

Autour du Transport Optimal et de ses Applications
en Apprentissage Supervisé

Jean ALAUX; Ulysse MARTEAU-FEREY

Introduction

La question du transport optimal est un problème d'optimisation sous contrainte qui a été soulevé pour la première fois par MONGE en 1781. La théorie a connu de nombreux développements au cours du 20ième siècle. Aujourd'hui, on lui découvre des applications nouvelles en informatique. Nous allons étudier en particulier son apport en traitement de l'image et en apprentissage supervisé.

La première partie introduit le cadre théorique du transport optimal. Dans les deux parties qui suivent, on développe les outils de calcul qui lui sont liés, à savoir la programmation linéaire et un algorithme de régularisation entropique.

On présente comment ces outils permettent de calculer le barycentre d'une famille d'images selon une métrique associée au transport optimal, pertinente pour certains problèmes informatiques.

Enfin on s'intéresse à une application en apprentissage supervisé de cette théorie. Si on se donne un graphe et sur certains de ces sommets une distribution de probabilité, la théorie du transport optimal permet d'affecter aux sommets vides une mesure de probabilité. Ce procédé d'affectation particulier permet de préserver une certaine cohérence que l'on expliquera.

Conventions et notations

- Si X est un espace munie d'une topologie connue, $\mathcal{C}_b^0(X)$ désigne l'espace des fonctions continues bornées sur X
- Si X et Y sont deux ensembles et si a et b sont des applications à valeurs réelles définies respectivement sur X et Y , alors pour $(x, y) \in X \times Y$, $a \oplus b(x, y) = a(x) + b(y)$.
- Si X est un espace mesurables, $\mathcal{M}_+(X)$ désigne l'espace des mesures positives sur X .
- Si X et Y sont deux espaces mesurés et si μ, ν sont des mesures définies respectivement sur X et Y , $\Pi(\mu, \nu)$ désigne l'ensemble des mesures sur $X \times Y$ dont les mesures marginales sont respectivement μ et ν .
- Si n, m sont des entiers non nuls, $M_{n,m}(\Omega)$ désigne l'espace vectoriel des matrices à n lignes et m colonnes et à coefficients dans Ω .
- Pour M une matrice, on note M^t sa transposée.
- On note \mathbb{R}_{++}^n ensemble des vecteurs de \mathbb{R}^n à coefficients strictement positifs.
- On note $M_{n,m}(\mathbb{R}_+)$ et $M_{n,m}(\mathbb{R}_{++})$ ensemble des matrices à coefficients respectivement positifs strictement positifs.
- On note $D_n(\mathbb{R})$ ensemble des matrices diagonales de taille n à valeurs dans \mathbb{R} (idem pour \mathbb{R}_{++}).
- Si $u \in \mathbb{R}^n$, $\text{diag}(u)$ est la matrice diagonale de taille n où $\text{diag}(u)_{ii} = u_i$ pour tout $i \in \llbracket 1, n \rrbracket$.
- $e = (1, \dots, 1)^t$.
- Si E est un espace vectoriel, on note E^* son dual et $\langle \cdot | \cdot \rangle$ le produit de dualité, que l'on confond avec le produit scalaire canonique dans les espaces vectoriels de dimension finie.
- $\|x\|$ désigne la norme infinie d'un vecteur x .

Table des matières

1	Outils de Transport Optimal	5
1.1	Problèmes de MONGE et KANTOROVITCH	5
1.2	Formulation Duale du problème de KANTOROVITCH	6
1.3	Démonstration du théorème de dualité	8
1.4	Cas particulier du transport optimal discret, introduction de problèmes linéaires	11
2	Programmation linéaire	13
2.1	Programmes linéaires	13
2.2	L'algorithme du simplexe	15
3	Application au traitement de l'image : barycentres de Wasserstein par descente de gradient	17
3.1	Potentiels de Kantorovitch comme sous gradients	17
3.2	Différentiabilité en le support des images	19
3.3	Un exemple simple	22
4	Calcul par régularisation entropique	25
4.1	Régularisée entropique et approximation de la solution	25
4.1.1	Définitions, formulation des problèmes primal et dual	26
4.1.2	Convergence des solutions du problème régularisé vers les solution du problème initial	29
4.2	Calcul effectif des solutions au problèmes primal et dual pour la régularisée entropique	29
4.2.1	Reformulation du problème	29
4.2.2	Résolution : l'algorithme de SINKHORN	30
5	Application à la diffusion de mesures de probabilité sur les sommets d'un graphe	38
5.1	Monotonicité c-cyclique	38
5.2	Idée de la preuve de l'expression de la distance de Wasserstein avec les fonctions de répartition	39
5.3	Éléments de calcul différentiel sur les graphes	41
5.3.1	Premières définitions sur un graphe	41

5.3.2	Calcul différentiel sur un sous-graphe	43
5.3.3	Formules	46
5.3.4	Laplacien, propriétés des fonctions harmoniques	48
5.4	Harmonisation de distributions de probabilités dans un graphe	54
5.5	Diffusion sur les graphes	59

Chapitre 1

Outils de Transport Optimal

L'objet de cette section est d'introduire le cadre théorique pour l'étude menée par la suite. On y présente le problème du transport optimal sous ses deux formes historiques : le problème de MONGE et le problème de KANTOROVITCH, et on démontre l'existence de solutions à ces problèmes. On présente la distance de Wasserstein sur les espaces de mesures finies. On démontre enfin un résultat important de dualité dans le cadre du transport optimal, qui généralise la dualité utilisée en programmation linéaire.

1.1 Problèmes de MONGE et KANTOROVITCH

On considère dans cette section deux espaces probabilisés (X, μ) et (Y, ν) . On se donne également une fonction $c : X \times Y \rightarrow \mathbb{R}$ appelée fonction de coût. Physiquement, le problème du transport optimal consiste à déterminer comment transporter de la "masse" située sur X suivant la répartition μ vers l'espace Y , de telles sortes qu'après le transport, la masse ait la répartition ν , et ceci avec un coût de transport minimal. Le coût de transport d'une unité de masse située en $x \in X$ vers $y \in Y$ est donné par $c(x, y)$. Le problème a été posé historiquement de deux manières différentes.

Dans un premier temps, on cherche une relation déterministe entre la répartition initiale de la masse et sa répartition finale. On entend donc envoyer toute la masse située en $x \in X$ en un unique $y \in Y$. Le problème ainsi posé est le suivant :

Problème de MONGE : Déterminer T réalisant :

$$\begin{aligned} & \inf_{\substack{T : X \rightarrow Y \\