0)^p]| \leq \mathbb{E} [|d(X_n, x_0)^p - d(X, x_0)^p|].$$

On fixe un ε et on applique le lemme 7.13 puis l'inégalité triangulaire dans (\mathcal{X}, d) :

$$\begin{aligned} \mathbb{E} [|d(X_n, x_0)^p - \mathbb{E} [d(X, x_0)^p]|] &\leq \varepsilon \mathbb{E} [d(X, x_0)^p] + C(p, \varepsilon) \mathbb{E} [|d(X_n, x_0) - d(X, x_0)|^p] \\ &\leq \varepsilon \mathbb{E} [d(X, x_0)^p] + C(p, \varepsilon) \mathbb{E} [d(X_n, X)^p]. \end{aligned}$$

Comme la dernière espérance vaut $W_p(\mu_n, \mu)^p$, qui tend vers 0 par hypothèse, on en déduit :

$$\limsup_n |\mathbb{E} [d(X_n, x_0)^p] - \mathbb{E} [d(X, x_0)^p]| \leq \varepsilon \mathbb{E} [d(X, x_0)^p].$$

Comme ε est arbitraire on a bien la convergence de $\mathbb{E} [d(X_n, x_0)^p]$ vers $\mathbb{E} [d(X, x_0)^p]$.

Sens réciproque. On suppose que μ_n converge faiblement vers μ et que les moments d'ordre p convergent. Soit π_n un couplage optimal de μ_n et μ , loi d'un couple (X_n, X) . Posons $Z_n = d(X_n, X)^p$: on veut montrer que $W_p(\mu_n, \mu)$ tend vers 0, ce qui revient à établir que $\mathbb{E} [Z_n]$ tend vers 0.

Supposons tout d'abord que π_n converge faiblement vers un couplage π_∞ .

Convergence en probabilité de Z_n vers 0. On peut montrer assez facilement que π_∞ a pour marginales μ et μ , et qu'il est optimal (en montrant que son support vérifie la propriété de monotonie cyclique) : c'est donc le couplage trivial, et (X_n, X) converge en loi vers (X, X) . La continuité de la distance entraîne que $Z_n = d(X_n, X)^p$ converge en loi vers 0 ; comme la limite est constante la convergence a également lieu en probabilité.

Uniforme intégrabilité. Posons $Y_n = d(X_n, x_0)^p$ et $Y = d(X, x_0)^p$. Comme X_n converge en loi vers X , Y_n converge en loi vers Y . De plus les Y_n sont positives et par hypothèse $\mathbb{E} [Y_n]$ converge vers $\mathbb{E} [Y]$. Par le théorème 7.12 elles sont donc uniformément intégrables.

Comme $Z_n \leq 2^{p-1}(Y_n + Y)$, on en déduit que la suite Z_n est uniformément intégrable. Comme les Z_n sont positives et convergent en probabilité vers 0, on peut appliquer le théorème 7.11 :

$$W_p(\mu_n, \mu)^p = \mathbb{E} [Z_n] \xrightarrow{n \rightarrow \infty} 0.$$

Cas général. Dans le cas général où π_n ne converge pas nécessairement, soit $\mu'_n = \mu_{\phi(n)}$ une sous-suite quelconque de μ_n , et π'_n les couplages correspondants. On peut montrer assez facilement (exercice!) que comme μ_n converge faiblement, l'ensemble de mesures

$$\bigcup_n \Pi(\mu_n, \mu)$$

est tendu. Il existe donc une sous-suite π''_n de π'_n qui converge faiblement. En appliquant les deux premières étapes, on sait que le long de cette sous-suite $W_p(\mu''_n, \mu)$ tend vers 0. Le théorème s'en déduit en appliquant le lemme 7.14.

7.4 Remarques bibliographiques

Notre présentation s'appuie principalement sur [AG]; pour les résultats sur la topologie faible on pourra consulter [Bil99].

Chapitre 8

Applications en probabilités

8.1 Une preuve du théorème limite central

a. Le résultat

Rappelons tout d'abord l'énoncé de ce résultat fondamental de probabilités.

Théorème 8.1 (TLC). *Si les X_i sont des variables i.i.d. admettant un moment d'ordre 2, alors leur somme centrée réduite converge en loi vers la gaussienne standard :*

$$\frac{\sum_{i=1}^n X_i - n\mathbb{E}[X_1]}{\sqrt{n\text{Var}(X_1)}} \xrightarrow{(d)} \mathcal{N}(0, 1).$$

En particulier si les X_i sont centrées et réduites¹,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i \xrightarrow{(d)} \mathcal{N}(0, 1).$$

La preuve habituelle de ce résultat passe par les fonctions caractéristiques, c'est-à-dire par une transformée de Fourier. D'autres preuves plus élémentaires peuvent être obtenues avec des arguments de couplage. Nous allons voir ici une preuve utilisant la distance de Wasserstein W_2 .

Définition 8.2 (Écart à la normale). *Pour X une variable aléatoire de carré intégrable, on note*

$$D(X) = W_2^2(X, \mathcal{N}(\mathbb{E}[X], \text{Var}(X)))$$

le coût de transport quadratique de X sur la loi normale de même moyenne et même variance.

Théorème 8.3 (TLC en distance de Wasserstein). *Si les X_i sont i.i.d. centrées réduites, alors*

$$D\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n X_i\right) \rightarrow 0,$$

donc $\frac{1}{\sqrt{n}} \sum_i X_i$ converge en loi vers une gaussienne $\mathcal{N}(0, 1)$.

1. C'est-à-dire $\mathbb{E}[X_i] = 0$ et $\text{Var}(X_i) = 1$.

b. Distance à la normalité

Théorème 8.4 (Propriétés de D).

D est quadratique. Pour tout $(a, b) \in \mathbb{R}^2$, $D(aX + b) = a^2D(X)$.

Écart à la normalité. Si X est de variance s_0^2 ,

$$\begin{aligned} \inf_{m \in \mathbb{R}, s > 0} W_2^2(X, \mathcal{N}(m, s^2)) &= D(X) \left(\frac{4s_0^2 - D(X)}{4s_0^2} \right) \\ &\in \left[\frac{1}{2}D(X), D(X) \right]. \end{aligned}$$

Convergence en loi. Si X_n converge en loi vers X , et si $\text{Var}(X_n)$ tend vers $\text{Var}(X)$, alors $D(X_n)$ tend vers $D(X)$.

Sous-additivité. Pour (X, Y) indépendantes,

$$D(X + Y) \leq D(X) + D(Y).$$

Il y a égalité si et seulement si X et Y suivent des lois normales.

Démonstration. Pour tout a, b , on montre facilement que $W_2^2(aX + b, aY + b) = a^2W_2^2(X, Y)$, ce qui donne le caractère quadratique de D .

Pour l'écart à la normalité, on peut supposer sans perte de généralité que X est centrée. Le fait que l'on regarde le coût $c(x, y) = |x - y|^2$ sur \mathbb{R} montre que le couplage croissant est toujours optimal. Par conséquent, si le couple (X, Z) est optimal, alors $(X, m + sZ)$ l'est aussi, donc

$$W_2^2(X, \mathcal{N}(m, s^2)) = \mathbb{E} [(X - m - sZ)^2] = s_0^2 - 2s\mathbb{E}[XZ] + s^2 + m^2.$$

Or $D = \mathbb{E} [(X - s_0Z)^2] = 2s_0^2 - 2s_0\mathbb{E}[XZ]$. Donc

$$W_2^2(X, \mathcal{N}(m, s^2)) = s_0^2 + \frac{s}{s_0}(D - 2s_0^2) + s^2 + m^2.$$

Le minimum en m et s de cette quantité est obtenu² pour $m = 0$ et $s = \frac{2s_0^2 - D}{2s_0} = s_0 - \frac{D}{2s_0}$. Il vaut

$$\begin{aligned} \inf_{m, s} W_2^2(X, \mathcal{N}(m, s^2)) &= s_0^2 - \frac{1}{2s_0^2}(D - 2s_0^2)^2 + \frac{D - 2s_0^2}{(2s_0)^2} \\ &= \frac{4s_0^4 - D^2 + 4s_0^2D - 4s_0^4}{4s_0^2} \\ &= D \frac{4s_0^2 - D}{4s_0^2}. \end{aligned}$$

Pour montrer la propriété de *convergence en loi*, considérons une suite $(X_n)_{n \in \mathbb{N}}$ qui converge en loi vers X , et telle que $\text{Var}(X_n)$ tend vers $\text{Var}(X)$. Nous avons vu au chapitre précédent que X_n converge alors nécessairement vers X en distance de Wasserstein. Notons m_n, s_n, m et s les moyennes et écarts-type de X_n et X . Par l'inégalité triangulaire sur W_2 ,

$$\sqrt{D(X_n)} \leq W_2(X_n, X) + W_2(X, m + sZ) + W_2(m + sZ, m_n + s_nZ),$$

2. Notons que la loi normale la plus proche de X a donc, si D est non-nul, un écart type strictement inférieur à s_0 .

donc³ $\limsup_n D(X_n) \leq D(X)$. On prouve de manière symétrique $\liminf_n D(X_n) \geq D(X)$, ce qui conclut la preuve.

Considérons maintenant la propriété de *sous-additivité*. Grâce au premier point on peut supposer que $\mathbb{E}[X] = \mathbb{E}[Y] = 0$. L'inégalité est facile à montrer. En effet, l'existence de couplages optimaux montre que l'on peut définir un quadruplet (X, Y, Z, T) de variables vérifiant :

- $Z \sim \mathcal{N}(0, \text{Var}(X))$, $T \sim \mathcal{N}(0, \text{Var}(Y))$,
- $D(X) = \mathbb{E}[(X - Z)^2]$, $D(Y) = \mathbb{E}[(Y - T)^2]$,
- (X, Z) est indépendant de (Y, T) .

On dispose alors d'une façon de coupler $(X + Y)$ avec une loi normale $\mathcal{N}(0, \text{Var}(X) + \text{Var}(Y))$: c'est le couple $(Z + T, X + Y)$. Par conséquent

$$\begin{aligned} D(X + Y) &\leq \mathbb{E}[(X + Y - Z - T)^2] = \mathbb{E}[(X - Z)^2] + \mathbb{E}[(Y - T)^2] + 2\mathbb{E}[(X - Z)(Y - T)] \\ &= D(X) + D(Y) \end{aligned}$$

puisque le dernier terme est nul par indépendance.

Le dernier point à montrer, le *cas d'égalité*, est le plus délicat. Supposons donc que $D(X, Y) = D(X) + D(Y)$: par l'équation précédente ceci implique que le couplage $(Z + T, X + Y)$ est optimal. On peut supposer que $X = a(Z)$ et $Y = b(T)$ avec a et b croissantes et continues à gauche, données par $a = F_X^{-1} \circ F_Z$ et $b = F_Y^{-1} \circ F_T$ (c'est le couplage croissant de la proposition 3.9, qui est optimal par le théorème 3.10). Notons π la loi de ce couplage optimal $(Z + T, a(Z) + b(T))$. Il s'agit alors de montrer que les fonctions a et b sont affines, ce qui impliquera la normalité de X et Y .

Les couples $(z_0 + t_0, a(z_0) + b(t_0))$ sont dans le support de π . Montrons que, pour tout (z_0, t_0) , le couple $(z_0 + t_0, a(z_0) + b(t_0))$ appartient au support du couplage. Supposons d'abord z_0 et t_0 sont des points de continuité pour a et b , et soit O un ouvert contenant $(z_0 + t_0, a(z_0) + b(t_0))$. L'application $\phi(z, t) \mapsto (z + t, a(z) + b(t))$ est alors continue en (z_0, t_0) , donc l'image réciproque $\phi^{-1}(O)$ est un ouvert, non-vide puisqu'il contient (z_0, t_0) . Par conséquent

$$\pi(O) = \mathbb{P}[\phi(Z, T) \in O] = \mathbb{P}[(Z, T) \in \phi^{-1}(O)] > 0.$$

puisque (Z, T) a une densité strictement positive sur tout \mathbb{R}^2 . Revenant au cas général, on peut écrire $z_0 = \lim z_n$ et $t_0 = \lim t_n$ avec (z_n) et (t_n) croissantes et telles que z_n, t_n sont des points de continuité pour a et b . Pour tout n , $(z_n + t_n, a(z_n) + b(t_n))$ est donc dans le support de π . Comme a et b sont continues à gauche et comme le support est un fermé, on en déduit que $(z_0 + t_0, a(z_0) + b(t_0))$ est bien dans le support.

Une équation fonctionnelle. Nous avons vu que le support d'un couplage optimal est nécessairement cycliquement monotone : ici, si (α_1, β_1) et (α_2, β_2) sont dans le support alors

$$\begin{aligned} &(\alpha_1 - \beta_1)^2 + (\alpha_2 - \beta_2)^2 \leq (\alpha_1 - \beta_2)^2 + (\alpha_2 - \beta_1)^2 \\ \iff &-2\alpha_1\beta_1 - 2\alpha_2\beta_2 \leq -2\alpha_1\beta_2 - 2\alpha_2\beta_1 \\ \iff &(\alpha_2 - \alpha_1)(\beta_2 - \beta_1) \geq 0. \end{aligned}$$

Par conséquent, si $\alpha_1 < \alpha_2$ alors $\beta_1 \leq \beta_2$.

L'argument de monotonie cyclique donné plus haut montre alors que pour tout (z_1, t_1, z_2, t_2) :

$$(z_1 + t_1 < z_2 + t_2) \implies (a(z_1) + b(t_1) \leq a(z_2) + b(t_2)).$$

3. La convergence du troisième terme vers 0 est laissée en exercice.

Comme a et b sont continues à gauche, on en déduit que l'implication est vraie avec une inégalité large à gauche, et que l'on a donc par symétrie :

$$(z_1 + t_1 = z_2 + t_2) \implies (a(z_1) + b(t_1) = a(z_2) + b(t_2)). \quad (8.1)$$

Conclusion. Supposons d'abord $a(0) = b(0) = 0$. Dans ce cas, pour $(z_1, t_1, z_2, t_2) = (z, 0, 0, z)$ on obtient $a = b$, puis pour $(z + t, 0, z, t)$ l'égalité d'additivité

$$a(z + t) = a(z) + a(t)$$

d'où l'on déduit classiquement (en utilisant la croissance) que $a(z) = b(z) = cz$. Dans le cas général on sait donc qu'il existe a_0, b_0 et c telles que : $a(z) = a_0 + cz$, $b(z) = b_0 + cz$. Comme $X = a(Z)$ et $Y = b(T)$, X et Y sont des fonctions affines de lois normales, et suivent donc toutes deux une loi normale. En considérant les espérances et les variances, on peut même montrer que $a_0 = b_0 = 0$ et $c = 1$, de sorte que $X = Z$ et $Y = T$. \square

c. Preuve du TLC.

La preuve utilise deux lemmes que nous démontrerons plus loin. Le premier est très classique :

Lemme 8.5 (Lemme sous-additif de Fekete). *Si la suite $(v_n)_{n \in \mathbb{N}}$ de nombres positifs vérifie :*

$$\forall m, n, \quad v_{m+n} \leq v_m + v_n,$$

alors (v_n/n) converge vers $\inf_{p \in \mathbb{N}} (v_p/p)$.

Le second est un résultat d'uniforme intégrabilité.

Lemme 8.6. *Soit $(X_i)_{i \in \mathbb{N}}$ une suite de variables i.i.d. centrées et soit $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$. Si X_1 est de carré intégrable, alors $\mathbb{E}[S_n] = \mathbb{E}[X_1^2]$ et la suite (S_n^2) est uniformément intégrable.*

Venons-en maintenant à la preuve du théorème ; nous procédons par étapes.

Convergence de D le long d'une sous-suite. Soit $T_n = S_{2^n} = \frac{1}{\sqrt{2^n}} \sum_{i=1}^{2^n} X_i$. Notons \tilde{T}_n une copie indépendante de T_n . En découpant :

$$T_{n+1} = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2^n}} \sum_{i=1}^{2^n} X_i + \frac{1}{\sqrt{2^n}} \sum_{i=2^n+1}^{2^{n+1}} X_i \right),$$

on voit que $T_{n+1} \stackrel{(d)}{=} \frac{1}{\sqrt{2}}(T_n + \tilde{T}_n)$, d'où

$$\begin{aligned} D(T_{n+1}) &= D\left(\frac{1}{\sqrt{2}}(T_n + \tilde{T}_n)\right) \\ &\leq \frac{1}{2} \left(D(T_n) + D(\tilde{T}_n) \right) \\ &\leq D(T_n). \end{aligned}$$

grâce à la sous-additivité et le fait que \tilde{T}_n a même loi que T_n . La suite $D(T_n)$ est donc décroissante, et comme elle est positive elle converge vers une limite $l \geq 0$.

Identification de la limite. La suite $(T_n)_{n \in \mathbb{N}}$ est tendue (car bornée dans L^2). On peut donc trouver une extraction ϕ telle que $(T_{\phi(n)})$ converge en loi vers une limite Z . Comme $T_{\phi(n)}$ est une sous-suite de S_n , $(T_{\phi(n)}^2)$ est uniformément intégrable, donc $\text{Var}(T_{\phi(n)})$ converge vers $\text{Var}(Z)$, ce qui implique que $D(T_{\phi(n)})$ tend vers $D(Z)$ par le théorème 8.4. Par le point précédent, $D(T_{\phi(n)})$ tend vers l , donc $D(Z) = l$. De même $D(\tilde{Z}) = l$. Comme les suites (T_n) et (\tilde{T}_n) sont indépendantes et de même loi, le couple $(T_{\phi(n)}, \tilde{T}_{\phi(n)})$ converge en loi vers (Z, \tilde{Z}) où \tilde{Z} est une copie indépendante de Z , donc $(T_{\phi(n)} + \tilde{T}_{\phi(n)})$ converge en loi vers $Z + \tilde{Z}$. Comme les variances convergent également, on peut passer à la limite dans l'égalité

$$\frac{1}{2}D(T_{\phi(n)} + \tilde{T}_{\phi(n)}) = D(T_{\phi(n)+1})$$

pour obtenir

$$\frac{1}{2}(D(Z + \tilde{Z})) = l,$$

en réutilisant la propriété de convergence du théorème 8.4. Comme on a aussi $l = \frac{1}{2}(D(Z) + D(\tilde{Z}))$, on en déduit que $D(Z + \tilde{Z}) = D(Z) + D(\tilde{Z})$, donc Z (et \tilde{Z}) suivent une loi normale ; par conséquent $l = D(Z) = 0$.

Conclusion. On sait maintenant que $D(T_n) = D(S_{2^n})$ tend vers 0, il s'agit de montrer que $D(S_n)$ tend vers 0.

Posons $v_n = nD(S_n) = D(\sum_{i=1}^n X_i)$. La sous-additivité de D montre que v_n est sous-additive, donc $D(S_n) = v_n/n$ converge vers $\inf D(S_p)$ par le lemme de Fekete. Comme $D(S_{2^n})$ tend vers 0, cette borne inférieure est nulle, et $D(S_n)$ tend bien vers 0.

d. Preuve des lemmes.

Preuve du lemme de Fekete Posons $u_n = v_n/n$. Pour cette suite,

$$\forall m, n, \quad u_{m+n} \leq \frac{1}{m+n}(v_m + v_n) \leq \frac{m}{m+n}u_m + \frac{n}{m+n}u_n.$$

On remarque d'abord que $u_{m+n} \leq \max(u_m, u_n)$, ce qui entraîne $u_{kn} \leq u_n$ par une récurrence immédiate (sur k). Fixons un entier p , et posons $n = pq_n + r_n$ la division euclidienne de n par p . Alors

$$\begin{aligned} u_n &= u_{pq_n+r_n} \leq \frac{pq_n}{pq_n+r_n}u_{pq_n} + \frac{r_n}{pq_n+r_n}u_{r_n} \\ &\leq u_{pq_n} + \frac{p}{pq_n+r_n}u_{r_n} \\ &\leq u_p + \frac{1}{q_n}u_{r_n} \\ &\leq u_p + \frac{1}{q_n}u_1. \end{aligned}$$

Comme $q_n \sim n/p$ tend vers l'infini on en déduit $\limsup_n u_n \leq u_p$. Ceci étant valable pour tout p ,

$$\limsup_n u_n \leq \inf_p u_p \leq \liminf_n u_n \leq \limsup_n u_n,$$

donc ces quantités sont égales, et $u_n = (v_n/n)$ converge vers $\inf_p u_p = \inf_p (v_p/p)$.

Preuve du lemme d'uniforme intégrabilité. Rappelons que $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$, et supposons sans perte de généralité que $\mathbb{E}[X_1^2] = 1$. En utilisant l'indépendance on a directement :

$$\mathbb{E}[S_n^2] = \frac{1}{n} \sum_i \mathbb{E}[X_i^2] = 1.$$

Pour montrer l'uniforme intégrabilité on utilise un procédé de troncature assez fin.

Lemme 8.7 (Décomposition). *Soit X une v.a. centrée réduite et $\varepsilon > 0$. Il existe des variables (B, R) telles que $X \stackrel{(d)}{=} B + R$ et :*

- B et R sont centrées,
- $\mathbb{E}[B^4] = M(\varepsilon) < \infty$ et $\text{Var}(R) \leq \varepsilon$.

Démonstration. Pour tout K posons X_K la troncature $X_K = (X \wedge K) \vee (-K)$. Comme X_K converge vers X pour $K \rightarrow \infty$ et est dominée par $|X|$ de carré intégrable, on a par convergence dominée $\text{Var}(X - X_K) \rightarrow 0$. Choisissons K tel que cette variance soit inférieure à ε . En posant $B = X_K - \mathbb{E}[X_K]$ et $R = X - B = X - X_K + \mathbb{E}[X_K]$, on obtient bien la décomposition voulue puisque B est bornée par $2K$. \square

Revenons à l'uniforme intégrabilité. Pour tout i on introduit la décomposition $X_i = B_i + R_i$, et on majore :

$$\begin{aligned} \mathbb{E}[S_n^2 \mathbf{1}_{|S_n| > R}] &= \frac{1}{n} \mathbb{E} \left[\left(\sum_{i=1}^n B_i + \sum_{i=1}^n R_i \right)^2 \mathbf{1}_{|S_n| > R} \right] \\ &\leq \frac{2}{n} \mathbb{E} \left[\left(\sum_{i=1}^n B_i \right)^2 \mathbf{1}_{|S_n| > R} \right] + \frac{2}{n} \mathbb{E} \left[\left(\sum_{i=1}^n R_i \right)^2 \mathbf{1}_{|S_n| > R} \right]. \end{aligned}$$

Dans le premier terme, appliquons l'inégalité de Cauchy-Schwarz ; dans le second, majorons l'indicatrice par 1 : on reconnaît alors la variance d'une somme puisque les R_i sont centrées. Par conséquent :

$$\begin{aligned} \mathbb{E}[S_n^2 \mathbf{1}_{|S_n| > R}] &\leq \frac{2}{n} \mathbb{E} \left[\left(\sum_{i=1}^n B_i \right)^4 \right]^{1/2} \mathbb{P}[|S_n| > R]^{1/2} + \frac{2}{n} \text{Var} \left(\sum_{i=1}^n R_i \right) \\ &\leq \frac{2}{Rn} \mathbb{E} \left[\left(\sum_{i=1}^n B_i \right)^4 \right]^{1/2} + 2\varepsilon, \end{aligned}$$

où l'on a utilisé l'inégalité de Bienaymé-Tchebicheff à la dernière ligne. Pour contrôler le moment d'ordre 4 on développe :

$$\mathbb{E} \left[\left(\sum_{i=1}^n B_i \right)^4 \right] = \sum_{i,j,k,l} \mathbb{E}[B_i B_j B_k B_l].$$

Dans cette somme, tous les termes contenant un indice non-répété disparaissent puisque les B_i sont indépendants et centrés. Par conséquent

$$\begin{aligned}
\mathbb{E} \left[\left(\sum_{i=1}^n B_i \right)^4 \right] &= \sum_i \mathbb{E} [B_i^4] + \sum_{i,j,j \neq i} \mathbb{E} [B_i^2] \mathbb{E} [B_j^2] \\
&= n \mathbb{E} [B_1^4] + n(n-1) \mathbb{E} [B_1^2]^2 \\
&\leq 2n^2 \mathbb{E} [B_1^4] = 2n^2 M(\varepsilon).
\end{aligned}$$

En reportant cette majoration, il vient

$$\mathbb{E} [S_n^2 \mathbf{1}_{|S_n| > R}] \leq \frac{2\sqrt{2}}{R} M(\varepsilon)^{1/2} + 2\varepsilon,$$

Le terme de droite ne dépend pas de n , donc

$$\limsup_{R \rightarrow \infty} \sup_n \mathbb{E} [S_n^2 \mathbf{1}_{|S_n| > R}] \leq 2\varepsilon,$$

ce qui montre l'uniforme intégrabilité puisque ε est arbitraire.

8.2 Convergence pour les chaînes de Markov par couplage

a. Le résultat classique

Si $(X_t)_{t \in \mathbb{N}}$ est une chaîne de Markov sur un espace d'états E dénombrable, on a le résultat classique suivant de convergence en loi.

Théorème 8.8 (Convergence en loi des chaînes de Markov). *Si la chaîne $(X_t)_{t \in \mathbb{N}}$ est irréductible, apériodique et récurrente positive, alors pour toute loi initiale μ_0 , X_t converge en loi vers la probabilité invariante ν :*

$$\mathbb{P}_{\mu_0} [X_t = x] \rightarrow \nu(x).$$

Une des preuves classiques de ce résultat se fait par *couplage* : on construit un processus (X_t, \tilde{X}_t) sur l'espace produit $E \times E$ tel que :

1. X_t a la loi de la chaîne d'origine, partant d'une loi initiale μ_0 ,
2. \tilde{X}_t a la loi de la chaîne d'origine, partant d'une loi initiale ν_0 ,
3. $\mathbb{P} [X_t \neq \tilde{X}_t]$ tend vers 0.

On en déduit que la distance en variation totale entre les lois de X_t et \tilde{X}_t tend vers 0. En prenant la loi invariante ν comme loi initiale pour \tilde{X}_0 , ceci implique que X_t converge vers ν en variation totale.

b. La chaîne M/M/ ∞

On modélise une file d'attente en temps discret de la façon suivante :

- Au départ, il y a x_0 personnes en attente.
- À chaque étape de temps, chaque personne en attente lance une pièce : avec probabilité p elle est servie et sort de la file, avec probabilité $q = 1 - p$ elle reste dans la file ; puis . . .
- de nouveaux clients se rajoutent à la file : le nombre de clients suit une loi de Poisson de paramètre λ .

Toutes les variables aléatoires sous-jacentes sont supposées indépendantes.

Remarque 8.1. Chaque client reste dans la file pendant un temps géométrique, indépendamment des autres clients ; tout se passe comme si chaque client était servi par une caisse différente, ou encore comme s'il y avait une infinité de caisses/serveurs disponibles pour servir les clients.

Les départs se font de manière Markovienne : la probabilité qu'un client parte entre t et $t+1$ ne dépend pas du temps qu'il a passé dans le système.

Les arrivées se font de manière Markovienne : on peut imaginer un processus de Poisson à temps continu sous-jacent.

Ces propriétés expliquent la terminologie barbare « $M/M/\infty$ » (départs et arrivées Markovienne avec une infinité de serveurs).

Théorème 8.9 (Probabilités de transition). Les probabilités de transition de la chaîne $M/M/\infty$ sont données par

$$p_{k,l} = \mathbb{P}[X_1 = l | X_0 = k] = \sum_{j=0}^{k \wedge l} \frac{k!}{j!(k-j)!(l-j)!} q^j p^{k-j} e^{-\lambda} \lambda^{l-j}.$$

Démonstration. On conditionne par les valeurs de Y , le nombre de personnes qui ne sont pas servies parmi les k présentes à l'instant 0 : si $Y = j$, il faut donc exactement $l-j$ arrivées de nouveaux clients pour avoir $X_1 = l$. Y suit la loi binomiale $\text{Bin}(k, q)$, donc

$$\begin{aligned} p_{k,l} &= \sum_{j=0}^k \mathbb{P}[X_1 = l | X_0 = k, Y = j] \mathbb{P}[Y = j | X_0 = k] \\ &= \sum_{j=0}^k \mathbf{1}_{j \leq l} e^{-\lambda} \frac{\lambda^{l-j}}{(l-j)!} \times \binom{k}{j} q^j p^{k-j} \end{aligned}$$

ce qui donne l'expression annoncée. □

Théorème 8.10. La chaîne est irréductible apériodique récurrente positive, de probabilité invariante $\nu = \text{Poi}(\lambda/p)$.

Démonstration. L'irréductibilité est claire, puisqu'on peut même passer de tout état x à tout état y en un seul coup. L'apériodicité est également claire puisque 0 mène à lui-même. Pour vérifier l'invariance il suffit de vérifier la réversibilité, on calcule donc

$$\begin{aligned} \nu(k) p_{kl} &= e^{-\lambda/p} \frac{\lambda^k}{p^k k!} p_{k,l} \\ &= e^{-\lambda/p - \lambda} \sum_{j=0}^{k \wedge l} \frac{\lambda^{k+l-j} p^{-j} q^j}{j!(k-j)!(l-j)!} \\ &= \nu(l) p_{lk} \end{aligned}$$

puisque l'expression à l'avant dernière ligne est symétrique en k et l . □

À ce stade, on sait par le résultat de convergence en loi que la chaîne converge vers la loi de Poisson ν , quelle que soit la loi initiale. En construisant explicitement un couplage, nous allons donner une vitesse de convergence pour la distance de Wasserstein.

Théorème 8.11 (Convergence exponentielle en distance de Wasserstein). Pour toute loi de départ μ_0 ayant un moment d'ordre 1, pour tout $t \in \mathbb{N}$,

$$W_1(\mu_t, \nu) \leq q^t W_1(\mu_0, \nu).$$

Démonstration. On va construire, partant d'une loi initiale (X_0, Y_0) , un couplage $(X_t, Y_t)_{t \in \mathbb{N}}$ de deux copies du processus. L'idée de ce couplage particulier est de choisir autant que possible le même aléa pour les deux processus.

Pour cela, donnons-nous une famille $(B_{t,n})$ de variables de Bernoulli indépendantes de paramètre q , et une famille (A_t) de variables de loi de Poisson de paramètre λ :

- A_t correspond au nombre d'arrivées au temps t ,
- $B_{t,n} = 0$ si le n^e client au temps t est servi, 1 sinon.

Soit (X_0, Y_0) un couple de variables quelconque à valeurs dans $\mathbb{N} \times \mathbb{N}$. Définissons par récurrence :

$$X_{t+1} = \sum_{k=1}^{X_t} B_{t+1,k} + A_{t+1},$$

$$Y_{t+1} = \sum_{k=1}^{Y_t} B_{t+1,k} + A_{t+1}.$$

On voit aisément que (X_t) et (Y_t) , pris seuls, sont des chaînes M/M/ ∞ . Remarquons que l'on utilise les mêmes variables A_{t+1} pour définir les arrivées dans les deux processus, et, dans la mesure du possible, les mêmes variables pour décider des départs. Ceci se traduit par la disparition de nombreux termes dans la différence :

$$|X_{t+1} - Y_{t+1}| = \sum_{k=\min(X_t, Y_t)+1}^{\max(X_t, Y_t)} B_{t+1,k}.$$

En notant \mathcal{F}_t la tribu engendrée par les $(B_{s,n}, A_s)_{s \leq t, n}$, on en déduit

$$\mathbb{E}[|X_{t+1} - Y_{t+1}| | \mathcal{F}_t] = q |X_t - Y_t|.$$

d'où, en reprenant l'espérance et après une récurrence immédiate,

$$\mathbb{E}[|X_t - Y_t|] = q^t \mathbb{E}[|X_0 - Y_0|].$$

Donnons-nous maintenant une loi initiale μ_0 admettant un moment d'ordre 1. Soit (X_0, Y_0) un couple réalisant la distance $W_1(\mu_0, \nu)$. Alors (X_t, Y_t) est un couplage de μ_t et ν (puisque ν est invariante), et

$$W_1(\mu_t, \nu) \leq \mathbb{E}[|X_t - Y_t|] \leq q^t \mathbb{E}[|X_0 - Y_0|] = q^t W_1(\mu_0, \nu).$$

□

8.3 Remarques bibliographiques

La preuve du théorème limite central *via* la distance de Wasserstein est due à Tanaka [Tan73]. Nous avons globalement repris la présentation de [JS05], qui mentionne la parenté de la méthode avec des résultats de Rényi. La preuve de l'uniforme intégrabilité de S_n^2 suit [Bro82, Lemme 5.2] On pourra consulter [CM89] pour une généralisation de l'approche en dimension quelconque, et [JS05] pour des généralisations à d'autres distances de Wasserstein et à des convergences vers des lois stables, ainsi que pour des considérations sur la vitesse de convergence de $D(S_n)$ vers 0.

Chapitre 9

Équation de Monge-Ampère et transport de Brenier

9.1 Introduction

L'objectif de cette dernière partie du cours est d'établir des inégalités fonctionnelles fondamentales sur \mathbb{R}^n , en utilisant le transport optimal comme un outil.

Nous établissons d'abord dans ce chapitre l'équation de Monge-Ampère qui résulte du changement de variable lié au transport de Brenier.

Le chapitre 10 concerne l'inégalité de Brunn-Minkowski et l'inégalité de Prékopa-Leindler, forme inverse ou duale de l'inégalité de Hölder.

Dans le chapitre 11 nous démontrons l'inégalité isopérimétrique et les inégalités de Sobolev L^p euclidiennes.

Enfin le chapitre 12 est consacré à la preuve de l'inégalité de Young optimale comme cas particulier de l'inégalité de Brascamp-Lieb. Nous montrons simultanément l'inégalité de Brascamp-Lieb inverse qui généralise l'inégalité de Prékopa-Leindler.

9.2 L'équation de Monge-Ampère dans un cas simple

Rappelons tout d'abord le théorème de Brenier. On se place dans l'espace euclidien \mathbb{R}^n et l'on considère le coût quadratique $c_2(x, y) = \frac{1}{2}|x - y|^2$, qui induit le coût de transport entre mesures :

$$\mathcal{T}_2(\mu, \nu) = \frac{1}{2} \min_{\pi \in \Pi(\mu, \nu)} \iint |x - y|^2 \pi(dx, dy)$$

Pour des mesures absolument continues, ce coût est réalisé par un transport déterministe donné par le gradient d'une fonction convexe :

Théorème 9.1 (Brenier). *Soient μ, ν deux mesures de probabilité sur \mathbb{R}^n équipé de sa norme euclidienne canonique admettant des moments d'ordre 2 ; si μ est absolument continue par rapport à la mesure de Lebesgue alors il existe un unique plan de transport optimal $\pi^* \in \Pi(\mu, \nu)$.*

Ce plan de transport optimal est déterministe : il existe une application de transport T^ de μ sur ν , unique μ presque sûrement, telle que $\pi^*(dxdy) = \mu(dx)\delta_{T^*(x)}(dy)$.*

$$\mathcal{T}_2(\mu, \nu) = \frac{1}{2} \int |x - T^*(x)|^2 \mu(dx).$$

De plus, cette application T^* s'écrit sous la forme $T^* = \nabla\phi$, où $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction convexe.

Puisque T^* transporte μ sur ν , $T^*\#\mu = \nu$, pour toute fonction borélienne bornée H sur \mathbb{R}^n ,

$$\int H(y)d\nu(y) = \int H(T^*(x))d\mu(x).$$

Supposons que μ et ν sont des mesures absolument continues par rapport à la mesure de Lebesgue, de densité respective f et g . L'égalité précédente s'écrit encore

$$\int H(y)g(y)dy = \int H(T^*(x))f(x)dx.$$

Faisons pour l'instant les deux hypothèses simplificatrices suivantes :

1. $T^* = \nabla\phi$ est bijective de \mathbb{R}^n dans \mathbb{R}^n ,
2. T^* est de classe \mathcal{C}_1 sur \mathbb{R}^n (autrement dit ϕ est de classe \mathcal{C}^2).

Le changement de variable $y = T^*(x)$ dans la première intégrale donne, pour toute fonction H mesurable bornée

$$\int H(T^*(x))g(T^*(x))|\det(J_{T^*}(x))|dx = \int H(T^*(x))f(x)dx,$$

où $J_{T^*}(x)$ est la matrice jacobienne de T^* au point x . Par conséquent, on obtient pour presque tout $x \in \mathbb{R}^n$,

$$f(x) = g(T^*(x))|\det(J_{T^*}(x))|.$$

Puisque $T^* = \nabla\phi$ et ϕ est une fonction convexe $J_{T^*}(x)$ est en fait la hessienne de ϕ , $\text{Hess}_x\phi$, une matrice semi-définie positive. Par conséquent l'application ϕ satisfait l'équation de Monge-Ampère suivante, pour presque tout x :

$$f(x) = g(\nabla\phi(x))\det(\text{Hess}_x\phi). \tag{9.1}$$

Cette équation sera un élément essentiel des preuves des inégalités fonctionnelles présentées dans les chapitres suivants.

Malheureusement nos hypothèses simplificatrices sont beaucoup trop restrictives, et ne sont en général pas réalisées. Le théorème de Brenier indique en effet uniquement que $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction convexe : ceci entraîne comme on l'a vu que $\nabla\phi$ est défini presque partout, mais ne dit rien sur sa bijectivité ni sa régularité. L'objet de la suite de ce chapitre est de montrer que l'on peut encore donner un sens à l'équation de Monge-Ampère (9.1) sous ces conditions plus faibles.

Théorème 9.2. *Soit μ et ν des mesures absolument continues par rapport à la mesure de Lebesgue de densités respectives f et g , et $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe telle que $\nu = \nabla\phi\#\mu$. Alors, pour μ -presque tout x , ϕ est deux fois différentiable au sens d'Aleksandrov en x et on a l'équation de Monge-Ampère :*

$$f(x) = g(\nabla\phi(x))\det(\text{Hess}_x\phi).$$

De plus l'application $x \rightarrow \det(\text{Hess}_x\phi)$ est localement intégrable sur l'intérieur de son domaine $\Omega = \text{int}(\text{dom}\phi)$, $\text{dom}\phi = \{x \in \mathbb{R}^n, \phi(x) < +\infty\}$.

La preuve de ce théorème est relativement technique et nécessite des outils d'analyse convexe et de théorie de la mesure. Ces outils sont présentés dans les trois prochaines sections. La preuve proprement dite est donnée au paragraphe 9.6 (cf. Théorème 9.21).

9.3 Outils d'analyse convexe : transformée de Legendre et dualité convexe

Définition 9.3 (Transformée de Legendre.). Soit $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, $\psi \neq +\infty$ ($\text{dom } \psi \neq \emptyset$). La transformée de Legendre¹ de ψ , notée $\psi^* : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est l'application définie pour $p \in \mathbb{R}^n$ par

$$\psi^*(p) = \sup_{x \in \mathbb{R}^n} \{p \cdot x - \psi(x)\}.$$

Les deux lemmes suivants rassemblent des résultats usuels autour de la transformée de Legendre et du sous-gradient.

Lemme 9.4 (Duale convexe d'une fonction arbitraire). Soit $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ telle $\psi \neq +\infty$. On note $\Omega = \text{int}(\text{dom } \psi)$.

1. La duale convexe ψ^* est convexe et s.c.i. Si $\psi^* \neq \infty$, on peut définir sa duale ψ^{**} et on a $\psi^{**} \leq \psi$.
2. Un point p est dans $\partial\psi(x)$ si et seulement si $\psi^*(p) + \psi(x) = x \cdot p$. L'ensemble $\partial\psi(x)$ est convexe fermé.
3. On a l'équivalence :

$$p \in \partial\psi(x) \iff (x \in \partial\psi^*(p) \text{ et } \psi^{**}(x) = \psi(x)).$$

Démonstration. Notons $B_r(x)$ la boule euclidienne de \mathbb{R}^n de centre $x \in \mathbb{R}^n$ et de rayon $r \geq 0$.

1. $\psi^* = \sup_{x \in \mathbb{R}^n} (px - \psi(x)) = \sup_{x \in \text{dom}(\psi)} (px - \psi(x))$ est un supremum de fonctions affines, donc elle est convexe et s.c.i.

Pour tout $x \in \mathbb{R}^n$, on a

$$\psi^{**}(x) = \sup_p \{p \cdot x - \psi^*(p)\} = \sup_p \inf_y \{p \cdot x - p \cdot y + \psi(y)\} \leq \psi(x).$$

où la dernière inégalité s'obtient en choisissant $y = x$.

2. Par définition, on a

$$\begin{aligned} \partial\psi(x) &= \{p \in \mathbb{R}^n, \forall y \in \mathbb{R}^n, \psi(y) \geq \psi(x) + (y - x) \cdot p\} \\ &= \{p \in \mathbb{R}^n, -\psi(x) \geq \psi^*(p) - x \cdot p\} \end{aligned}$$

$\partial\psi(x)$ est un ensemble de niveau de l'application s.c.i. convexe $y \rightarrow \psi^*(p) - x \cdot p$. Cet ensemble est donc convexe et fermé. Il reste à observer que $x \cdot p - \psi(x) \geq \psi^*(p)$ signifie que $x \cdot p - \psi(x) = \psi^*(p)$ puisque l'inégalité inverse est toujours vraie.

3. En appliquant le point précédent à ψ^* , on sait que $x \in \partial\psi^*(p)$ si et seulement si $\psi^{**}(x) + \psi^*(p) = x \cdot p$. On a donc l'implication de droite à gauche. Réciproquement si $p \in \partial\psi(x)$, alors $\psi^{**}(x) \geq px - \psi^*(p) = \psi(x)$, donc il y a bien égalité et x est bien dans $\partial\psi^*(p)$. \square

Si la fonction de départ est convexe, on peut dire beaucoup plus de sa transformée. Dans toute la suite, pour tout sous-ensemble M de \mathbb{R}^n et $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, on notera

$$\partial\psi(M) = \bigcup_{x \in M} \partial\psi(x).$$

1. Cette fonction est aussi appelée *duale convexe* ou encore transformée de Fenchel.

Lemme 9.5 (Duale convexe d'une fonction convexe). Soit $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe telle $\psi \neq +\infty$. On note $\Omega = \text{int}(\text{dom } \psi)$.

1. La fonction ψ est lipschitzienne sur tout compact inclus dans Ω . En particulier elle est continue et localement lipschitzienne sur Ω .
2. Si $K \subset \Omega$ est compact et non-vide alors $\partial\psi(K)$ est compact et non-vide. En particulier $\partial\psi(x)$ est convexe compact et non-vide quand $x \in \Omega$.
3. Continuité du sous-gradient : Soit x_0 un point de différentiabilité de ψ . Alors l'application $x \mapsto \sup_{p \in \partial\psi(x)} |p - \nabla\psi(x_0)|$ est continue en x_0 .
4. $\psi^* \neq \infty$ et sa duale ψ^{**} est bien définie. ψ est s.c.i en x_0 si et seulement $\psi^{**}(x_0) = \psi(x_0)$. En particulier, pour tout $x \in \Omega$, $\psi^{**}(x) = \psi(x)$ et on a l'équivalence

$$p \in \partial\psi(x) \iff x \in \partial\psi^*(p);$$

si ψ est s.c.i. ces propriétés sont vraies partout.

5. Pour tout $A \subset \Omega$ alors $\partial\psi(A) \subset B \implies A \subset \partial\psi^*(B)$.

Démonstration. 1. Montrons d'abord que ψ est localement bornée sur Ω . Soit x un point de Ω et r assez petit pour avoir $B_r^\infty(x) = \prod_i]x_i - r, x_i + r[\subset \Omega$. Tout point de cette boule étant clairement combinaison convexe des 2^n points extrémaux, ψ est majorée sur cette boule par sa valeur maximale M en ces points extrémaux. Il reste à minorer ψ . Pour y quelconque dans la boule, le symétrique $\tilde{y} = 2x - y$ de y par rapport à x est encore dans la boule. Par convexité, $\psi(x) \leq \frac{1}{2}(\psi(y) + \psi(\tilde{y}))$ donc $\psi(y) \geq 2\psi(x) - M$. La fonction ψ est donc bien localement bornée sur Ω ; par compacité, elle est donc bornée sur tout compact inclus dans Ω .

Soit maintenant K un compact inclus dans Ω et montrons que ψ est Lipschitzienne sur K .

Montrons d'abord le cas unidimensionnel. Soit $K = [a, b]$ un compact de la droite, inclus dans l'intervalle Ω . Pour ε suffisamment petit, $a_\varepsilon = a - \varepsilon$ et $b_\varepsilon = b + \varepsilon$ sont dans Ω . Pour tout $y \leq z$ dans K , on compare les pentes de ψ sur les trois intervalles $[a_\varepsilon, z]$, $[y, z]$ et $[y, b_\varepsilon]$: par convexité,

$$\frac{\psi(z) - \psi(a_\varepsilon)}{z - a_\varepsilon} \leq \frac{\psi(z) - \psi(y)}{z - y} \leq \frac{\psi(b_\varepsilon) - \psi(y)}{b_\varepsilon - y},$$

d'où

$$|\psi(z) - \psi(y)| \leq \frac{2M}{\varepsilon} |z - y|$$

où M , borne de $|\psi|$ sur $[a_\varepsilon, b_\varepsilon]$, existe par le point précédent.

Revenons au cas général. Quitte à remplacer K par son enveloppe convexe, qui est un convexe compact² inclus dans Ω ³, on peut supposer que K est convexe compact.

Soit K_ε le grossissement $K_\varepsilon = \{x \in \mathbb{R}^n, d(x, K) \leq \varepsilon\}$. Cet ensemble est compact, et est inclus dans Ω pour ε petit (l'application $x \mapsto d(x, \Omega^c)$ étant soit infinie, soit continue sur le compact K et donc minorée); ψ y est donc majorée par une constante M par le point précédent.

Soit x, y deux points de K . Paramétrons l'unique droite D qui passe par x et y par $t \mapsto x_t = x + tv$ où v est unitaire. Alors $D \cap K$ est convexe compact, c'est donc un segment de la forme $[x(a), x(b)]$, et on a $[x(a - \varepsilon), x(b + \varepsilon)] \subset D \cap K_\varepsilon$. Comme $\phi : t \mapsto \psi(x(t))$ est convexe et bornée par M sur l'intervalle $[a - \varepsilon, b + \varepsilon]$, la conclusion suit en appliquant le cas unidimensionnel (notons que M et ε ne dépendent pas du choix de la droite D).

2. Pour voir que $\text{co}(K)$ est compacte, on peut faire appel au théorème de Carathéodory, qui affirme que $\text{co}(K) = \{\sum_{i=0}^n \lambda_i x_i, \lambda \in S, x_0, \dots, x_n \in K\}$ où S est le simplexe $\{\lambda = (\lambda_0, \dots, \lambda_n) \in \mathbb{R}_+^{n+1}, \sum \lambda_i = 1\}$. L'ensemble $\text{co}(K)$ est donc bien compact comme image continue de $S \times K^{n+1}$ par $f : (\lambda, x_0, \dots, x_n) \mapsto \sum_{i=0}^n \lambda_i x_i$.

3. puisque Ω est convexe

2. Soit K un compact inclus dans Ω .

Montrons que $\partial\psi(K)$ est borné. Comme précédemment on peut supposer que K est convexe, et introduire un grossissement K_ε tel que $K \subset K_\varepsilon \subset \Omega$. Par le point précédent on sait que ψ est Lipschitzienne sur K_ε pour une certaine constante C . Soit $p \in \partial\psi(x)$ un élément de $\partial\psi(K)$. Par définition,

$$\forall y \in \mathbb{R}^n, \psi(y) \geq \psi(x) + p(y - x).$$

Pour $y = x + tp$, et pour $|t| \leq \varepsilon$, on obtient $t|p|^2 \leq C|t||p|$ puisque x et y sont dans K_ε . On en déduit $|p| \leq C$: $\partial\psi(K)$ est bien borné.

Montrons maintenant que $\partial\psi(K)$ est fermé. Soit p_n une suite de points de $\partial\psi(K)$ qui converge vers p , et notons x_n des points de K tels que $p_n \in \partial\psi(x_n)$. Quitte à extraire une sous-suite on peut supposer que x_n converge vers un point $x \in K$. Par définition du sous-différentiel on a pour tout n

$$\forall y \in \mathbb{R}^n, \psi(y) \geq \psi(x_n) + p_n(y - x_n).$$

La fonction ψ étant continue sur K , on peut passer à la limite et obtenir

$$\forall y \in \mathbb{R}^n, \psi(y) \geq \psi(x) + p(y - x),$$

donc $p \in \partial\psi(x)$ appartient bien à $\partial\psi(K)$.

Le caractère non-vide sera établi plus bas, dans la première étape de la preuve du point 4.

3. Soit x_0 un point de différentiabilité de ψ . On cherche à montrer que la fonction $S : x \mapsto \sup_{p \in \partial\psi(x)} |p - \nabla\psi(x_0)|$ est continue en x_0 . Soit x_n une suite qui tend vers x_0 : on peut supposer que x_n est à valeurs dans une petite boule fermée $K \subset \Omega$. L'ensemble $\partial\psi(x_n)$ est compact par le point précédent donc le sup dans la définition de $S(x_n)$ est atteint en un certain p_n . Il nous reste à montrer que p_n tend vers $\nabla\psi(x_0)$.

Pour cela écrivons une fois de plus la définition :

$$\forall y \in \mathbb{R}^n, \psi(y) \geq \psi(x_n) + p_n(y - x_n).$$

Comme p_n est à valeurs dans $\partial\psi(K)$ qui est compact par le point précédent, on peut extraire de toute sous-suite de (p_n) une sous-sous-suite convergente. La limite p d'une telle sous-sous-suite vérifie

$$\forall y \in \mathbb{R}^n, \psi(y) \geq \psi(x) + p(y - x)$$

en passant à la limite dans l'inégalité précédente, donc nécessairement $p \in \partial\psi(x_0)$. Comme ψ est supposée différentiable en x_0 cet ensemble est le singleton $\{\nabla\psi(x_0)\}$. On en déduit que la suite p_n d'origine converge vers $\nabla\psi(x_0)$, ce qu'on voulait établir.

4. *1^{re} étape.* Nous allons d'abord montrer que si $\psi(x_0) < \infty$ et si ψ est s.c.i. en x_0 , alors $\psi^* \neq \infty$ et $\psi^{**}(x_0) = \psi(x_0)$. Au vu du dernier point du lemme 9.4, il suffit pour cela de montrer que le sous-différentiel $\partial\psi(x_0)$ n'est pas vide.

Rappelons que l'épigraphe de ψ est l'ensemble des points de $\mathbb{R}^n \times \mathbb{R}$ « au dessus du graphe de ψ », c'est-à-dire :

$$\text{epi}(\psi) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}, \psi(x) \leq \alpha\}.$$

Notons $F = \overline{\text{epi}(\psi)}$ la fermeture de l'épigraphe de ψ . Puisque ψ est convexe, F est un convexe fermé de $\mathbb{R}^n \times \mathbb{R}$. Pour $\alpha < \psi(x_0)$, l'ensemble fermé F ne rencontre pas le compact $\{(x_0, \alpha)\}$:

si c'était le cas, il existerait une suite (x_k, β_k) de points de $\text{epi}(\psi)$ tendant vers (x_0, α) , et par s.c.i. de ψ en x_0 on aurait

$$\psi(x_0) \leq \liminf_k \psi(x_k) \leq \lim_k \beta_k = \alpha,$$

ce qui est exclu.

Posons alors $\alpha_n = \psi(x_0) - 1/n$. D'après le théorème de Hahn-Banach, il existe un hyperplan affine de $\mathbb{R}^n \times \mathbb{R}$ qui sépare F de $\{(x_0, \alpha_n)\}$. Ceci se traduit par l'existence de $p_n \in \mathbb{R}^n$, $\gamma_n \in \mathbb{R}$ et $\varepsilon_n > 0$ tel que la forme affine $(x, \alpha) \mapsto (x - x_0) \cdot (-p_n) + (\alpha - \alpha_n)\gamma_n$ qui s'annule au point (x_0, α_n) , soit supérieure ou égale à ε_n sur tout l'ensemble F (un hyperplan séparateur étant alors donné par l'image réciproque de $\varepsilon_n/2$ par la forme).

Comme $\psi(x_0)$ est fini, $(x_0, \psi(x_0))$ est dans l'épigraphe donc $(\psi(x_0) - \alpha_0)\gamma_n \geq \varepsilon > 0$. Par conséquent γ_n est strictement positif, et quitte à normaliser p_n et ε_n on peut supposer $\gamma_n = 1$. On obtient alors pour tout x , $(x - x_0) \cdot (-p_n) + (\psi(x) - \alpha_n) \geq 0$, ou encore

$$\psi(x) \geq \alpha_n + p_n(x - x_0).$$

En termes géométriques, ψ est au dessus du plan de pente p passant par (x_0, α_0) . Pour δ assez petit, ψ est bornée (disons par M) sur $B(x_0, \delta)$. En appliquant l'inégalité précédente à $x = x_0 + \delta \frac{p_n}{|p_n|}$ il vient $|p_n| \leq 2M + \frac{1}{\delta}$, donc p_n est bornée. On en extrait une sous-suite qui converge, disons vers p_0 . En passant à la limite le long de la sous-suite on montre que p_0 est dans $\partial\psi(x_0)$, qui n'est donc pas vide.

Deuxième étape. Si Ω n'est pas vide, on peut appliquer la première étape en un point x de Ω pour montrer que $\psi^* \neq \infty$. Si Ω est vide, le domaine de ψ est un convexe d'intérieur vide. On peut montrer qu'il est inclus dans un sous-espace affine strict de \mathbb{R}^n . En raisonnant dans ce sous-espace, on se ramène au cas où Ω n'est pas vide, ce qui permet d'établir $\psi^* \neq \infty$. À ce point de la démonstration, on sait donc que $\psi^* \neq \infty$ dans tous les cas, et que ψ est minorée par une fonction affine.

Troisième étape. On suppose toujours ψ s.c.i. en x_0 , mais avec $\psi(x_0) = \infty$. Comme ψ est s.c.i elle est identiquement infinie sur un voisinage de x_0 . En appliquant Hahn-Banach dans \mathbb{R}^n à $\{x_0\}$ et $\bar{\text{dom}} \psi$, on obtient l'existence de p_0 et ε tels que pour tout $x \in \text{dom} \psi$,

$$p_0 \cdot (x - x_0) \geq \varepsilon.$$

Par ailleurs l'étape précédente donne l'existence d'un p_1 tel que $\psi(x) \geq p_1 \cdot x + C$. Posons $p_\lambda = p_1 + \lambda p_0$ pour tout $\lambda > 0$. Pour tout $x \in \text{dom} \psi$,

$$p_\lambda(x_0 - x) + \psi(x) \geq p_1(x_0 - x) + \lambda p_0(x_0 - x) + p_1 x + C = p_1 x_0 + C + \lambda \varepsilon.$$

Par conséquent

$$\begin{aligned} \psi^{**}(x_0) &= \sup_p \inf_{x \in \text{dom} \psi} \{p \cdot x_0 - p \cdot x + \psi(x)\} \\ &\geq \inf_{x \in \text{dom} \psi} \{p_\lambda \cdot x_0 - p_\lambda \cdot x + \psi(x)\} \\ &\geq p_1 x_0 + C + \lambda \varepsilon. \end{aligned}$$

Ceci étant vrai pour un $\lambda > 0$ arbitraire, $\psi^{**}(x_0) = \infty = \psi(x_0)$.

Quatrième étape : réciproque Supposons $\psi^{**}(x_0) = \psi(x_0)$. Soit (x_n) une suite de points de \mathbb{R}^n telle que $\lim_n x_n = x_0$. Puisque $\psi \geq \psi^{**}$ et ψ^{**} est une fonction s.c.i. on a

$$\liminf_n \psi(x_n) \geq \liminf_n \psi^{**}(x_n) \geq \psi^{**}(x_0) = \psi(x_0).$$

La fonction ψ est donc s.c.i. en x_0 .

Cinquième étape : conclusion Les dernières propriétés annoncées sont faciles à établir : par le premier point du lemme, ψ est continue (et *a fortiori* s.c.i.) sur Ω , et il ne reste plus qu'à appliquer le point 3 du lemme 9.4.

5. Supposons $\partial\psi(A) \subset B$. Soit $x \in A$. Puisque $A \subset \Omega$, il existe $p \in \partial\psi(x)$, pour lequel $x \in \partial\psi^*(p)$. Comme $\partial\psi(A) \subset B$, p est dans B donc $x \in \partial\psi^*(B)$. □

Pour conclure cette section, montrons le théorème des accroissements finis annoncé précédemment :

Lemme 9.6. *Soit ϕ une fonction convexe, Ω l'intérieur du domaine de ϕ et f une fonction \mathcal{C}^1 sur Ω . Alors, pour tout x, y dans Ω , il existe $z \in [x, y]$ et $p \in \partial_\phi(z)$ tels que :*

$$(\phi + f)(y) - (\phi + f)(x) = (p + \nabla f(z)) \cdot (y - x).$$

Démonstration. On paramètre le segment $[x, y]$ par $[0, 1]$ en posant $\tilde{\phi}(t) = \phi((1-t)x + ty)$, et on définit de même $\tilde{f}(t) = f((1-t)x + ty)$. On suit l'idée habituelle de preuve en introduisant la fonction auxiliaire

$$\psi : t \mapsto \tilde{\phi}(t) + \tilde{f}(t) - t \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right].$$

Dans le cas usuel où $(\tilde{\phi} + \tilde{f})$ est dérivable sur $]0, 1[$, on lui applique le théorème de Rolle pour conclure. Ici \tilde{f} est bien dérivable mais $\tilde{\phi}$ n'est que continue. On peut toutefois dire que, comme $\psi(0) = \psi(1)$, ψ admet un extremum en un point intérieur t_0 .

Si c'est un minimum, on a pour tout $t \in [0, 1]$

$$\tilde{\phi}(t) + \tilde{f}(t) - t \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right] \geq \tilde{\phi}(t_0) + \tilde{f}(t_0) - t_0 \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right],$$

soit encore

$$\tilde{\phi}(t) \geq \tilde{\phi}(t_0) - (\tilde{f}(t) - \tilde{f}(t_0)) + (t - t_0) \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right].$$

On en déduit (par la caractérisation locale du sous-gradient) que

$$-\tilde{f}'(t_0) + (\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \in \partial\tilde{\phi}(t_0).$$

Si $\psi(t_0)$ est un maximum, soit p_0 un élément du sous-gradient de $\tilde{\phi}(t_0)$ (un tel élément existe bien puisque t_0 est intérieur). On a alors

$$\begin{aligned} \tilde{\phi}(t_0) + p_0(t - t_0) + \tilde{f}(t) - t \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right] &\leq \tilde{\phi}(t) + \tilde{f}(t) - t \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right] \\ &\leq \tilde{\phi}(t_0) + \tilde{f}(t_0) - t_0 \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right], \end{aligned}$$

ce qui entraîne

$$p_0(t - t_0) \leq -(\tilde{f}(t) - \tilde{f}(t_0)) + (t - t_0) \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right].$$

En prenant successivement les limites à gauche et à droite en t_0 on obtient que

$$p_0 = -\tilde{f}'(t_0) + \left[(\tilde{\phi} + \tilde{f})(1) - (\tilde{\phi} + \tilde{f})(0) \right].$$

Cet élément p_0 appartient donc bien dans tous les cas à $\partial\tilde{\phi}(t_0)$.

À ce moment de la preuve on a donc obtenu l'existence de t_0 et d'un réel $p_0 \in \partial\tilde{\phi}(t_0)$ tel que :

$$\phi(y) + f(y) - (\phi(x) + f(x)) = p_0 + \tilde{f}'(t_0).$$

Comme f est \mathcal{C}^1 , $\tilde{f}'(t_0) = \nabla f(z_0) \cdot (y - x)$ en posant $z_0 = (1 - t_0)x + t_0y$. Il reste donc à montrer que p_0 s'obtient comme produit scalaire d'un élément de $\partial\phi(z_0)$ avec $(y - x)$. Pour cela on remarque que $p_0 \in \partial\tilde{\phi}(t_0)$ signifie que pour tout z dans la droite affine (de \mathbb{R}^n) passant par x et y ,

$$\phi(z) \geq \phi(z_0) + \frac{(z - z_0, y - x)}{\|y - x\|^2} p_0.$$

Par le théorème de Hahn Banach, la forme affine $z \mapsto \frac{(z - z_0, y - x)}{\|y - x\|^2} p_0$ peut être étendue à tout \mathbb{R}^n en conservant la majoration par ψ . En d'autres termes il existe q orthogonal à $y - x$ et $p = p_0 \frac{y - x}{\|y - x\|^2} + q$ tels que $p \in \partial\phi(z_0)$. Alors p_0 s'écrit $p \cdot (y - x)$ pour un $p \in \partial\phi(z_0)$, et l'on conclut

$$\begin{aligned} \phi(y) + f(y) - (\phi(x) + f(x)) &= p_0 + \tilde{f}'(t_0) \\ &= (p + \nabla f(z_0)) \cdot (y - x), \end{aligned}$$

ce qui termine la preuve. □

9.4 Outils d'analyse convexe : une généralisation de la Hessienne

La première observation est que si ϕ n'est pas deux fois différentiable, elle n'admet pas de Hessienne au sens classique du terme. Avant de définir la Hessienne généralisée, rappelons tout d'abord quelques notations.

Le domaine d'une fonction convexe $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ est l'ensemble des points où elle est finie. On notera toujours Ω l'intérieur du domaine de ϕ . L'ensemble des points où ϕ est différentiable est noté Ω_{diff} , il est inclus dans Ω , et l'on a vu que $\lambda(\text{dom}(\phi) \setminus \Omega_{\text{diff}}) = 0$: ϕ est presque partout différentiable sur son domaine.

Le sous-différentiel de f en chaque point $x \in \mathbb{R}^n$ est défini par

$$p \in \partial\phi(x) \forall y \in \mathbb{R}^n, \phi(y) \geq \phi(x) + p \cdot (y - x),$$

c'est géométriquement l'ensemble des pentes⁴ telles que ϕ est au-dessus du plan de pente p passant par $(x, \phi(x))$.

Cet ensemble est vide si $x \notin \text{dom}\phi$. La fonction ϕ est différentiable en x si et seulement si $\partial\phi(x) = \{\nabla\phi(x)\}$.

Dans l'introduction du chapitre, nous avons besoin de supposer que $\nabla\phi$ est différentiable. Nous allons remplacer cette hypothèse de différentiabilité par une hypothèse de régularité du sous-différentiel au voisinage de chaque point x_0 pour lequel $\nabla\phi(x_0)$ existe.

Définition 9.7 (Hessienne au sens d'Aleksandrov). *Soit ϕ une application convexe de \mathbb{R}^n dans $\mathbb{R} \cup \{+\infty\}$, et x_0 un point de $\Omega = \text{int}(\text{dom}\phi)$. On dit que ϕ est deux fois différentiable au sens d'Aleksandrov en x_0 si les deux conditions suivantes sont vérifiées :*

Différentiabilité : ϕ est différentiable (au sens classique) en x_0 ,

DL d'ordre 1 pour le sous-gradient : Il existe une matrice symétrique Λ de taille n telle que :

$$\forall \varepsilon, \exists \delta, \quad |x - x_0| \leq \delta \implies \sup_{p \in \partial\phi(x)} |p - \nabla\phi(x_0) - \Lambda(x - x_0)| < \varepsilon |x - x_0|.$$

La matrice Λ sera notée $\text{Hess}_{x_0}\phi$ et appelée Hessienne de ϕ au sens d'Aleksandrov au point x_0 .

4. Par *plan de pente p passant par (x, α)* on entend le plan affine dans \mathbb{R}^{n+1} , qui passe par (x, α) et est orthogonal à $(p, -1)$. Si ϕ est différentiable en x de gradient $p = \nabla\phi(x)$, on retrouve bien sûr le plan tangent.

Exercice 9.1. Montrer que puisque ϕ est une application convexe, $\text{Hess}_{x_0}\phi$ est semi-définie positive.

L'intérêt de cette notion de différentiabilité au sens d'Aleksandrov pour les fonctions convexes est le résultat suivant.

Théorème 9.8 (Aleksandrov). *Toute fonction convexe $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est presque partout deux fois différentiable au sens d'Aleksandrov sur son domaine.*

On admet ce résultat qui nous permettra d'établir l'équation de Monge-Ampère (9.1) presque partout.

La différentiabilité au sens d'Aleksandrov peut également se traduire par l'existence d'un développement limité d'ordre 2.

Proposition 9.9 (DL d'ordre 2 pour une fonction convexe). *Si une application convexe $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ est deux fois différentiable au sens d'Aleksandrov en un point x_0 , alors*

$$\phi(x_0 + h) = \phi(x_0) + h \cdot \nabla\phi(x_0) + \frac{1}{2} h \cdot \text{Hess}_{x_0}\phi(h) + o(|h|^2)$$

quand $|h|$ tend vers 0.

Nous allons établir ce résultat en admettant provisoirement la généralisation suivante du théorème des accroissements finis.

Lemme 9.10. *Soit ϕ une fonction convexe, Ω l'intérieur du domaine de ϕ et f une fonction \mathcal{C}^1 sur Ω . Alors, pour tout x, y dans Ω , il existe $z \in [x, y]$ et $p \in \partial_\phi(z)$ tels que :*

$$(\phi + f)(y) - (\phi + f)(x) = (p + \nabla f(z)) \cdot (y - x).$$

Preuve de la proposition 9.9. Commençons par nous ramener par transformations affines à un cas plus simple. Comme $x_0 \in \Omega$, ϕ est finie au voisinage de x_0 . Posons pour $h \in \mathbb{R}^n$

$$\tilde{\phi}(h) = \phi(x_0 + h) - \phi(x_0) - \nabla\phi(x_0) \cdot h.$$

La fonction $\tilde{\phi}$ est convexe, finie dans un voisinage de 0, nulle en 0 et différentiable en 0 de gradient nul.

On vérifie facilement que $\partial\tilde{\phi}(h) = \partial\phi(x_0 + h) - \nabla\phi(x_0)$. Comme ϕ est deux fois différentiable en x_0 au sens d'Aleksandrov, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que si $|x - x_0| \leq \delta$ alors

$$\sup_{p \in \partial\phi(x)} |p - \nabla\phi(x_0) - \Lambda(x - x_0)| < \varepsilon|x - x_0|.$$

Autrement dit, en écrivant x comme $x_0 + h$, on a

$$\sup_{p \in \partial\tilde{\phi}(h)} |p - \Lambda(h)| < \varepsilon|h|. \tag{9.2}$$

$\tilde{\psi}$ est donc deux fois différentiable en 0 au sens d'Aleksandrov et $\text{Hess}_0\tilde{\psi} = \Lambda$.

On peut donc sans perte de généralité supposer $x_0 = 0$, $\phi(0) = 0$ et $\nabla\phi(0) = 0$. Notons $\Lambda = \text{Hess}_0\phi$: il s'agit de montrer que $\phi(h) - \frac{1}{2} h \cdot \Lambda(h) = o(|h|^2)$.

On applique le lemme 9.10 à la fonction ϕ et à $f : x \mapsto -\frac{1}{2}x \cdot \Lambda(x)$, entre les points 0 et h . Il existe donc $u \in [0, h]$ et $p \in \partial\phi(u)$ tels que

$$\phi(h) - \frac{1}{2}h \cdot \Lambda(h) = (p - \Lambda(u)) \cdot h.$$

Soit $\varepsilon > 0$. Puisque ϕ est deux fois différentiable au sens d'Aleksandrov en 0, il existe $\delta > 0$ tels que si $|u| \leq \delta$ alors pour tout $q \in \partial\phi(u)$, $|q - \Lambda(u)| < \varepsilon|u|$. Pour $|h| \leq \delta$, on a donc

$$\begin{aligned} \left| \phi(h) - \frac{1}{2}h \cdot \Lambda(h) \right| &= |p - \Lambda(u)) \cdot h| \\ &\leq \varepsilon|u| |h| \leq \varepsilon|h|^2, \end{aligned}$$

ce qui termine la démonstration de la Proposition 9.9. \square

À cet instant, on peut définir pour une fonction ϕ convexe une suite d'ensembles où ϕ est « de plus en plus régulière » : $\text{dom}(\phi)$ est l'ensemble des points où ϕ est finie, Ω_{diff} l'ensemble des points où elle est différentiable, Ω_{A1} l'ensemble des points où elle est deux fois différentiable, et enfin X l'ensemble des points où sa hessienne $\text{Hess } \phi$ (au sens d'Aleksandrov) est inversible. On a donc la chaîne d'inclusions

$$X \subset \Omega_{\text{A1}} \subset \Omega_{\text{diff}} \subset \Omega = \text{int}(\text{dom}(\phi)) \subset \text{dom}(\phi),$$

On notera avec une étoile les ensembles définis similairement pour la fonction duale ϕ^* .

Remarque 9.1. *Tous les ensembles de Ω_{A1} à $\text{dom}(\phi)$ sont introduits pour des raisons techniques : pour ϕ de classe \mathcal{C}^2 ils sont tous confondus, et dans le cas général ils ne diffèrent que par un ensemble Lebesgue-négligeable, par le théorème d'Aleksandrov.*

Le rôle de X est en revanche bien différent comme le montre l'exemple ci-dessous.

Exemple 9.11. *Soit ϕ l'application de \mathbb{R} dans \mathbb{R} définie par $\phi(x) = \mathbf{1}_{x>0}x^2/2$. Déterminer les différents ensembles, de X à $\text{dom}(\phi)$. Déterminer la duale convexe ϕ^* de ϕ . Déterminer l'image par $\nabla\phi$ de la mesure de Lebesgue sur $[-2, 2]$.*

Proposition 9.12 (Une CNS d'inversibilité). *Soit $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe deux fois différentiable au sens d'Aleksandrov au point x_0 . Notons $\Lambda = \text{Hess}_{x_0} \psi$. Alors Λ est inversible si et seulement si ψ^* est deux fois différentiable au sens d'Aleksandrov au point $p_0 = \nabla\psi(x_0)$; quand c'est le cas on a $\text{Hess}_{p_0} \psi^* = \Lambda^{-1}$.*

Corollaire 9.13. *Soit $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ une fonction convexe. Alors*

- $\nabla\psi$ induit une bijection de X sur X^* .
- $\nabla\psi(\Omega_{\text{A1}} \setminus X) \subset (\text{dom}(\psi^*) \setminus \Omega_{\text{A1}}^*)$.

Remarque 9.2. *La fonction $\nabla\phi$ fonctionne d'une certaine manière comme un inverse : si la hessienne est définie et inversible, son inverse l'est aussi, mais si elle est définie et dégénérée, son inverse n'est pas défini.*

Preuve de la proposition 9.12. Préliminaire. On commence par se ramener à un cas simple en posant, comme dans la preuve de la proposition 9.9, $\tilde{\psi}(h) = \psi(x_0 + h) - \psi(x_0) - h \cdot \nabla\psi(x_0)$. Cette fonction est nulle en 0 et y est deux fois différentiable au sens d'Aleksandrov, de gradient nul et de hessienne Λ . De plus, on vérifie simplement que sa duale est donnée par

$$\tilde{\psi}^*(p) = \psi^*(p + p_0) - (p + p_0) \cdot x_0 + \psi(x_0), \quad p \in \mathbb{R}^n,$$

en posant $p_0 = \nabla\psi(x_0)$. Puisque $p_0 \in \partial\psi(x_0) = \{\nabla\psi(x_0)\}$, $\psi(x_0) + \psi^*(p_0) = p_0 \cdot x_0$ donc

$$\tilde{\psi}^*(0) = \psi^*(p_0) - p_0 \cdot x_0 + \psi(x_0) = 0.$$

Quitte à remplacer ψ par $\tilde{\psi}$ on peut donc supposer $x_0 = 0$ et $\psi(0) = \psi^*(0) = \nabla\psi(0) = 0$.

Sens direct — différentiabilité. Supposons la matrice Λ inversible, et montrons que ψ^* est deux fois différentiable au sens d'Aleksandrov, de hessienne Λ^{-1} .

Dans un premier temps, nous devons établir que $\nabla\psi^*(0)$ existe et vaut 0, c'est à dire $\partial\psi^*(0) = \{0\}$. L'inclusion de droite à gauche est claire : comme $0 \in \Omega$ et $0 \in \partial\psi(0)$, on a bien $0 \in \partial\psi^*(0)$.

Inversement si $x \in \partial\psi^*(0)$, montrons que nécessairement $x = 0$. L'ensemble $\partial\psi^*(0)$ est convexe et contient 0 et x . Par suite pour tout $t \in [0, 1]$, $tx \in \partial\psi^*(0)$. Comme $tx \in \Omega$ pour tout t suffisamment petit, ceci entraîne $0 \in \partial\psi(tx)$. Soit $\varepsilon > 0$. Puisque ψ vérifie la propriété (9.2), en choisissant $y = 0 \in \partial\psi(tx)$, pour t suffisamment petit, $|tx| \leq \delta$, on obtient

$$|\Lambda(tx)| < \varepsilon|tx|, \quad |\Lambda(x)| < \varepsilon|x|.$$

Lorsque $\varepsilon \rightarrow 0$, on obtient $\Lambda(x) = 0$, c'est à dire $x = 0$ car on a supposé la matrice Λ inversible.

Sens direct — ordre 2. Soit $\varepsilon > 0$. Pour montrer que ψ^* est deux fois différentiable au sens d'Aleksandrov, sachant que $\nabla\psi^*(0) = 0$, on applique dans un premier temps la propriété de continuité du sous-différentiel (point 3 du lemme 9.5) à ψ^* au point 0. Il existe $\eta > 0$ tel que si $|p| < \eta$ et $x \in \partial\psi^*(p)$ alors $|x| < \delta$ (on peut choisir δ tel que $B_\delta(0) \subset \Omega$). Soit $|p| < \eta$ et $x \in \partial\psi^*(p)$. Puisque $x \in \Omega$, on a $p \in \partial\psi(x)$ et d'après (9.2), $|p - \Lambda(x)| < \varepsilon|x|$. On obtient ainsi

$$|x - \Lambda^{-1}(p)| \leq \|\Lambda^{-1}\| |\Lambda(x) - p| < \varepsilon \|\Lambda^{-1}\| |x| \leq \varepsilon \|\Lambda^{-1}\| (|x - \Lambda^{-1}(p)| + \|\Lambda^{-1}\| |p|)$$

d'où $(1 - \varepsilon \|\Lambda^{-1}\|) |x - \Lambda^{-1}(p)| \leq \varepsilon \|\Lambda^{-1}\|^2 |p|$. Par conséquent, pour tout $|p| \leq \eta$,

$$\sup_{x \in \partial\psi^*(p)} |x - \Lambda^{-1}(p)| \leq \frac{\varepsilon \|\Lambda^{-1}\|^2}{1 - \varepsilon \|\Lambda^{-1}\|} |p|,$$

et ψ^* est bien deux fois différentiable au sens d'Aleksandrov en 0, de hessienne Λ^{-1} .

Réciproque. Nous montrons maintenant que si ψ^* est deux fois différentiable, alors Λ est inversible. Supposons l'inverse. Soit x un point du noyau de Λ , $x_n = \frac{1}{n}x$, et p_n un élément de $\partial\psi(x_n)$. Pour n suffisamment grand $|x_n| < \delta$, par conséquent (9.2) donne

$$|p_n| = |p_n - \Lambda(x_n)| \leq \varepsilon|x_n|,$$

ce qui entraîne en particulier que p_n tend vers 0.

Par hypothèse ψ^* est différentiable en 0 ; comme $\partial\psi^*(0) \ni 0$, on a $\nabla\psi^*(0) = 0$. Comme ψ^* est deux fois différentiable au sens d'Aleksandrov, il existe une matrice symétrique Λ' et $\delta' > 0$ tels que si $|p| < \delta'$,

$$\sup_{x \in \partial\psi^*(p)} |x - \Lambda'(p)| < \varepsilon|p|.$$

Puisque $x_n \in \partial\psi^*(p_n)$ et $p_n \rightarrow 0$, on a pour n suffisamment grand,

$$|x_n| \leq |x_n - \Lambda'(p_n)| + |\Lambda'(p_n)| \leq (\varepsilon + \|\Lambda'\|) |p_n| \leq (\varepsilon + \|\Lambda'\|) \varepsilon |x_n|,$$

Comme $x_n \neq 0$, on obtient $1 \leq (\varepsilon + \|\Lambda'\|) \varepsilon$, ce qui est absurde pour ε proche de 0. □

Preuve du corollaire. Le premier point est une réécriture de la proposition : si $x \in \Omega_{\text{Al}} \setminus X$, alors $\nabla\psi(x)$ est bien dans $\text{dom}(\psi^*)$, mais ψ^* n'est pas deux fois différentiable en $\nabla\psi(x)$.

Notons d'abord que si $x \in X$ et $y \in \Omega_{\text{diff}}$, alors

$$\nabla\psi(x) = \nabla\psi(y) \implies x = y.$$

En effet, d'après la proposition, ψ^* est deux fois différentiable en $\nabla\psi(x)$ et *a fortiori* différentiable, donc $\partial\psi^*(\nabla\psi(x))$ est le singleton $\{x\}$. Or $\nabla\psi(y) = \nabla\psi(x)$ implique que $y \in \partial\psi^*(\nabla\psi(x))$, donc y est bien égal à x .

La fonction $\nabla\psi$ est donc injective sur X . Toujours d'après la proposition, $\nabla\psi(X) \subset X^*$. Réciproquement, si $p \in X^*$, $\nabla\psi^*(p)$ existe et ψ^{**} y est deux fois différentiable, de hessienne inversible. Appelons x ce point : x est nécessairement dans $\text{int}(\text{dom}(\psi^{**})) = \text{int}(\text{dom}(\psi))$. Sur l'ouvert Ω , $\psi = \psi^{**}$ donc ψ est elle-même deux fois différentiable en x de hessienne inversible, autrement dit $x \in X$. De plus si $q = \nabla\psi(x)$, $q \in X^*$ par le point précédent, donc $x = \nabla\psi^*(q) = \nabla\psi^*(p)$ donc $p = q$. Autrement dit $p = \nabla\psi\nabla\psi^*(p)$, ce qui achève la preuve de surjectivité et la formule d'inversion. \square

9.5 Outils de théorie de la mesure

Rappelons quelques résultats standard de théorie de la mesure, issu du Théorème de Radon-Nikodym-Lebesgue. Pour plus de détails, nous renvoyons au livre de W. Rudin [Rud87].

Soit ν une mesure borélienne positive définie sur un ouvert Ω de \mathbb{R}^n . Elle se décompose en une mesure ν_{ac} absolument continue par rapport à la mesure de Lebesgue, de densité g , et une mesure singulière ν_{sing} dont le support est de mesure nulle pour la mesure de Lebesgue λ sur \mathbb{R}^n ,

$$\nu = \nu_{ac} + \nu_{sing}.$$

On dit que ν est σ -finie si \mathbb{R}^n s'écrit comme une union dénombrable d'ensemble de ν -mesure finie ; c'est en particulier le cas si ν est localement finie (c'est-à-dire finie sur les compacts).

Pour une mesure ν absolument continue par rapport à la mesure de Lebesgue, la densité de Radon-Nikodym notée souvent $d\nu/d\lambda$ peut se construire à partir du rapport des mesures de petits ensembles autour de x . Avant d'énoncer un résultat précis, donnons deux définitions. Dans toute la suite, $B_r(x)$ désigne la boule euclidienne centrée en x et de rayon r ; son volume usuel $\lambda(B_r(x))$ sera aussi noté $|B_r(x)| = |B_r|$.

Définition 9.14 (Convergence propre). *Une famille de boréliens $(M_r(x))_{r>0}$ converge proprement vers $\{x\}$ si $M_r(x) \subset B_{R(r)}(x)$ avec $\lim_{r \rightarrow 0} R(r) = 0$, et si il existe $C, \delta > 0$ telle que tout $0 < r < \delta$, $|M_r(x)|/|B_{R(r)}(x)| \geq C$.*

Définition 9.15 (Différentiation de Lebesgue, point de Lebesgue). *Si le rapport $\nu(B_r(x))/|B_r(x)|$ converge quand r tend vers 0, on dit que ν est différentiable par rapport à λ en x , ou que x est un point de Lebesgue pour ν , et on note $D\nu(x)$ la limite du rapport.*

Théorème 9.16. *Soit ν une mesure localement finie, de décomposition de Radon-Nikodym $\nu = f\lambda + \nu_{sing}$. Alors pour λ presque tout x , ν est différentiable par rapport à λ en x , c'est à dire*

$$f(x) = D\nu(x) = \lim_{r \rightarrow 0} \frac{\nu(B_r(x))}{|B_r(x)|} \quad \lambda\text{-presque partout.} \quad (9.3)$$

La dernière égalité reste vraie si on remplace $B_r(x)$ par une famille $M_r(x)$ qui converge proprement vers $\{x\}$. De plus on a pour ν_{sing} -presque tout x

$$\lim_{r \rightarrow 0} \frac{\nu(B_r(x))}{|B_r(x)|} = +\infty.$$

Dans le théorème usuel de changement de variable dans une intégrale, le terme jacobien s'interprète comme rapport entre le volume d'une boule et de son image par une application linéaire. La même idée est traduite dans notre cadre par la proposition suivante.

Proposition 9.17. Soit $\psi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe deux fois différentiable au sens d'Aleksandrov au point x_0 . Notons $\Lambda = \text{Hess}_{x_0}\psi$, alors

$$\lim_{r \rightarrow 0} \frac{|\partial\psi(B_r(x_0))|}{|B_r(x_0)|} = \det \Lambda.$$

Par ailleurs si Λ est inversible, $\partial\psi(B_r(x_0))$ converge proprement vers $\{\nabla\psi(x_0)\}$.

Démonstration. Comme dans la démonstration de la proposition 9.9, on se ramène par translation au cas où ψ est deux fois différentiable au sens d'Aleksandrov en $x_0 = 0$ avec $\nabla\psi(0) = 0$. Notons $B_r = B_r(0)$ et ΛB_r son image par Λ (c'est un ellipsoïde).

Supposons dans un premier temps que Λ est inversible. Etant donné $\varepsilon > 0$, d'après (9.2), il existe $\delta > 0$ tel que pour tout $0 < r < \delta$, $x \in B_r$ et $p \in \partial\psi(x)$,

$$|x - \Lambda^{-1}(p)| \leq \|\Lambda^{-1}\| |\Lambda(x) - p| < \varepsilon \|\Lambda^{-1}\| |x|$$

et donc

$$|x - \Lambda^{-1}(p)| \leq (1 + \varepsilon \|\Lambda^{-1}\|) |x|$$

Par conséquent, pour tout $p \in \partial\psi(x)$, $p = \Lambda(\Lambda^{-1}(p)) \in \Lambda((1 + \varepsilon \|\Lambda^{-1}\|)B_r) = (1 + \varepsilon \|\Lambda^{-1}\|)\Lambda B_r$. On a donc pour r suffisamment petit

$$\partial\psi(B_r) \subset (1 + \varepsilon \|\Lambda^{-1}\|)\Lambda B_r. \quad (9.4)$$

Par ailleurs, d'après la Proposition 9.12, puisque Λ est inversible, ψ^* est deux fois différentiable au sens d'Aleksandrov. Plus précisément il existe $\delta' > 0$ tel que pour tout $|p| \leq \delta'$ et pour tout $x \in \partial\psi^*(p)$,

$$|x - \Lambda^{-1}(p)| \leq \varepsilon |p|, \quad |x| \leq \varepsilon |p| + |\Lambda^{-1}(p)| \leq (\varepsilon \|\Lambda\| + 1) |\Lambda^{-1}(p)|.$$

Choisissons $0 < r < \delta' / \|\Lambda\|$. Si $p \in \Lambda B_r$, alors $p = \Lambda z$, $z \in B_r$ et $|p| \leq \delta'$ donc si $x \in \partial\psi^*(p)$, alors $|x| \leq (\varepsilon \|\Lambda\| + 1) |z| \leq (\varepsilon \|\Lambda\| + 1) r$. Par conséquent, on a $\partial\psi^*(\Lambda B_r) \subset (\varepsilon \|\Lambda\| + 1) B_r = B_{(\varepsilon \|\Lambda\| + 1)r}$ et par suite d'après le point 5 du Lemme 9.5, $\Lambda B_r \subset \partial\psi(B_{(\varepsilon \|\Lambda\| + 1)r})$ (car pour r suffisamment petit $B_{(\varepsilon \|\Lambda\| + 1)r} \subset \Omega$). Par linéarité, on obtient pour r suffisamment petit,

$$\frac{1}{\varepsilon \|\Lambda\| + 1} \Lambda B_r \subset \partial\psi(B_r). \quad (9.5)$$

Les inclusions (9.4) et (9.5) induisent les inégalités suivantes

$$\frac{1}{(\varepsilon \|\Lambda\| + 1)^n} \frac{|\Lambda B_r|}{|B_r|} \leq \frac{|\partial\psi(B_r)|}{|B_r|} \leq (1 + \varepsilon \|\Lambda^{-1}\|)^n \frac{|\Lambda B_r|}{|B_r|}.$$

Par ailleurs, on a

$$|\Lambda B_r| = \int \mathbf{1}_{\Lambda B_r}(p) dp = \int \mathbf{1}_{B_r}(\Lambda^{-1}(p)) dp = \int \mathbf{1}_{B_r}(x) \det(\Lambda) dx = \det(\Lambda) |B_r|.$$

Ainsi puisque $\lim_{\varepsilon \rightarrow 0} (\varepsilon \|\Lambda\| + 1)^n = \lim_{\varepsilon \rightarrow 0} (1 + \varepsilon \|\Lambda^{-1}\|)^n = 1$, l'inégalité précédente établit que

$$\lim_{r \rightarrow 0} \frac{|\partial\psi(B_r)|}{|B_r|} = \det(\Lambda).$$

L'inclusion (9.4) implique que $\partial\psi(B_r) \subset B_{R(r)}$ avec $R(r) = (1 + \varepsilon \|\Lambda^{-1}\|) \|\Lambda\| r$, et par ailleurs

$$\frac{|\partial\psi(B_r)|}{|B_{R(r)}|} = \frac{r^n}{R(r)^n} \frac{|\partial\psi(B_r)|}{|B_r|} \geq \frac{\det(\Lambda)}{(\varepsilon \|\Lambda\| + 1)^n (1 + \varepsilon \|\Lambda^{-1}\|)^n \|\Lambda\|^n},$$

ce qui montre que $\partial\psi(B_r)$ converge proprement vers $\{0\}$.

On suppose maintenant que Λ n'est pas inversible, il reste à montrer que

$$\lim_{r \rightarrow 0} \frac{|\partial\psi(B_r)|}{|B_r|} = 0 = \det(\Lambda).$$

On sait qu'il existe $\delta > 0$ tel que pour $0 < r < \delta$, si $|x| < r$ alors pour tout $y \in \partial\psi(x)$,

$$|y - \Lambda(x)| < \varepsilon|x|, \quad y \in \Lambda B_r + \varepsilon B_r.$$

Par conséquent on a $\partial\psi(B_r) \subset \Lambda B_r + \varepsilon B_r$. Comme Λ est symétrique réelle on peut la diagonaliser dans une base orthonormée (e_1, \dots, e_n) de vecteurs propres, associés aux valeurs propres $(\lambda_1, \dots, \lambda_n)$.

La boule euclidienne B_r est incluse dans la boule de rayon r pour la norme infinie dans le système de coordonnées relatif à la base (e_1, \dots, e_n) : notons B'_r cette boule. Le choix de la base assure que

$$\Lambda B'_r = \{x, |(x, e_i)| \leq \lambda_i r\},$$

donc

$$\Lambda B'_r + \varepsilon B'_r \subset \{x, |(x, e_i)| \leq (\lambda_i + \varepsilon)r\}.$$

L'ensemble de droite est un pavé, donc

$$|\partial\psi(B_r)| \leq |\Lambda B'_r + \varepsilon B'_r| \leq r^n \prod_{i=1}^n (\lambda_i + \varepsilon).$$

En divisant par le volume de la boule euclidienne $|B_r| = C_n r^n$, et en prenant la limite en r on obtient

$$\limsup_{r \rightarrow 0} \frac{|\partial\psi(B_r)|}{|B_r|} \leq \frac{1}{C_n} \prod_{i=1}^n (\lambda_i + \varepsilon).$$

Quand ε tend vers 0, le terme de droite tend vers $\frac{1}{C_n} \det(\Lambda)$, qui est nul puisque Λ n'est pas inversible, ce qui conclut la preuve. \square

9.6 Équation de Monge-Ampère, formules de changement de variables

Soit μ une mesure de densité f par rapport à la mesure de Lebesgue et ϕ une fonction convexe. On s'intéresse à la mesure image de μ par l'application $\nabla\phi$ définie presque partout sur Ω , $\nu = \nabla\phi\#\mu$: pour que ceci ait un sens on suppose donc que μ est portée par Ω . Comme $\nabla\phi$ n'est définie que sur Ω_{diff} , on a par définition, pour tout borélien M ,

$$\begin{aligned} \nabla\phi^{-1}(M) &= \{x \in \Omega_{\text{diff}}, \nabla\phi(x) \in M\} \\ &= \{x \in \Omega_{\text{diff}}, \exists y \in M, \nabla\phi(x) = y\} \\ &= \{x \in \Omega_{\text{diff}}, \exists y \in M, x \in \partial\phi^*(y)\} \\ &= \partial\phi^*(M), \end{aligned}$$

où l'avant dernière équation provient du point 4 du lemme 9.5. Par conséquent, pour tout borélien,

$$\nu(M) = \mu(\partial\phi^*(M)). \quad (9.6)$$

Nous allons établir l'équation de Monge-Ampère en plusieurs étapes. Le premier résultat montre que l'équation a lieu pour certains points de l'ensemble de départ.

Proposition 9.18. Soit μ une mesure absolument continue par rapport à la mesure de Lebesgue de densité f . Soit $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe et $\nu = \nabla\phi\#\mu$. On suppose que ν est localement finie.

Si $x \in X$ est un point de Lebesgue pour f , alors ν est Lebesgue-différentiable en $\nabla\phi(x)$ et on a

$$D\nu(\nabla\phi(x)) = \lim_{r \rightarrow 0} \frac{\nu(B_r(\nabla\phi(x)))}{|B_r|} = \frac{f(x)}{\det(\text{Hess}_x \phi)}.$$

Démonstration. Soit x un point de X qui est aussi un point de Lebesgue pour f ; on pose $y = \nabla\phi(x)$. Comme ν est localement finie, on a d'après (9.6), pour tout $r > 0$,

$$\frac{\nu(B_r(y))}{|B_r|} = \frac{\mu(\partial\phi^*(B_r(y)))}{|\partial\phi^*(B_r(y))|} \cdot \frac{|\partial\phi^*(B_r(y))|}{|B_r|}.$$

Nous allons passer à la limite dans cette égalité lorsque $r \rightarrow 0$. Pour le premier quotient du membre de droite, on applique la proposition 9.17 à la fonction ϕ^* qui, d'après la condition d'inversibilité de la Proposition 9.12, est deux fois différentiable au sens d'Aleksandrov au point $y = \nabla\phi(x)$ puisque $x \in X$. On obtient

$$\lim_{r \rightarrow 0} \frac{\mu(\partial\phi^*(B_r(\nabla\phi(x))))}{|\partial\phi^*(B_r(\nabla\phi(x)))|} = \det(\text{Hess}_{\nabla\phi(x)} \phi^*) = \det(\text{Hess}_x^{-1} \phi) = \frac{1}{\det(\text{Hess}_x \phi)}.$$

Pour le second quotient, $\partial\phi^*(B_r(\nabla\phi(x)))$ converge proprement vers $\{\nabla\phi^*(\nabla\phi(x))\} = \{x\}$, et comme x est un point de Lebesgue pour f ,

$$\lim_{r \rightarrow 0} \frac{|\partial\phi^*(B_r(\nabla\phi(x)))|}{|B_r|} = f(x).$$

On obtient donc

$$\lim_{r \rightarrow 0} \frac{\nu(B_r(\nabla\phi(x)))}{|B_r|} = \frac{f(x)}{\det(\text{Hess}_x \phi)},$$

ce qui établit la Lebesgue-différentiabilité de ν par rapport à λ en $\nabla\phi(x)$. \square

La deuxième étape établit que l'application $x \mapsto \det(\text{Hess}_x \phi)$ est localement intégrable sur Ω . La preuve de ce lemme s'appuie sur la Proposition 9.18.

Lemme 9.19. Soit $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe alors la mesure $\omega = \nabla\phi^*\#\lambda$ (où λ désigne la mesure de Lebesgue sur \mathbb{R}^n) est une mesure localement finie sur $\Omega = \text{int}(\text{dom}(\phi))$ dont la partie absolument continue ω_{ac} admet pour densité : pour presque tout $x \in \Omega$

$$D\omega(x) = \det(\text{Hess}_x \phi).$$

Par suite l'application $x \mapsto \det(\text{Hess}_x \phi)$ est localement intégrable sur Ω . Par ailleurs, tous les points de X sont des points de Lebesgue pour ω .

Démonstration. D'après (9.6), pour tout borélien M dans \mathbb{R}^n ,

$$\omega(M) = \lambda(\partial\phi^{**}(M)).$$

Si de plus $M \subset \Omega = \text{int}(\text{dom} \phi)$ alors puisque $\phi(x) = \phi^{**}(x)$ pour $x \in \Omega$ (cf. 4. Lemme 9.5),

$$\omega(M) = \lambda(\partial\phi(M)).$$

Lorsque M est compact, $\partial\phi(M)$ est compact (cf. 2. Lemme 9.5) donc $\lambda(\partial\phi(M)) < +\infty$, ce qui fait de la mesure ω une mesure localement finie sur Ω .

Soit $y \in X^*$ tel que $\nabla\phi^*(y) \in \Omega$. Puisque tout point de \mathbb{R}^n est un point de Lebesgue pour la mesure de Lebesgue λ , d'après la proposition 9.18,

$$1 = D\omega(\nabla\phi^*(y)) \det(\text{Hess}_y\phi^*).$$

Choisissons $x \in X$ et $y = \nabla\phi(x)$. D'après le Corollaire 9.13, y est un élément de X^* et d'après la Proposition 9.12, $\det(\text{Hess}_y\phi^*) = [\det(\text{Hess}_x\phi)]^{-1}$. Par conséquent, on a

$$D\omega(x) = \det(\text{Hess}_x\phi).$$

En particulier, x est un point de Lebesgue pour ω . Il reste à montrer que $\omega_{ac}(\Omega \setminus X) = 0$. Puisque d'après le Théorème d'Aleksandrov

$$\omega_{ac}(\Omega \setminus X) = \omega_{ac}(\Omega_{Al} \setminus X) \leq \omega(\Omega_{Al} \setminus X).$$

D'après ce qui précède,

$$\omega(\Omega_{Al} \setminus X) = \lambda(\partial\phi(\Omega_{Al} \setminus X)) = \lambda(\nabla\phi(\Omega_{Al} \setminus X)).$$

D'après le Corollaire 9.13 et le Théorème d'Aleksandrov, on a

$$\lambda(\nabla\phi(\Omega_{Al} \setminus X)) \leq \lambda(\text{dom}(\phi^*) \setminus \Omega_{Al}^*) = 0,$$

et par conséquent $\omega_{ac}(\Omega \setminus X) = 0$. Ceci termine la démonstration. \square

Nous allons pouvoir maintenant établir le résultat principal de ce chapitre, l'équation de Monge-Ampère annoncée dans le Théorème 9.2. Cette équation n'est rien d'autre qu'une formule de changement de variables. Avant cela, nous allons établir une autre formule de changement de variables, qui sera très utile dans les chapitres suivants. Par ailleurs, le résultat qui suit complète en un certain sens le Lemme 9.19 que nous venons de montrer.

Lemme 9.20. *Soit $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe et μ_ϕ la mesure de densité (localement intégrable)*

$$f_\phi(x) = \det(\text{Hess}_x\phi) \mathbf{1}_\Omega(x), \quad x \in \mathbb{R}^n,$$

où $\Omega = \text{int}(\text{dom } \phi)$. Alors on a

$$\nabla\phi \# \mu_\phi = \lambda|_{\nabla\phi(X)},$$

où $\lambda|_{\nabla\phi(X)}$ est la mesure de Lebesgue restreinte à $\nabla\phi(X)$.

Ce lemme donne le changement de variable suivant : pour toute fonction $G : \mathbb{R}^n \rightarrow \mathbb{R}$ mesurable bornée,

$$\int G d\lambda|_{\nabla\phi(X)} = \int_{\nabla\phi(X)} G(y) dy = \int_{\Omega} G(\nabla\phi(x)) \det(\text{Hess}_x\phi) dx. \quad (9.7)$$

Démonstration. Appliquons la Proposition 9.18 à la mesure $\mu = \mu_\phi$ de densité

$$f_\phi(x) = \det(\text{Hess}_x\phi) \mathbf{1}_\Omega(x), \quad x \in \mathbb{R}^n.$$

D'après le Lemme 9.19, cette densité est bien localement intégrable sur Ω et tout point de X est un point de Lebesgue pour μ_ϕ .

Notons $\nu = \nabla\phi\#\mu_\phi$. D'après (9.6) et par définition de μ_ϕ

$$\nu(\nabla\phi(X)) = \nu(\partial\phi(X)) = \mu_\phi(\partial\phi^*(\partial\phi(X))) = \mu_\phi(\nabla\phi^*(\nabla\phi(X))) = \mu_\phi(X) = \mu_\phi(\mathbb{R}^n).$$

La mesure ν est donc portée par $\nabla\phi(X)$.

Soit $K \subset \Omega^*$ un compact. Alors

$$\nu(K) = \nu(K \cap \nabla\phi(X)) = \mu_\phi(\partial\phi^*(K \cap \nabla\phi(X))) \leq \mu_\phi\left(\partial\phi^*\left(\overline{K \cap \nabla\phi(X)}\right)\right).$$

Puisque $\overline{K \cap \nabla\phi(X)} \subset K \subset \Omega^*$, l'ensemble $\overline{K \cap \nabla\phi(X)}$ est un compact de Ω^* . Par suite d'après le point 2. du Lemme 9.5, l'ensemble $\partial\phi^*(K \cap \nabla\phi(X))$ est compact et

$$\mu_\phi\left(\partial\phi^*\left(\overline{K \cap \nabla\phi(X)}\right)\right) < \infty.$$

Par conséquent, ν est une mesure localement finie sur Ω^* .

D'après la Proposition 9.18, pour tout $x \in X$, $\nabla\phi(x)$ est un point de Lebesgue pour ν et

$$\det(\text{Hess}_x\phi)\mathbf{1}_\Omega(x) = f_\phi(x) = D\nu(\nabla\phi(x)) \det(\text{Hess}_x\phi).$$

Par conséquent $1 = D\nu(\nabla\phi(x))$ pour tout $x \in X$. La mesure ν est donc la mesure de Lebesgue sur $\nabla\phi(X)$. \square

Nous allons maintenant établir l'équation de Monge-Ampère énoncée dans la version plus complète suivante du Théorème 9.2.

Théorème 9.21. *Soient μ et ν deux probabilités sur \mathbb{R}^n absolument continues par rapport à la mesure de Lebesgue de densité respectives f et g . Soit $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe telle que $\nabla\phi\#\mu = \nu$ ($\nabla\phi$ est le transport de Brenier de μ sur ν). Alors les propriétés suivantes sont vérifiées.*

- La mesure μ est portée par l'ensemble X des points x tels que $\text{Hess}_x\phi$ existe au sens d'Aleksandrov et est inversible.
- La mesure ν est portée par $\partial\phi(X) = \nabla\phi(X)$.
- L'application $\nabla\phi^*$ est le transport de Brenier de ν sur μ , $\nabla\phi^*\#\nu = \mu$.
- L'équation de Monge-Ampère suivante est vérifiée, pour presque tout x dans $\text{dom}(\phi)$ ou pour μ -presque tout x dans \mathbb{R}^n ,

$$f(x) = g(\nabla\phi(x)) \det(\text{Hess}_x\phi),$$

et l'application $x \rightarrow \det(\text{Hess}_x\phi)$ est localement intégrable sur Ω .

Sous les hypothèses du théorème, on a donc : pour toute fonction H mesurable bornée sur \mathbb{R}^n ,

$$\begin{aligned} \int H(y)g(y)dy &= \int H d\nu = \int_{\nabla\phi(X)} H d\nu = \int_X H(\nabla\phi) d\mu \\ &= \int_X H(\nabla\phi(x))f(x)dx = \int H(\nabla\phi(x))g(\nabla\phi(x)) \det(\text{Hess}_x\phi)dx. \end{aligned}$$

Démonstration du Théorème 9.21. Supposons maintenant que μ et ν sont deux mesures de densités respectives f et g par rapport à la mesure de Lebesgue et ϕ

Puisque $\mu \ll \lambda$, d'après le Théorème d'Aleksandrov,

$$\mu(\Omega \setminus \Omega_{Al}) = \lambda(\Omega \setminus \Omega_{Al}) = 0.$$

Par ailleurs d'après le Corollaire 9.13,

$$\partial\phi(\Omega_{Al} \setminus X) \subset \text{dom}(\phi^*) \setminus \Omega_{Al}^*.$$

et par conséquent, d'après le point 5. du Lemme 9.5

$$\Omega_{Al} \setminus X \subset \partial\phi^*(\text{dom}(\phi^*) \setminus \Omega_{Al}^*).$$

Par suite, d'après (9.6), on obtient

$$\mu(\Omega_{Al} \setminus X) \leq \mu(\partial\phi^*(\Omega^* \setminus \Omega_{Al}^*)) = \nu(\Omega^* \setminus \Omega_{Al}^*).$$

Puisque $\mu \ll \lambda$, d'après le Théorème d'Aleksandrov,

$$\nu(\Omega^* \setminus \Omega_{Al}^*) = \lambda(\Omega^* \setminus \Omega_{Al}^*) = 0.$$

On obtient ainsi $\mu(\Omega_{Al} \setminus X) = 0$ et par conséquent

$$\mu(\mathbb{R}^n \setminus X) = 0,$$

La mesure μ est donc portée par l'ensemble X .

D'après la Proposition 9.18, l'équation de Monge-Ampère

$$f(x) = g(\nabla\phi(x)) \det(\text{Hess}_x\phi),$$

est satisfaite pour presque tout x dans $\text{dom}(\phi)$ ou pour μ presque tout x dans \mathbb{R}^n .

Par ailleurs, puisque

$$\nu(\nabla\phi(X)) = \mu(\nabla\phi^*(\nabla\phi(X))) = \mu(X) = \mu(\mathbb{R}^n),$$

la mesure ν est portée par $\nabla\phi(X) = \partial\phi(X) = (\nabla\phi^*)^{-1}(X) \subset \Omega^*$. Par suite pour tout sous-ensemble mesurable $M \subset \mathbb{R}^n$ mesurable,

$$\begin{aligned} \nabla\phi^*\#\nu(M) &= \nu((\nabla\phi^*)^{-1}(M)) \\ &= \nu((\nabla\phi^*)^{-1}(M) \cap (\nabla\phi^*)^{-1}(X)) \\ &= \nu((\nabla\phi^*)^{-1}(M \cap X)) \\ &= \mu(\partial\phi^*((\nabla\phi^*)^{-1}(M \cap X))) \\ &= \mu(M \cap X) \\ &= \mu(M), \end{aligned}$$

où la 4-ème égalité est une conséquence du fait que $\nabla\phi\#\mu = \nu$ et de (9.6). L'application $\nabla\phi^*$ transporte donc ν sur μ , $\nabla\phi^*\#\nu = \mu$. \square

9.7 Remarques bibliographiques

Les preuves présentées dans ce chapitre de l'équation de Monge-Ampère dans le Théorème 9.21 ainsi que celles des résultats annexes ou préliminaires des Propositions 9.12 et 9.17 et des Lemmes 9.19 et 9.20 sont celles de R. J. McCann [McC97].

Chapitre 10

Inégalités de Brunn-Minkowski et de Prékopa-Leindler

L'objectif de ce paragraphe est de démontrer une inégalité géométrique classique sur les ensembles, l'inégalité de Brunn-Minkowski, et une inégalité fonctionnelle l'inégalité de Prékopa-Leindler. Ces deux inégalités seront démontrées en utilisant le transport de Brenier. L'inégalité de Prékopa-Leindler peut être interprétée comme une version fonctionnelle plus générale de l'inégalité de Brunn-Minkowski.

10.1 Inégalité de Brunn-Minkowski

Etant donné deux sous-ensembles A et B de \mathbb{R}^n , on note $A + B$ la somme de ces ensembles définie par

$$A + B = \{a + b, a \in A, b \in B\}.$$

L'inégalité de Brunn-Minkowski est une inégalité optimale qui compare le volume euclidien de la somme $A + B$ à celle de A et B lorsque ces ensembles sont mesurables.

Si B est l'ensemble vide alors l'ensemble $A + B$ est vide. Cependant dès que les deux ensembles A et B ne sont pas vide il est clair que $|A + B| \geq |A + \{b\}| = |A|$ pour $b \in B$, ainsi $|A + B| \geq \max(|A|, |B|)$. Imaginons que B contienne deux points éloignés l'un de l'autre, on voit simplement sur un dessin, que $A + B$ peut être un ensemble beaucoup plus gros que A , son volume peut doubler.

L'inégalité de Brunn-Minkowski donne une minoration optimale du volume de $A + B$ en fonction de ceux de A et B .

Théorème 10.1. (*Inégalité de Brunn-Minkowski*) Soient A et B des sous-ensembles mesurables non vides de \mathbb{R}^n , alors

$$|A + B|^{1/n} \geq |A|^{1/n} + |B|^{1/n}.$$

Cette inégalité admet une formulation multiplicative équivalente. Pour $s \in \mathbb{R}$ et $A \subset \mathbb{R}^n$, notons $sA = \{sa, a \in A\}$. Soit $t \in [0, 1]$, par homogénéité, en remplaçant A par $(1 - t)A$ et B par tB , on obtient

$$|(1 - t)A + tB|^{1/n} \geq (1 - t)|A|^{1/n} + t|B|^{1/n}.$$

En utilisant l'inégalité arithmético-géométrique $(1 - t)u + tv \geq u^{1-t}v^t$ (conséquence de la concavité de la fonction logarithme), on obtient la formulation suivante de l'inégalité de Brunn-Minkowski, pour tout $t \in [0, 1]$

$$|(1 - t)A + tB| \geq |A|^{1-t}|B|^t. \tag{10.1}$$

Pour retrouver la forme additive, il suffit de choisir $A = \frac{A'}{|A'|^{1/n}}$, $B = \frac{B'}{|B'|^{1/n}}$ et $t = \frac{|A'|^{1/n}}{|A'|^{1/n} + |B'|^{1/n}}$.

Démonstration du Théorème 10.1. On suppose que A et B sont des ensembles compacts. D'après ce qui précède, il suffit de démontrer l'inégalité (10.1) lorsque $|A| = |B| = 1$, c'est à dire

$$|(1-t)A + tB| \geq 1, \quad t \in [0, 1].$$

L'inégalité de Brunn-Minkowski se retrouve en choisissant $A = \frac{A'}{|A'|^{1/n}}$, $B = \frac{B'}{|B'|^{1/n}}$ et $t = \frac{|A'|^{1/n}}{|A'|^{1/n} + |B'|^{1/n}}$.

Soient λ_A et λ_B les mesures de Lebesgue restreintes aux ensembles A et B . Ces mesures sont des probabilités. Soit $\nabla\phi$ l'application de transport de Brenier telle que $\nabla\phi\#\lambda_A = \lambda_B$. Considérons l'application convexe $\phi_t : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ définie par

$$\phi_t(x) = (1-t)\frac{|x|^2}{2} + t\phi(x), \quad x \in \mathbb{R}^n, \quad t \in [0, 1].$$

Cette application a même domaine que ϕ , on note $\Omega = \text{int}(\text{dom } \phi) = \text{int}(\text{dom } \phi_t)$. Notons respectivement X et X_t l'ensemble des points x où $\text{Hess}_x\phi$ et respectivement $\text{Hess}_x\phi_t$ est inversible. Il est clair que $X \subset X_t$ car pour presque tout $x \in \Omega$,

$$\text{Hess}_x\phi_t = (1-t)I + t\text{Hess}_x\phi,$$

et $\text{Hess}_x\phi$ est une matrice positive. Pour tout $x \in A \cap X$, puisque $\nabla\phi(x) \in B$ et

$$\nabla\phi_t(x) = (1-t)x + t\nabla\phi(x),$$

on obtient

$$\nabla\phi_t(A \cap X) \subset (1-t)A + tB.$$

Par conséquent d'après le Lemme 9.20 et plus particulièrement (9.7),

$$\begin{aligned} |(1-t)A + tB| &\geq \int \mathbf{1}_{\nabla\phi_t(A \cap X)}(y) dy \\ &= \int_{\nabla\phi_t(X_t)} \mathbf{1}_{\nabla\phi_t(A \cap X)}(y) dy \\ &= \int_{\Omega} \mathbf{1}_{\nabla\phi_t(A \cap X)}(\nabla\phi_t(x)) \det(\text{Hess}_x\phi_t) dx \\ &= \int_X \mathbf{1}_A(x) \det((1-t)I + t\text{Hess}_x\phi) dx. \end{aligned}$$

La suite de la démonstration s'appuie sur la propriété de log-concavité du déterminant sur l'ensemble des matrices symétriques positives. Dans le lemme qui suit nous donnons deux inégalités qui résultent de cette propriété de log-concavité. La seconde sera utile dans le chapitre suivant.

Lemme 10.2. (*Log-concavité du déterminant*) Soit $\Lambda, \Lambda_0, \Lambda_1$ des matrices symétriques positives sur \mathbb{R}^n et $t \in [0, 1]$.

1. $\det((1-t)I + t\Lambda) \geq (\det \Lambda)^t$.
2. $\det((1-t)\Lambda_0 + t\Lambda_1) \geq (\det \Lambda_0)^{1-t} (\det \Lambda_1)^t$.
3. $(\det \Lambda)^{1/n} \leq \frac{\text{Tr } \Lambda}{n}$.

Ce lemme se démontre en diagonalisant la matrice symétrique Λ dans une base orthogonale. Le déterminant s'écrit alors simplement en fonction des valeurs propres. Le résultat découle ensuite de la propriété de concavité de la fonction logarithme.

En appliquant le premier point de ce lemme au résultat précédent on obtient

$$|(1-t)A + tB| \geq \int_X \mathbf{1}_A(x) [\det(\text{Hess}_x \phi)]^t dx.$$

D'après le Théorème 9.21, puisque $\nabla \phi \# \lambda_A = \lambda_B$, la mesure λ_A est portée par X , la mesure λ_B est portée par $\nabla \phi(X)$ et l'équation de Monge-Ampère indique que pour presque tout $x \in A \cap X$,

$$1 = \det(\text{Hess}_x \phi).$$

On obtient ainsi le résultat attendu,

$$|(1-t)A + tB| \geq \int_X \mathbf{1}_A(x) dx = 1.$$

□

10.2 Inégalité de Prékopa-Leindler

L'inégalité de Prékopa-Leindler est une forme inverse de l'inégalité de Hölder écrite sous la forme suivante : si f et g sont des fonctions intégrables positives sur \mathbb{R}^n alors pour tout $t \in [0, 1]$,

$$\int f^{1-t}(x)g^t(x) \leq \left(\int f(x)dx \right)^{1-t} \left(\int g(x)dx \right)^t.$$

L'inégalité de Prékopa-Leindler majore le membre de droite par l'intégrale d'une fonction qui peut être choisie de façon "optimale".

Théorème 10.3. (*Inégalité de Prékopa-Leindler*) Soient f, g et h des fonctions positives intégrables sur \mathbb{R}^n telles que pour tout $x, y \in \mathbb{R}^n$,

$$f(x)^{1-t}g(y)^t \leq h((1-t)x + ty),$$

alors,

$$\left(\int f(x)dx \right)^{1-t} \left(\int g(y)dy \right)^t \leq \int h(z)dz.$$

Remarques :

— Etant données f et g , la meilleure fonction h qu'on puisse choisir est

$$h(z) = \sup_{x,y: z=(1-t)x+ty} f(x)^{1-t}g(y)^t.$$

— Etant donné deux sous-ensembles mesurables A et B , appliquons ce résultat aux fonctions $f = \mathbf{1}_A$, $g = \mathbf{1}_B$ et $h = \mathbf{1}_{(1-t)A+tB}$. Par définition de l'ensemble $(1-t)A + tB$, l'hypothèse

$$\mathbf{1}_A(x)^{1-t}\mathbf{1}_B(y)^t \leq \mathbf{1}_{(1-t)A+tB}((1-t)x + ty),$$

est bien satisfaite. L'inégalité de Prékopa-Leindler donne alors

$$|A|^{1-t}|B|^t \leq |(1-t)A + tB|.$$

Cette inégalité est la forme multiplicative de l'inégalité de Brunn-Minkowski. Ainsi l'inégalité de Prékopa-Leindler est une inégalité fonctionnelle plus générale que l'inégalité de Brunn-Minkowski.

En fait, l'inégalité de Prékopa-Leindler en dimension n se déduit de l'inégalité de Brunn-Minkowski en dimension $n+1$ en considérant les ensembles $A = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}, 0 \leq t \leq f(x)\}$, $B = \{(y, u) \in \mathbb{R}^n \times \mathbb{R}, 0 \leq u \leq g(y)\}$ et $C = \{(z, v) \in \mathbb{R}^n \times \mathbb{R}, 0 \leq v \leq h(z)\}$ (exercice).

Démonstration du Théorème 10.3. Par convergence monotone, on peut se ramener à des fonctions à support borné (en considérant par exemple les fonctions $f_n = f\mathbf{1}_{B_n}$, $g_n = g\mathbf{1}_{B_n}$ et $h_n = h\mathbf{1}_{B_n}$). Par homogénéité, on peut supposer que $\int f = \int g = 1$. Il suffit donc de démontrer que $\int h \geq 1$ sous l'hypothèse du théorème.

Soit μ et ν les probabilités de densités respectives f et g (elles admettent un moment d'ordre 2 car sont à support borné). Soit $\nabla\phi$ l'application de transport de Brenier telle que $\nabla\phi\#\mu = \nu$. Comme dans la preuve de l'inégalité de Brunn-Minkowski, nous considérons la mesure μ_t définie par

$$\mu_t = \nabla\phi_t\#\mu, \quad t \in [0, 1],$$

où pour presque tout $x \in \Omega = \text{int}(\text{dom } \phi)$,

$$\nabla\phi_t(x) = (1-t)x + t\nabla\phi(x).$$

Les ensembles X ou X_t sont définis comme dans la preuve de l'inégalité de Brunn-Minkowski. D'après le Lemme 9.20 et plus particulièrement (9.7),

$$\int h(x)dx \geq \int_{\nabla\phi_t(X_t)} h(y) dy = \int_{\Omega} h(\nabla\phi_t(x)) \det(\text{Hess}_x\phi_t)dx.$$

Par hypothèse, pour tout $x \in \Omega$

$$h(\nabla\phi_t(x)) = h((1-t)x + t\nabla\phi(x)) \geq f(x)^{1-t}g(\nabla\phi(x))^t,$$

et d'après le Lemme 10.2

$$\det(\text{Hess}_x\phi_t) = \det((1-t)I + t\text{Hess}_x\phi) \geq [\det(\text{Hess}_x\phi)]^t,$$

par conséquent on obtient

$$\int h(x)dx \geq \int_{\Omega} f(x)^{1-t} [g(\nabla\phi(x)) \det(\text{Hess}_x\phi)]^t dx.$$

Puisque μ et ν sont absolument continues, d'après le Théorème 9.21, l'équation de Monge-Ampère suivante est satisfaite pour presque tout $x \in \Omega$,

$$f(x) = g(\nabla\phi(x)) \det(\text{Hess}_x\phi).$$

On en déduit que

$$\int h(x)dx \geq \int f(x)dx = 1.$$

□

Chapitre 11

Inégalité isopérimétrique et inégalités de Sobolev sur \mathbb{R}^n .

L'inégalité isopérimétrique sur \mathbb{R}^n indique que parmi les ensembles compacts de volume donné, ceux dont la mesure de surface est minimale sont les boules euclidiennes.

Nous verrons que ce résultat est aussi la conséquence d'une inégalité de Sobolev optimale sur \mathbb{R}^n . De manière générale, étant donné $p \geq 1$, une inégalité de Sobolev donne l'exposant critique p^* pour lequel l'espace de Sobolev $W^{1,p}(\mathbb{R}^n) = \{f \in L^p(\mathbb{R}^n), \nabla f \in L^p(\mathbb{R}^n)\}$ est inclus dans $L^{p^*}(\mathbb{R}^n)$.

11.1 Inégalité isopérimétrique sur \mathbb{R}^n .

Résoudre un principe isopérimétrique dans un espace donné consiste à trouver les ensembles compacts qui minimisent leur surface à volume fixé. Sur l'espace euclidien, étant donné un sous-ensemble $A \subset \mathbb{R}^n$, la mesure de surface naturelle du bord $\partial A = \overline{A} \setminus \text{int}(A)$, notée $|\partial A|$, est définie par

$$|\partial A| = \liminf_{\varepsilon \rightarrow 0} \frac{|A + \varepsilon B_1| - |A|}{\varepsilon}.$$

Observons que pour tout $\alpha > 0$, $|\partial(\alpha A)| = \alpha^{n-1} |\partial A|$. Soit $V > 0$ un volume fixé et α tel que $V = \alpha^n |B_1|$. Par homogénéité,

$$\inf_{A, |A|=V} |\partial A| = \inf_{A, |A/\alpha|=|B_1|} |\partial A| = \inf_{A', |A'|=|B_1|} |\partial(\alpha A')| = \left(\frac{|B_1|}{V}\right)^{(n-1)/n} \inf_{A', |A'|=|B_1|} |\partial A'|.$$

Par conséquent, pour résoudre le problème isopérimétrique, il suffit de déterminer les ensembles A' de même volume que la boule euclidienne B_1 dont la surface est minimale.

Dans l'espace euclidien, les boules euclidiennes sont les ensembles de mesure de surface minimale. Ce résultat est une conséquence directe de l'inégalité de Brunn-Minkowski.

Théorème 11.1. (*Inégalité isopérimétrique euclidienne*) *Pour tout sous-ensemble mesurable A tel que $|A| = |B_1|$, on a*

$$|\partial A| \geq |\partial B_1|, \tag{11.1}$$

ou encore, de façon équivalente, pour tout sous-ensemble A compact,

$$|\partial A| \geq n |B_1|^{1/n} |A|^{(n-1)/n}.$$

Démonstration. L'équivalence entre les deux énoncés du théorème est une conséquence des propriétés d'homogénéité pour les mesures de volume et de surface et du fait que $|\partial B_1| = n|B_1|$ (puisque $B_1 + \varepsilon B_1 = (1 + \varepsilon)B_1$).

Soit A un sous-ensemble tel que $|A| = |B_1|$. Alors d'après l'inégalité de Brunn-Minkowski,

$$\begin{aligned} |A + \varepsilon B_1| &\geq \left(|A|^{1/n} + \varepsilon |B_1|^{1/n} \right)^n \\ &= (1 + \varepsilon)^n |B_1| \\ &= |(1 + \varepsilon)B_1| \\ &= |B_1 + \varepsilon B_1| \end{aligned}$$

Puisque $|A| = |B_1|$, l'inégalité isopérimétrique (11.1) découle ensuite de la définition de $|\partial A|$ et $|\partial B_1|$. \square

Dans le paragraphe suivant, nous allons établir une inégalité fonctionnelle appelée inégalité de Sobolev L^1 qui redonne l'inégalité isopérimétrique.

11.2 Inégalités de Sobolev

Pour $p \geq 1$, l'espace de Sobolev $W^{1,p}(\mathbb{R}^n)$ est l'ensemble des fonctions f dans $L^p(\mathbb{R}^n)$, différentiables et telles que $\nabla f \in L^p(\mathbb{R}^n)$. Pour $p \in [1, n)$, définissons

$$p^* = \frac{np}{n-p}.$$

On a $p^* \geq p$. Notons $\|f\|_p = (\int |f|^p)^{1/p}$.

Les inégalités de Sobolev que nous établissons dans ce paragraphe donnent l'existence d'une constante $S_n(p) > 0$ telle que pour tout $f \in W^{1,p}(\mathbb{R}^n)$,

$$\|f\|_{p^*} \leq S_n(p) \|\nabla f\|_p,$$

Par conséquent on a $W^{1,p}(\mathbb{R}^n) \subset L^{p^*}(\mathbb{R}^n)$. L'objectif de ce paragraphe est d'établir ces inégalités de Sobolev avec la constante optimale notée $S_n(p)$.

a. Inégalité de Sobolev L^1

Pour $p = 1$, $p^* = \frac{n}{n-1}$ et l'inégalité de Sobolev L^1 optimale est la suivante.

Théorème 11.2. (Inégalité de Sobolev L^1) Pour toute fonction f dans $W^{1,1}(\mathbb{R}^n)$,

$$\|f\|_{n/(n-1)} \leq \frac{1}{n|B_1|^{1/n}} \|\nabla f\|_1.$$

Soit A un sous-ensemble compact dont la mesure de Lebesgue du bord $\lambda(\partial A)$ est nulle. Appliquons ce résultat à la famille de fonctions lipschitziennes $f_\varepsilon, \varepsilon > 0$ définies par

$$f_\varepsilon(x) = \left[1 - \frac{d(x, A)}{\varepsilon} \right]_+, \quad x \in \mathbb{R}^n,$$

où $[h]_+ = \max(0, h)$ et $d(x, A) = \min_{y \in A} |y - x|$. D'après le Théorème de Rademacher, f_ε est presque partout différentiable. Pour presque tout x , on a

$$|\nabla f_\varepsilon(x)| \leq \frac{\mathbf{1}_{\overline{A_\varepsilon} \setminus \text{int}(A)}(x)}{\varepsilon}.$$

Par conséquent, pour tout $\varepsilon > 0$,

$$\|f_\varepsilon\|_{n/(n-1)} \leq \frac{1}{n|B_1|^{1/n}} \frac{|\overline{A_\varepsilon} \setminus \text{int}(A)|}{\varepsilon} = \frac{1}{n|B_1|^{1/n}} \frac{|\overline{A + \varepsilon B_1}| - |A|}{\varepsilon}.$$

Pour tout $x \in \mathbb{R}^n$, on a $\lim_{\varepsilon \rightarrow 0} f_\varepsilon(x) = \mathbf{1}_A(x)$. Ainsi par passage à la limite lorsque ε tend vers 0, le théorème de Lebesgue donne

$$|A|^{(n-1)/n} \leq \frac{1}{n|B_1|^{1/n}} |\partial A|,$$

Cette inégalité est exactement l'inégalité isopérimétrique optimale sur \mathbb{R}^n . Ceci montre par ailleurs l'optimalité de la constante dans l'inégalité de Sobolev L^1 , $S_n(1) = \frac{1}{n|B_1|^{1/n}}$.

Démonstration du Théorème 11.2. Par des arguments de densité, il suffit de démontrer le résultat pour des fonctions de classe \mathcal{C}^2 à support compact. Puisque presque sûrement $|\nabla f| = |\nabla|f||$ on peut considérer uniquement des fonctions mesurables positives. Par homogénéité, on peut se restreindre au cas où $\int f^{n/(n-1)} = 1$.

Soit μ la probabilité de densité $f^{n/(n-1)}$ et ν la loi uniforme sur B_1 de densité $\mathbf{1}_{B_1}/|B_1|$. Soit $\nabla\phi$ le transport de Brenier de μ sur ν , $\nabla\phi\#\mu = \nu$. Pour μ -presque tout x , $\nabla\phi(x) \in B_1$ donc $|\nabla\phi(x)| \leq 1$. D'après le Théorème 9.21, l'équation de Monge-Ampère suivante est satisfaite pour $-\mu$ -presque tout x ,

$$f^{n/(n-1)}(x) = \frac{\mathbf{1}_{B_1}(\nabla\phi(x))}{|B_1|} \det \text{Hess}_x \phi.$$

D'après le Lemme 10.2, on a

$$(\det \text{Hess}_x \phi)^{1/n} \leq \frac{\text{Tr}(\text{Hess}_x \phi)}{n} = \frac{\Delta_A \phi(x)}{n},$$

($\Delta_A \phi$ défini ainsi s'interprète comme le laplacien de ϕ au sens d'Aleksandrov). Par conséquent, on a pour μ -presque tout x

$$|B_1|^{1/n} f^{1/(n-1)}(x) \leq \frac{\Delta_A \phi(x)}{n},$$

et ainsi en intégrant cette inégalité par rapport à la mesure μ ,

$$|B_1|^{1/n} \int f f^{1/(n-1)} \leq \frac{1}{n} \int f \Delta_A \phi.$$

A ce point de la démonstration, lorsque $\Delta_A \phi$ correspond au laplacien usuel, on peut intégrer par partie le membre de droite de l'inégalité précédente. Le lemme suivant montre que le résultat ainsi obtenu par intégration par partie reste valable lorsque $\Delta_A \phi$ est le laplacien de ϕ au sens d'Aleksandrov, $\Delta_A \phi(x) = \text{Tr}(\text{Hess}_x \phi)$. La preuve de ce lemme est donnée à la fin de ce paragraphe.

Lemme 11.3. *Soit $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe. Alors pour toute fonction f positive, de classe \mathcal{C}^2 , à support compact inclus dans $\Omega = \text{int}(\text{dom } \phi)$, on a*

$$\int f \Delta_A \phi \leq - \int \nabla f \cdot \nabla \phi.$$

En appliquant ce lemme et le fait que $|\nabla\phi| \leq 1$, nous obtenons l'inégalité de Sobolev L^1 ,

$$|B_1|^{1/n} \int f^{n/(n-1)} \leq -\frac{1}{n} \int \nabla f \cdot \nabla \phi \leq \frac{1}{n} \int |\nabla f|.$$

□

Démonstration du Lemme 11.3. Soit f une fonction positive, de classe \mathcal{C}^2 , à support compact K inclus dans Ω . Soit (e_1, \dots, e_n) la base canonique de \mathbb{R}^n . Etant donné $i \in \{1, \dots, n\}$, g une fonction sur \mathbb{R}^n et $h > 0$, on définit la fonction $g_h = g_h^i$ par

$$g_h(x) = g_h^i(x) = \frac{g(x + he_i) + g(x - he_i) - 2g(x)}{h^2}, \quad x \in \mathbb{R}^n.$$

Notons $\partial_{ii}\phi(x) = \text{Hess}_x\phi(e_i) \cdot e_i$ la dérivée seconde au sens d'Aleksandrov dans la direction e_i . Cette dérivée existe presque sûrement sur $\text{dom } \phi$ d'après le théorème d'Aleksandrov. D'après la Proposition 9.9, pour presque tout $x \in \Omega$

$$\lim_{h \rightarrow 0} \phi_h^i(x) = \partial_{ii}\phi(x).$$

Soit $\varepsilon > 0$ tel que $\overline{K_\varepsilon} \subset \Omega$. La fonction ϕ est continue donc bornée sur le compact $\overline{K_\varepsilon}$. Pour $|h| < \varepsilon$, si $x \notin \overline{K_\varepsilon}$ alors $f_h(x) = 0$, et par suite

$$\int \phi f_h = \int_{\overline{K_\varepsilon}} \phi f_h.$$

Puisque pour $|h| \leq \varepsilon$, $K \subset \overline{K_\varepsilon} + he_i$ et $f(x) = 0$ pour $x \notin K$, on a

$$\int_{\overline{K_\varepsilon}} \phi(x) f(x + he_i) dx = \int_{\overline{K_\varepsilon} + he_i} \phi(x - he_i) f(x) dx = \int \phi(x - he_i) f(x) dx,$$

et

$$\int_{\overline{K_\varepsilon}} \phi(x) f(x - he_i) dx = \int_{\overline{K_\varepsilon} - he_i} \phi(x + he_i) f(x) dx = \int \phi(x - he_i) f(x) dx.$$

On obtient donc, pour tout $|h| \leq \varepsilon$,

$$\int \phi f_h = \int_{\overline{K_\varepsilon}} \phi f_h = \int_K \phi_h f.$$

Puisque ϕ est convexe, $\phi_h(x) f(x) \geq 0$ pour tout $x \in K$, et le Lemme de Fatou donne

$$\int_K f \partial_{ii}\phi \leq \liminf_{h \rightarrow 0} \int_K \phi_h f = \liminf_{h \rightarrow 0} \int_{\overline{K_\varepsilon}} \phi f_h. \quad (11.2)$$

On sait que pour tout $x \in \mathbb{R}^n$, $\lim_{h \rightarrow 0} f_h(x) = \partial_{ii}f(x)$. La fonction f_h est bornée sur \mathbb{R}^n , en effet

$$|f_h(x)| = \frac{1}{h^2} \left| \int_0^h \left(\int_{-t}^t \partial_{ii}f(x + ue_i) du \right) dt \right| \leq \sup_{z \in \mathbb{R}^n} |\partial_{ii}f(z)| < \infty.$$

Puisque

$$\int_{\overline{K_\varepsilon}} |\phi f_h| \leq |\overline{K_\varepsilon}| \sup_{z \in \mathbb{R}^n} |\partial_{ii}f(z)| \sup_{z \in \overline{K_\varepsilon}} |\phi(z)| < \infty, \quad |h| < \varepsilon,$$

en appliquant le théorème de Lebesgue, on obtient

$$\lim_{h \rightarrow 0} \int_{\overline{K_\varepsilon}} \phi f_h = \int_{\overline{K_\varepsilon}} \phi \partial_{ii}f = \int \phi \partial_{ii}f,$$

et ainsi en utilisant l'inégalité (11.2) et après intégration par partie,

$$\int f \partial_{ii}\phi \leq \int \phi \partial_{ii}f = - \int \partial_i \phi \partial_i f.$$

La démonstration du Lemme se termine en sommant cette inégalité pour i entre 1 et n . □

b. Inégalité de Sobolev L^p , $1 < p < n$.

Une preuve similaire à celle de l'inégalité de Sobolev L^1 permet d'obtenir l'inégalité de Sobolev L^p dans \mathbb{R}^n . Dans cette preuve la densité qui remplace la densité de la loi uniforme sur B_1 est la densité h_p définie par

$$h_p(x) = \frac{1}{(\sigma_p + |x|^{p'})^{(n-p)/p}}, \quad x \in \mathbb{R}^n,$$

où $p' = \frac{p}{p-1}$ est l'exposant conjugué de p et σ_p la constante telle que $\int h_p^{p^*} = 1$ avec $p^* = \frac{np}{n-p}$,

$$\sigma_p = \left[\int \frac{dy}{(\sigma_p + |x|^{p'})^{(n-p)/p}} \right]^{\frac{p}{n+1}}.$$

Théorème 11.4. *Soit $p \in (1, n)$ ($n \geq 2$), et $p^* = \frac{np}{n-p}$. Pour toute fonction $f \in W^{1,p}(\mathbb{R}^n)$,*

$$\|f\|_{p^*} \leq \frac{\|\nabla f\|_p}{\|\nabla h_p\|_p},$$

avec égalité si $f = h_p$. Par conséquent $S_n(p) = 1/\|\nabla h_p\|_p$ est la constante optimale de cette inégalité de Sobolev.

Démonstration. Comme dans la preuve de l'inégalité de Sobolev L^1 , on peut supposer que la fonction f est positive de classe \mathcal{C}^2 et à support compact. Par homogénéité de l'inégalité de Sobolev, on peut aussi supposer que $\int f^{p^*} = 1$. On veut donc montrer que $\|\nabla h_p\|_p \leq \|\nabla f\|_p$.

Soit μ la loi de densité $F = f^{p^*}$ et ν la loi de densité $H = h_p^{p^*}$. On peut vérifier simplement que ν admet un moment d'ordre 2. Soit $\nabla\phi$ le transport de Brenier de μ sur ν , $\nabla\phi\#\mu = \nu$. L'équation de Monge-Ampère s'écrit pour μ -presque tout x ,

$$F(x) = H(\nabla\phi(x)) \det \text{Hess}_x \phi.$$

Puisque, d'après le Lemme 10.2,

$$(\det \text{Hess}_x \phi)^{1/n} \leq \frac{\Delta_A \phi(x)}{n}, \tag{11.3}$$

on a pour μ -presque tout x ,

$$H^{-1/n}(\nabla\phi(x)) \leq F^{-1/n}(x) \frac{\Delta_A \phi(x)}{n},$$

et par conséquent

$$\int h_p^{p^*(1-1/n)} = \int H^{-1/n} d\nu = \int H^{-1/n}(\nabla\phi) d\mu \leq \frac{1}{n} \int f^{p^*(1-1/n)} \Delta_A \phi.$$

D'après le Lemme 11.3, et puisque $p^*(1 - 1/n) - 1 = p^*/p'$

$$\int f^{p^*(1-1/n)} \Delta_A \phi \leq - \int \nabla(f^{p^*(1-1/n)}) \cdot \nabla\phi = -p^*(1 - 1/n) \int f^{p^*/p'} \nabla f \cdot \nabla\phi.$$

D'après l'inégalité de Hölder, et puisque $\nabla\phi\#\mu = \nu$, on a

$$\begin{aligned} - \int f^{p^*/p'} \nabla f \cdot \nabla\phi &\leq \int f^{p^*/p'} |\nabla\phi| |\nabla f| \\ &\leq \left(\int |\nabla\phi|^{p'} d\mu \right)^{1/p'} \left(\int |\nabla f|^p \right)^{1/p} \\ &= \left(\int |y|^{p'} h_p^{p^*}(y) dy \right)^{1/p'} \|\nabla f\|_p. \end{aligned}$$

Finalement on obtient

$$\int h_p^{p^*(1-1/n)} \leq \frac{p^*(1-1/n)}{n} \left(\int |y|^{p'} h_p^{p^*}(y) dy \right)^{1/p'} \|\nabla f\|_p.$$

Pour obtenir le résultat souhaité, il suffit de montrer qu'il y a égalité dans toute la suite des inégalités précédentes lorsque $f = h_p$. Dans ce cas, on a nécessairement $\nabla\phi(x) = x$ et $\text{Hess}_x\phi = Id$ et par conséquent l'inégalité (11.3) est une égalité. L'autre inégalité utilisée est l'inégalité de Hölder. La propriété fondamentale de l'application h_p est de réaliser l'égalité dans cette inégalité de Hölder. Plus précisément, on a

$$-\nabla h_p(x) \cdot x = |\nabla h_p(x)| |x|,$$

donc

$$- \int h_p^{p^*/p'}(x) \nabla h_p(x) \cdot x dx = \int h_p^{p^*/p'}(x) |x| |\nabla h_p(x)| dx.$$

Par ailleurs, l'égalité est réalisée dans l'inégalité de Hölder lorsque l'une des fonctions intégrée dans le membre de droite est un multiple de l'autre. Ici on peut vérifier que

$$|\nabla h_p(x)|^p = (cste) |x|^{p'} h_p^{p^*}(x).$$

et donc

$$\int h_p^{p^*/p'}(x) |x| |\nabla h_p(x)| dx = \left(\int |x|^{p'} h_p^{p^*}(x) dx \right)^{1/p'} \|\nabla h_p\|_p.$$

□

Chapitre 12

Inégalités de Young optimales et de Brascamp-Lieb

L'objectif de ce chapitre est de présenter les inégalités de Young optimales et leur lien avec les inégalités de Brascamp-Lieb.

12.1 Inégalités de Young optimale

Soient f et g des fonctions intégrables sur \mathbb{R}^n . Le produit de convolution de ces deux fonctions est la fonction définie par

$$f * g(x) = \int f(x-y)g(y)dy, \quad x \in \mathbb{R}^n.$$

L'inégalité de Young la plus largement connue est la suivante, pour tous $p, q, r \in [1, +\infty]$, tels que

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r},$$

on a

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Remarquons que si $r < +\infty$, alors nécessairement $p \leq r$ et $q \leq r$, autrement dit l'intérêt de cette inégalité est de montrer qu'il suffit que $f \in L^q(\mathbb{R}^n)$ et $g \in L^p(\mathbb{R}^n)$ pour que $f * g \in L^r(\mathbb{R}^n)$ et pourtant $r \geq \max(p, q)$.

Cette inégalité de Young s'obtient simplement à partir de l'inégalité de Hölder en suivant les étapes de l'exercice suivant.

Exercice 12.1. 1. En utilisant deux fois l'inégalité de Hölder, montrer que pour tous réels positifs α, β, γ tels que $\alpha + \beta + \gamma = 1$, pour toutes fonctions intégrables positives f, g, h sur \mathbb{R}^n ,

$$\int f^\alpha g^\beta h^\gamma \leq \left(\int f \right)^\alpha \left(\int g \right)^\beta \left(\int h \right)^\gamma.$$

2. Soient $p, q, r \in [1, +\infty]$, tels que $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. En choisissant $\alpha = 1/r, \beta = 1/p - 1/r, \gamma = 1/q - 1/r$, montrer que pour tout $x \in \mathbb{R}^n$,

$$f * g(x) \leq \|f\|_p^{1-p/r} \|g\|_q^{1-q/r} (f^p * g^q(x))^{1/r}.$$

3. En déduire l'inégalité de Young $\|f * g\|_r \leq \|f\|_p \|g\|_q$.

Excepté pour quelques valeurs de p, q, r , l'inégalité de Young que nous venons de présenter n'est pas optimale la constante 1 peut être améliorée comme le montre le théorème suivant.

Théorème 12.1. (Inégalité de Young optimale) Pour tous $p, q, r \in [1, +\infty]$ tels que

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r},$$

on a

$$\|f * g\|_r \leq \left(\frac{C_p C_q}{C_r} \right)^n \|f\|_p \|g\|_q,$$

avec $C_\infty = C_1 = 1$ et pour $p \in]1, +\infty[$,

$$C_p = \sqrt{\frac{p^{1/p}}{p^{1/p'}}}, \quad \frac{1}{p} + \frac{1}{p'} = 1,$$

p' est appelé exposant conjugué de p .

Nous verrons plus loin que la constante $\left(\frac{C_p C_q}{C_r} \right)^n$ est optimale donc toujours inférieur à 1.

Voici une autre formulation de l'inégalité de Young optimale. D'après l'inégalité de Hölder, on sait que

$$\|f * g\|_r = \sup_h \frac{|\int f * g h|}{\|h\|_{r'}}.$$

Par conséquent, puisque $C_r = 1/C_{r'}$, l'inégalité de Young optimale s'écrit encore, pour toute fonction f, g, h

$$\left| \iint f(x-y)g(y)h(x) dx dy \right| = \left| \int f * g h \right| \leq (C_p C_q C_{r'})^n \|f\|_p \|g\|_q \|h\|_{r'}.$$

Sous cette forme l'inégalité de Young est en fait un cas particulier de l'inégalité de Brascamp-Lieb, inégalité que nous présentons dans la partie qui suit.

12.2 Inégalités de Brascamp-Lieb et de Brascamp-Lieb inverse

Théorème 12.2. (Inégalité de Brascamp-Lieb) Soient N, n_1, n_2, \dots, n_m des entiers et c_1, \dots, c_m des réels positifs tels que $n_i \leq N$ et

$$\sum_{i=1}^m c_i n_i = N.$$

Soient $B_i : \mathbb{R}^N \rightarrow \mathbb{R}^{n_i}$ des applications linéaires surjectives telles que

$$\bigcap_{i=1}^m \ker B_i = \{0\}.$$

Etant données f_1, \dots, f_m des fonctions intégrables positives, $f_i \in L^1(\mathbb{R}^{n_i})$, on note

$$I(f_1, \dots, f_m) = \int_{\mathbb{R}^N} \prod_i f_i^{c_i}(B_i x) dx.$$

Soit \bar{I} la constante optimale dans l'inégalité suivante (dite de Brascamp-Lieb)

$$I(f_1, \dots, f_m) \leq \bar{I} \prod_i \left(\int_{\mathbb{R}^{n_i}} f_i \right)^{c_i}.$$

Alors cette constante optimale \bar{I} est atteinte sur l'ensemble des densités gaussiennes centrées, c'est à dire

$$\begin{aligned} \bar{I} &:= \sup_{A_1, \dots, A_m} \left\{ I(\gamma_{A_1^{-1}}, \dots, \gamma_{A_m^{-1}}), \forall x_i \in \mathbb{R}^{n_i}, \gamma_{A_i^{-1}}(x_i) = \frac{e^{-\frac{1}{2}x_i \cdot A_i x_i}}{(2\pi)^{n_i/2} \sqrt{\det(A_i^{-1})}} \right\} \\ &= \sup_{A_1, \dots, A_m} \left\{ \frac{I(\gamma_1, \dots, \gamma_m)}{\prod_i \left(\int_{\mathbb{R}^{n_i}} \gamma_i \right)^{c_i}}, \forall x_i \in \mathbb{R}^{n_i}, \gamma_i(x_i) = e^{-x_i \cdot A_i x_i} \right\}, \end{aligned}$$

où les matrices A_i sont symétriques définies positives de taille n_i . Plus précisément, on a $\bar{I} = 1/\sqrt{D}$ avec

$$D = \inf_{A_1, \dots, A_m} \frac{\det \left(\sum_i c_i B_i^t A_i B_i \right)}{\prod_i (\det A_i)^{c_i}}.$$

Le dernier point de ce théorème est une simple conséquence du fait que pour toute matrice A symétrique définie positive de taille n ,

$$\int e^{-x \cdot A x} dx = \frac{\pi^{n/2}}{\sqrt{\det A}} = \pi^{n/2} \sqrt{\det A^{-1}}.$$

Par suite $\int \gamma_i = \pi^{n_i/2} \sqrt{\det A_i}$ et

$$\begin{aligned} I(\gamma_1, \dots, \gamma_m) &= \int_{\mathbb{R}^N} \exp \left(- \sum_i c_i B_i(x) \cdot A_i B_i(x) \right) dx \\ &= \int_{\mathbb{R}^N} \exp \left(- \sum_i c_i x \cdot B_i^t A_i B_i(x) \right) dx \\ &= \int_{\mathbb{R}^N} \exp \left(-x \cdot \left(\sum_i c_i B_i^t A_i B_i(x) \right) \right) dx \\ &= \frac{\pi^{N/2}}{\sqrt{\det \left(\sum_i c_i B_i^t A_i B_i \right)}}. \end{aligned}$$

La dernière égalité est vraie car la matrice $A = \sum_i c_i B_i^t A_i B_i$, de taille N , est symétrique définie positive puisque $\bigcap_{i=1}^m \ker B_i = \{0\}$.

Démonstration de l'inégalité de Young optimale à partir de l'inégalité de Brascamp-Lieb.

Il suffit de montrer que pour toute fonction f, g, h positive sur \mathbb{R}^n et pour tout $p, q, r' \geq 1$ tels que

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r'} = 2,$$

on a

$$\iint f(x-y)g(y)h(x) dx dy \leq (C_p C_q C_{r'})^n \|f\|_p \|g\|_q \|h\|_{r'}.$$

En posant $c_1 = \frac{1}{p}$, $c_2 = \frac{1}{q}$, $c_3 = \frac{1}{r}$, $n_1 = n_2 = n_3 = n$ et $N = 2n$, on a bien

$$c_1 n_1 + c_2 n_2 + c_3 n_3 = N.$$

Par ailleurs, si $f_1 = f^p$, $f_2 = g^q$, $f_3 = h^r$ et

$$B_1(x) = x_1 - x_2, \quad B_2(x) = x_2, \quad B_3(x) = x_1, \quad x = (x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^N,$$

alors $\ker B_1 \cap \ker B_2 \cap \ker B_3 = \{0\}$ et l'inégalité de Young s'écrit

$$I(f_1, f_2, f_3) \leq (C_p C_q C_{r'})^n \left(\int_{\mathbb{R}^n} f_1 \right)^{c_1} \left(\int_{\mathbb{R}^n} f_2 \right)^{c_2} \left(\int_{\mathbb{R}^n} f_3 \right)^{c_3}.$$

Pour obtenir l'inégalité de Young optimale à partir de l'inégalité de Brascamp-Lieb, il suffit donc de vérifier que

$$\bar{I} = (C_p C_q C_{r'})^n.$$

Pour simplifier, nous allons seulement vérifier cette identité pour $n = 1$. On aura ainsi obtenu l'inégalité de Young optimale en dimension 1. L'inégalité de Young optimale en dimension n peut ensuite se déduire par tensorisation.

Puisque $\bar{I} = 1/\sqrt{D}$, nous devons déterminer

$$D = \inf_{A_1, A_2, A_3 \in \mathbb{R}^+} \frac{\det(\sum_i c_i B_i^t A_i B_i)}{A_1^{c_1} A_2^{c_2} A_3^{c_3}}.$$

D'après la définition des B_i ,

$$B_1 = (1, -1), \quad B_2 = (0, 1), \quad B_3 = (1, 0),$$

et par conséquent,

$$\sum_i c_i B_i^t A_i B_i = \begin{pmatrix} c_1 A_1 + c_3 A_3 & -c_1 A_1 \\ -c_1 A_1 & c_1 A_1 + c_2 A_2 \end{pmatrix},$$

et

$$D = \inf_{A_1, A_2, A_3 \in \mathbb{R}^+} \frac{c_1 c_2 A_1 A_2 + c_1 c_3 A_1 A_3 + c_2 c_3 A_2 A_3}{A_1^{c_1} A_2^{c_2} A_3^{c_3}}.$$

Après dérivation, sachant que $c_1 + c_2 + c_3 = 2$, on obtient que les extrémums sont les points (A_1, A_2, A_3) tels que $(1 - c_1)c_1 A_1 = (1 - c_2)c_2 A_2 = (1 - c_3)c_3 A_3$, c'est à dire

$$(A_1, A_2, A_3) = \left(\frac{C}{(1 - c_1)c_1}, \frac{C}{(1 - c_2)c_2}, \frac{C}{(1 - c_3)c_3} \right).$$

On a donc

$$\begin{aligned} D &= \frac{[(1 - c_1)(1 - c_2)]^{-1} + [(1 - c_1)(1 - c_3)]^{-1} + [(1 - c_2)(1 - c_3)]^{-1}}{[c_1(1 - c_1)]^{-c_1} [c_2(1 - c_2)]^{-c_2} [c_3(1 - c_3)]^{-c_3}} \\ &= \frac{c_1^{c_1} c_2^{c_2} c_3^{c_3}}{(1 - c_1)^{1-c_1} (1 - c_2)^{1-c_2} (1 - c_3)^{1-c_3}} \\ &= (C_p C_q C_{r'})^2. \end{aligned}$$

□

La preuve par transport de l'inégalité de Brascamp-Lieb que nous présentons donne simultanément l'inégalité de Brascamp-Lieb présentée dans le théorème qui suit..

Théorème 12.3. (*Inégalité de Brascamp-Lieb inverse*) Avec les mêmes notations et hypothèses que dans le théorème de Brascamp-Lieb précédent, étant données des fonctions intégrables positives g_1, \dots, g_m , $g_i \in L^1(\mathbb{R}^{n_i})$, soit

$$J(g_1, \dots, g_m) = \int_{\mathbb{R}^N} \left[\sup_{x_1, \dots, x_m / x = \sum c_i B_i^t(x_i)} \prod_i g_i^{c_i}(x_i) \right] dx.$$

Alors la constante optimale \underline{J} dans l'inégalité (dite de Brascamp-Lieb inverse)

$$J(g_1, \dots, g_m) \geq \underline{J} \prod_i \left(\int_{\mathbb{R}^{n_i}} g_i \right)^{c_i}$$

est atteinte pour des fonctions gaussiennes centrées,

$$\begin{aligned} \underline{J} &:= \inf_{A_1, \dots, A_m} \left\{ J(\gamma_{A_1^{-1}}, \dots, \gamma_{A_m^{-1}}), \forall x_i \in \mathbb{R}^{n_i}, \gamma_{A_i^{-1}}(x_i) = \frac{e^{-\frac{1}{2}x_i \cdot A_i x_i}}{(2\pi)^{n_i/2} \sqrt{\det(A_i^{-1})}} \right\} \\ &= \inf_{A_1, \dots, A_m} \left\{ \frac{J(\gamma_1, \dots, \gamma_m)}{\prod_i \left(\int_{\mathbb{R}^{n_i}} \gamma_i \right)^{c_i}}, \forall x_i \in \mathbb{R}^{n_i}, \gamma_i(x_i) = e^{-x_i \cdot A_i x_i} \right\}, \end{aligned}$$

et de plus,

$$\underline{J} = \frac{1}{\bar{I}} = \sqrt{D}.$$

Pour illustrer ces les théorèmes précédents, choisissons $n_1 = n_2 = n$, $c_1 = 1 - t$, $c_2 = t$, $t \in [0, 1]$, $N = n$, $g_1 = f$, $g_2 = g$, et pour tout $x \in \mathbb{R}^n$, $B_1(x) = B_2(x) = x$. On a alors

$$D = \inf_{A_1, A_2} \frac{\det(tA_1 + (1-t)A_2)}{(\det A_1)^t (\det A_2)^{1-t}}.$$

D'après le Lemme 10.2, la propriété de log-concavité du déterminant implique que $D = 1$.

L'inégalité de Brascamp-Lieb inverse s'écrit alors

$$\int \sup_{x, y / z = tx + (1-t)y} [f(x)^{1-t} g(y)^t] dz \geq \left(\int f(x) dx \right)^{1-t} \left(\int g(y) dy \right)^t.$$

Cette inégalité correspond à l'inégalité de Prékopa-Leindler.

L'inégalité de Brascamp-Lieb correspond quant à elle à l'inégalité de Hölder,

$$\int f^{1-t}(x) g^t(x) \leq \left(\int f(x) dx \right)^{1-t} \left(\int g(x) dx \right)^t.$$

Ainsi, on peut interpréter l'inégalité de Prékopa-Leindler, comme une forme duale de l'inégalité de Hölder.

12.3 Démonstration des inégalités de Brascamp-Lieb

Pour démontrer les deux inégalités de Brascamp-Lieb, par densité et homogénéité, on peut supposer que toutes les fonctions positives f_1, \dots, f_m et g_1, \dots, g_m sont continues à support compact et d'intégrale 1.

La preuve des inégalités de Brascamp-Lieb se fait alors en deux étapes. L'objectif est alors d'obtenir que pour toutes ces fonctions intégrables

$$I(f_1, \dots, f_m) \leq \bar{I} = \frac{1}{\sqrt{D}} \quad \text{et} \quad \sqrt{D} = \underline{J} \leq J(g_1, \dots, g_m)$$

autrement dit

$$DI(f_1, \dots, f_m) \leq D\bar{I} = \underline{J} \leq J(g_1, \dots, g_m).$$

La première étape algébrique consiste à montrer que

$$\bar{I}\underline{J} = 1. \tag{12.1}$$

Puisque $\bar{I} = 1/\sqrt{D}$, on a alors $D\bar{I} = \underline{J}$.

La seconde étape consiste à établir par méthode de transport que pour toutes fonctions positives d'intégrale 1, f_1, \dots, f_m et g_1, \dots, g_m ,

$$DI(f_1, \dots, f_m) \leq J(g_1, \dots, g_m) \tag{12.2}$$

Ainsi, en optimisant sur toutes les fonctions g_i gaussiennes d'intégrale 1, par définition de \underline{J} , (12.2) donne l'inégalité de Brascamp-Lieb

$$DI(f_1, \dots, f_m) \leq \underline{J} = D\bar{I}.$$

Ou alors, en optimisant sur toutes les fonctions f_i gaussiennes d'intégrale 1, par définition de \underline{I} , (12.2) donne l'inégalité de Brascamp-Lieb inverse

$$\underline{J} = D\bar{I} \leq J(g_1, \dots, g_m).$$

Preuve de l'inégalité (12.2) par transport.

Soient μ_i et ν_i les probabilités de densités respectives f_i et g_i sur \mathbb{R}^{n_i} et $\nabla\phi_i$ l'application de transport de Brenier telle que $\nabla\phi_i\#\mu_i = \nu_i$. Soit $\Omega_i = \text{int}(\text{dom}\phi_i)$ et X_i l'ensemble des points où l'application convexe ϕ_i est deux fois différentiable au sens d'Aleksandrov et la Hessienne est inversible. D'après le Theorème 9.21, μ_i est portée par X_i et ν_i est portée par $\nabla\phi_i(X_i)$ donc on peut supposer que f_i est nulle en dehors de X_i et g_i est nulle en dehors de $\nabla\phi_i(X_i)$. Par ailleurs, puisque f_i est continue, tout point de \mathbb{R}^n est un point de Lebesgue pour μ_i . Par suite la Proposition 9.18 indique que l'équation de Monge-Ampère est satisfaite pour tout $x_i \in X_i$,

$$f_i(x_i) = g(\nabla\phi_i(x_i)) \det(\text{Hess}_{x_i}\phi_i), \quad i = 1, \dots, m.$$

Les applications de transport $\nabla\phi_i$ interviennent dans le schéma suivant, qui met en relation l'espace \mathbb{R}^{n_i} et l'espace \mathbb{R}^N ,

$$\begin{array}{ccccccc} B_i^{-1}(X_i) \subset \mathbb{R}^N & \xrightarrow{B_i} & X_i \subset \mathbb{R}^{n_i} & \xrightarrow{\nabla\phi_i} & \nabla\phi_i(X_i) \subset \mathbb{R}^{n_i} & \xrightarrow{B_i^t} & \mathbb{R}^N \\ x & \mapsto & B_i(x) = x_i & \mapsto & \nabla\phi_i(x_i) = y_i & \mapsto & B_i^t(y_i) \end{array} .$$

Soit $Y = \bigcap_{i=1}^m B_i^{-1}(X_i)$. Les équations de Monge-Ampère précédentes indiquent que pour tout $x \in Y$,

$$f_i(B_i(x)) = g(\nabla\phi_i(B_i(x))) \det(\text{Hess}_{B_i(x)}\phi_i), \quad i = 1, \dots, m. \quad (12.3)$$

Considérons l'application convexe $\Phi : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$ définie par

$$\Phi(x) = \sum_{i=1}^m c_i \phi_i(B_i(x)), \quad x \in \mathbb{R}^N.$$

Son domaine de définition dans \mathbb{R}^N est

$$\text{dom } \Phi = \bigcap_{i=1}^m B_i^{-1}(\text{dom } \phi_i).$$

Notons $\Omega = \text{int}(\text{dom } \Phi)$. Si $B_i(j) \in \mathbb{R}^{n_i}$ représente la j -ième colonne de la matrice B_i , $\phi(B_i(x)) = \phi\left(\sum_{j=1}^N \bar{x}_j B_i(j)\right)$ pour $x = (\bar{x}_1, \dots, \bar{x}_N)$ et les dérivées partielles de $\phi_i \circ B_i$, si elles existent, sont données par

$$\partial_j(\phi_i(B_i(x))) = \nabla\phi_i(B_i(x)) \cdot B_i(j), \quad j = 1, \dots, N.$$

Par suite, en tout point x pour lequel, pour tout i , ϕ_i est différentiable au point $B_i(x)$, en particulier pour tout $x \in Y \subset \Omega$, on a

$$\nabla\Phi(x) = \sum_i c_i B_i^t(\nabla\phi_i(B_i(x))), \quad x \in \mathbb{R}^N. \quad (12.4)$$

Par ailleurs, Φ est deux fois différentiable au sens d'Aleksandrov sur Y , pour tout $x \in Y$,

$$\text{Hess}_x \Phi = \sum_i c_i B_i^t(\text{Hess}_{B_i(x)}\phi_i) B_i.$$

Notons X l'ensemble des points $x \in \mathbb{R}^N$ où $\text{Hess}_x \Phi$ est inversible ($Y \subset X$). D'après le Lemme 9.20 et plus particulièrement (9.7), on a

$$\begin{aligned} J(g_1, \dots, g_m) &= \int_{\mathbb{R}^N} \left[\sup_{y_1, \dots, y_m / y = \sum c_i B_i^t(y_i)} \prod_i g_i^{c_i}(y_i) \right] dy \\ &\geq \int_{\nabla\Phi(X)} \left[\sup_{y_1, \dots, y_m / y = \sum c_i B_i^t(y_i)} \prod_i g_i^{c_i}(y_i) \right] dy \\ &= \int_{\Omega} \left[\sup_{y_1, \dots, y_m / \nabla\Phi(x) = \sum c_i B_i^t(y_i)} \prod_i g_i^{c_i}(y_i) \right] \det(\text{Hess}_x \Phi) dx \\ &\geq \int_Y \left[\sup_{y_1, \dots, y_m / \nabla\Phi(x) = \sum c_i B_i^t(y_i)} \prod_i g_i^{c_i}(y_i) \right] \det(\text{Hess}_x \Phi) dx. \end{aligned}$$

Par définition de la constante D , on a alors pour tout $x \in Y$,

$$\det(\text{Hess}_x \Phi) \geq D \prod_i [\det(\text{Hess}_{B_i(x)}\phi_i)]^{c_i}.$$

Par conséquent, puisque les égalités (12.4) et (12.3) sont satisfaites pour tout $x \in Y$, on obtient

$$\begin{aligned}
J(g_1, \dots, g_m) &\geq D \int_Y \prod_i [g_i(\nabla \phi_i(B_i(x))) \det(\text{Hess}_{B_i(x)} \phi_i)]^{c_i} dx \\
&= D \int_Y \prod_i f_i(B_i(x))^{c_i} dx \\
&= D \int_{\mathbb{R}^N} \prod_i f_i(B_i(x))^{c_i} dx \\
&= D I(f_1, \dots, f_m).
\end{aligned}$$

L'avant dernière égalité découle du fait que $\prod_i f_i(B_i(x))^{c_i}$ est nulle pour tout point x en dehors de Y . Ceci termine la preuve de l'inégalité (12.2). \square

Preuve de l'égalité (12.1).

Soit $A = \sum_i c_i B_i^t A_i B_i$, A est une matrice symétrique définie positive A . Soit Q_A et \tilde{Q} les formes quadratiques définies par

$$Q_A(y) = \frac{1}{2} A(y) \cdot y, \quad y \in \mathbb{R}^n,$$

et

$$\tilde{Q}(x) = \frac{1}{2} \inf_{x_1, \dots, x_m / x = \sum c_i B_i^t(x_i)} \left\{ \sum_i c_i A_i^{-1}(x_i) \cdot x_i \right\}, \quad x \in \mathbb{R}^n.$$

La preuve de l'égalité (12.1) repose sur le lemme suivant démontré à la fin de cette section.

Lemme 12.4. $Q_{A^{-1}} = \tilde{Q}$.

Notons $\gamma_{A_1}, \dots, \gamma_{A_m}$ les densités des lois gaussiennes centrées de matrices de covariance inversibles A_i , alors le Lemme 12.4 donne

$$\begin{aligned}
J(\gamma_{A_1^{-1}}, \dots, \gamma_{A_m^{-1}}) &= \frac{1}{(2\pi)^{N/2} \prod_i (\det A_i^{-1})^{c_i/2}} \int_{\mathbb{R}^N} e^{-\tilde{Q}(x)} dx \\
&= \frac{1}{(2\pi)^{N/2} \prod_i (\det A_i^{-1})^{c_i/2}} \int_{\mathbb{R}^N} e^{-Q_{A^{-1}}(x)} dx \\
&= \frac{\sqrt{\det A^{-1}}}{\prod_i (\det A_i^{-1})^{c_i/2}} \\
&= \frac{\prod_i (\det A_i)^{c_i/2}}{\sqrt{\det A}} \\
&= \left(\frac{1}{(2\pi)^{N/2} \prod_i (\det A_i)^{c_i/2}} \int_{\mathbb{R}^N} e^{-\frac{1}{2} x \cdot A(x)} dx \right)^{-1} \\
&= \frac{1}{I(\gamma_{A_1}, \dots, \gamma_{A_m})}
\end{aligned}$$

Par suite

$$\underline{J} = \inf_{A_1, \dots, A_m} J(\gamma_{A_1}, \dots, \gamma_{A_m}) = \inf_{A_1, \dots, A_m} J(\gamma_{A_1^{-1}}, \dots, \gamma_{A_m^{-1}}) = \frac{1}{\sup_{A_1, \dots, A_m} I(\gamma_{A_1}, \dots, \gamma_{A_m})} = \frac{1}{\bar{I}}.$$

\square

Démonstration du Lemme 12.4. Montrons dans un premier temps que $Q_{A^{-1}} = Q_A^*$ où Q_A^* désigne la transformée de Legendre de Q_A , puis que $Q_A^* = \tilde{Q}$.

Considérons $A_{1/2}$ la matrice symétrique définie positive telle que $A_{1/2}^2 = A$, on a alors

$$Q_A(y) = \frac{1}{2}|A_{1/2}(y)|^2.$$

Soit Q_A^* la transformée de Legendre de Q_A . Puisque les matrices A^{-1} et $A_{1/2}^{-1}$ sont symétriques, pour tout $x \in \mathbb{R}^n$

$$\begin{aligned} Q_A^*(x) &= \sup_{y \in \mathbb{R}^n} \left\{ x \cdot y - \frac{1}{2}|A_{1/2}(y)|^2 \right\} \\ &= \sup_{z \in \mathbb{R}^n} \left\{ x \cdot A_{1/2}^{-1}(z) - \frac{1}{2}|z|^2 \right\} \\ &= \sup_{z \in \mathbb{R}^n} \left\{ A_{1/2}^{-1}(x) \cdot z - \frac{1}{2}|z|^2 \right\} \\ &= \frac{1}{2}|A_{1/2}^{-1}(x)|^2 \\ &= Q_{A^{-1}}(x). \end{aligned}$$

Montrons que la fonction \tilde{Q} est convexe semi-continue inférieurement. La convexité est une conséquence de la convexité de la fonction $(x_1, \dots, x_m) \rightarrow \sum_i c_i A_i^{-1}(x_i) \cdot x_i$ et de la linéarité de la condition $x = \sum c_i B_i^t(x_i)$. Soit $x \in \mathbb{R}^n$ et $y_x = A^{-1}x$. Montrons que \tilde{Q} est semi-continue inférieurement. En choisissant $x_i = A_i B_i(y_x)$, puisque $\sum c_i B_i^t(x_i) = A y_x = x$, on obtient

$$\tilde{Q}(x) \leq \sum_i c_i B_i(y_x) \cdot A_i B_i(y_x) = Q_A(y_x) = Q_A(A^{-1}x).$$

Soit K_x et K_x^1 les compacts définis par

$$K_x = \left\{ (x_1, \dots, x_m) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_m}, x = \sum c_i B_i^t(x_i), \sum_i c_i A_i^{-1}(x_i) \cdot x_i \leq Q_A(A^{-1}x) \right\},$$

et

$$K_x^1 = \left\{ (x_1, \dots, x_m) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_m}, \left| x - \sum c_i B_i^t(x_i) \right| \leq 1, \sum_i c_i A_i^{-1}(x_i) \cdot x_i \leq Q_A(A^{-1}x) + 1 \right\},$$

On a

$$\tilde{Q}(x) = \inf_{(x_1, \dots, x_m) \in K_x} \sum_i c_i A_i^{-1}(x_i) \cdot x_i = \inf_{(x_1, \dots, x_m) \in K_x^1} \sum_i c_i A_i^{-1}(x_i) \cdot x_i.$$

Soit $(x^{(n)})$ une suite de point de \mathbb{R}^n qui tend vers x . Quitte à extraire une sous-suite de $(x^{(n)})$, on peut supposer de plus que

$$\liminf_n \tilde{Q}(x^{(n)}) = \lim_n Q(x^{(n)}).$$

Soit $(x_1^{(n)}, \dots, x_m^{(n)}) \in K_{x^{(n)}}$ tel que

$$\tilde{Q}(x^{(n)}) = \sum_i c_i A_i^{-1}(x_i^{(n)}) \cdot x_i^{(n)}. \quad (12.5)$$

Puisque $x^{(n)} \rightarrow x$, par continuité on obtient que $K_{x^{(n)}} \subset K_x^1$ pour tout n suffisamment grand, $n \geq n_0$. La suite $(x_1^{(n)}, \dots, x_m^{(n)})_{n \geq n_0}$ est une suite du compact K_x^1 donc elle admet une sous suite convergente $(x_1^{(n_k)}, \dots, x_m^{(n_k)})_k$ vers un point (x_1, \dots, x_m) . Par continuité, par passage à la limite lorsque $k \rightarrow \infty$ dans l'égalité (12.5) et l'égalité

$$x^{(n_k)} = \sum c_i B_i^t(x_i^{(n_k)}),$$

on obtient

$$x = \sum c_i B_i^t(x_i),$$

et

$$\liminf_n \tilde{Q}(x^{(n)}) = \lim_k \tilde{Q}(x^{(n_k)}) = \sum_i c_i A_i^{-1}(x_i) \cdot x_i \geq \tilde{Q}(x).$$

La fonction \tilde{Q} est donc semi-continue inférieurement en tout point $x \in \mathbb{R}^N$.

La transformée de Legendre de \tilde{Q} est

$$\begin{aligned} \tilde{Q}^*(y) &= \sup_{x \in \mathbb{R}^n} \left\{ y \cdot x - \tilde{Q}(x) \right\} \\ &= \sup_{x \in \mathbb{R}^n} \sup_{x_1, \dots, x_m / x = \sum_i c_i B_i^t(x_i)} \left\{ y \cdot x - \frac{1}{2} \sum_i c_i A_i^{-1}(x_i) \cdot x_i \right\} \\ &= \sup_{x_1, \dots, x_m} \left\{ y \cdot \left(\sum_i c_i B_i^t(x_i) \right) - \frac{1}{2} \sum_i c_i A_i^{-1}(x_i) \cdot x_i \right\} \\ &= \sum_i c_i \sup_{x_i} \left\{ y \cdot B_i^t(x_i) - \frac{1}{2} |(A_i)_{1/2}^{-1}(x_i)|^2 \right\} \\ &= \sum_i c_i \sup_{z_i} \left\{ y \cdot B_i^t(A_i)_{1/2}(z_i) - \frac{1}{2} |z_i|^2 \right\} \\ &= \sum_i c_i \sup_{z_i} \left\{ (A_i)_{1/2} B_i(y) \cdot z_i - \frac{1}{2} |z_i|^2 \right\} \\ &= \frac{1}{2} \sum_i c_i |(A_i)_{1/2} B_i(y)|^2 \\ &= \frac{1}{2} \sum_i c_i B_i^t A_i B_i(y) \cdot y \\ &= \frac{1}{2} A(y) \cdot y \\ &= Q_A(y). \end{aligned}$$

Ainsi, puisque \tilde{Q} est convexe et continue inférieurement, on obtient

$$\tilde{Q} = (\tilde{Q}^*)^* = Q_A^* = Q_{A^{-1}}.$$

Bibliographie

- [AG] Luigi Ambrosio and Nicola Gigli, *A user's guide to optimal transport*.
- [BB99] Jean-David Benamou and Yann Brenier, *A numerical method for the optimal time-continuous mass transport problem and related problems*, Monge Ampère equation : applications to geometry and optimization (Deerfield Beach, FL, 1997), Contemp. Math., vol. 226, Amer. Math. Soc., Providence, RI, 1999, pp. 1–11. MR 1660739 (99j :65151)
- [BDM09] Rainer Burkard, Mauro Dell'Amico, and Silvano Martello, *Assignment problems*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2009. MR 2488749 (2010c :90002)
- [Ber01] Dimitri Bertsekas, *Auction algorithms*, pp. Vol. I : xxviii+539 pp. ; Vol. II : viii+549 pp. ; Vol. III : viii+580 pp. ; Vol. IV : viii+503 pp. ; Vol. V : viii+557 pp. ; Vol. VI : xxxii+258, Kluwer Academic Publishers, Dordrecht, 2001. MR 1865755 (2002h :00006)
- [Bil99] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics : Probability and Statistics, John Wiley & Sons, Inc., New York, 1999, A Wiley-Interscience Publication. MR 1700749 (2000e :60008)
- [Bro82] Lawrence D. Brown, *A proof of the central limit theorem motivated by the Cramér-Rao inequality*, Statistics and probability : essays in honor of C. R. Rao, North-Holland, Amsterdam-New York, 1982, pp. 141–148. MR 659464 (83h :60021)
- [CM89] Juan Antonio Cuesta and Carlos Matrán, *Notes on the Wasserstein metric in Hilbert spaces*, Ann. Probab. **17** (1989), no. 3, 1264–1276. MR 1009457 (90k :60029)
- [JS05] Oliver Johnson and Richard Samworth, *Central limit theorem and convergence to stable laws in Mallows distance*, Bernoulli **11** (2005), no. 5, 829–845. MR 2172843 (2006k :60037)
- [McC97] Robert J. McCann, *A convexity principle for interacting gases*, Advances in Mathematics **128** (1997), no. 1, 153 – 179.
- [MG07] Jiri Matousek and Bernd Gärtner, *Understanding and using linear programming*, Springer Science & Business Media, 2007.
- [MO14] Quentin Mérigot and Édouard Oudet, *Handling convexity-like constraints in variational problems*, SIAM J. Numer. Anal. **52** (2014), no. 5, 2466–2487. MR 3268615
- [PPO14] Nicolas Papadakis, Gabriel Peyré, and Edouard Oudet, *Optimal transport with proximal splitting*, SIAM J. Imaging Sci. **7** (2014), no. 1, 212–238. MR 3158785
- [RPDB11] J. Rabin, G. Peyré, J. Delon, and M. Berton, *Wassertein barycenter and its applications to texture mixing*, LNCS, Proc. SSVM'11, vol. 6667, Springer, 2011, pp. 435–446.
- [Rud87] Walter Rudin, *Real and complex analysis, 3rd ed.*, McGraw-Hill, Inc., New York, NY, USA, 1987.
- [Tan73] Hiroshi Tanaka, *An inequality for a functional of probability distributions and its application to Kac's one-dimensional model of a Maxwellian gas*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **27** (1973), 47–52. MR 0362442 (50 #14883)

Voici quelques résultats relatifs à l'application de transport $\nabla\phi_t$ et à la mesure μ_t .

Lemme 12.5. *Soit μ une mesure absolument continue par rapport à la mesure de Lebesgue. Soit $\phi : \mathbb{R}^n \rightarrow \mathbb{R} \cup +\infty$ une application convexe et Ω_{dif} l'ensemble des points où $\nabla\phi$ est différentiable. Pour tout $x \in \mathbb{R}^n$, on considère l'application convexe,*

$$\phi_t(x) = (1-t)\frac{|x|^2}{2} + t\phi(x).$$

1. L'application $\nabla\phi_t : \Omega_{dif} \rightarrow \mathbb{R}^n$ est injective pour tout $t \in [0, 1[$.
2. L'application $(\nabla\phi_t)^{-1} : \nabla\phi_t(\Omega_{dif}) \rightarrow \Omega_{dif}$ est $(1-t)^{-1}$ -lipschitzienne pour tout $t \in [0, 1[$.
3. Pour tout $t \in [0, 1[$, la mesure $\mu_t = \nabla\phi_t\mu$ est absolument continue par rapport à la mesure de Lebesgue, $\mu_t(\mathbb{R}^n) = \mu(\mathbb{R}^n)$.

La démonstration de ce lemme est donnée à la fin de cette partie. On déduit de ce lemme que $(\mu_t)_{t \in [0, 1[}$ est un chemin de probabilité absolument continue par rapport à la mesure de Lebesgue. Ce chemin interpole entre $\mu_0 = \mu$ et $\mu_1 = \nu$. Soit g_t la densité μ_t par rapport à la mesure de Lebesgue.

Démonstration du Lemme 12.5. Puisque ϕ est convexe et

$$\phi_t(x) = t\phi(x) + (1-t)\frac{|x|^2}{2}, \quad x \in \mathbb{R}^n,$$

l'application ϕ_t est strictement convexe pour $t \in [0, 1)$. Ceci implique que l'application $\nabla\phi_t : \mathcal{D} \rightarrow \mathbb{R}^n$ est injective sur son domaine de définition \mathcal{D} . En effet, supposons que $\nabla\phi_t(x) = \nabla\phi_t(y)$. Par convexité, on a pour tout $t \in (0, 1)$

$$\nabla\phi_t(y) \cdot (x - y) \leq \frac{\phi_t(tx + (1-t)y) - \phi_t(y)}{t} \leq \frac{\phi_t(x) - \phi_t(tx + (1-t)y)}{1-t} \leq \nabla\phi_t(x) \cdot (x - y).$$

Par suite pour tout $t \in (0, 1)$, $\phi_t(tx + (1-t)y) = \phi_t(x) = \phi_t(y)$. Puisque ϕ_t est strictement convexe, nécessairement $x = y$.

On peut donc considérer l'application inverse $\nabla\phi_t^{-1}$ définie sur $\mathcal{D}' = \nabla\phi_t(\mathcal{D})$. Montrons que cette application est $(1-t)^{-1}$ -lipschitzienne. D'après l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned} |x - y| |\nabla\phi_t(x) - \nabla\phi_t(y)| &\geq (x - y) \cdot (\nabla\phi_t(x) - \nabla\phi_t(y)) \\ &= (1-t)|x - y|^2 + t(x - y) \cdot (\nabla\phi(x) - \nabla\phi(y)). \end{aligned}$$

Puisque ϕ est une fonction convexe pour tout x, y tels que $\nabla\phi(x)$ et $\nabla\phi(y)$ existe, on a

$$\nabla\phi(y) \cdot (x - y) \leq \phi(x) - \phi(y) \leq \nabla\phi(x) \cdot (x - y),$$

et donc

$$(x - y) \cdot (\nabla\phi(x) - \nabla\phi(y)) \geq 0.$$

On obtient ainsi pour tous $x, y \in \mathcal{D}$,

$$|\nabla\phi_t(x) - \nabla\phi_t(y)| \geq (1-t)|x - y|,$$

ce qui montre que $\nabla\phi_t^{-1}$ est $(1-t)^{-1}$ -lipschitzienne sur \mathcal{D}' .

Les résultats de théorie de la mesure de Lebesgue indiquent alors que pour tout sous-ensemble mesurable M de \mathbb{R}^n ,

$$|\nabla\phi_t^{-1}(M)| = |\nabla\phi_t^{-1}(M \cap \mathcal{D}')| \leq (1-t)^{-n} |M \cap \mathcal{D}'| \leq (1-t)^{-n} |M|.$$

(Ce résultat est simple à montrer si M est une boule.)

Pour montrer que μ_t est absolument continue par rapport à la mesure de Lebesgue, il suffit de montrer que μ_t ne charge pas les ensembles M de mesure de Lebesgue nul. Si $|M| = 0$ alors l'inégalité précédente donne $|\nabla\phi_t^{-1}(M)| = 0$. Par définition de la mesure μ_t , on obtient alors

$$\mu_t(M) = \mu_0(\nabla\phi_t^{-1}(M)) = 0,$$

car μ_0 est absolument continue par rapport à la mesure de Lebesgue. □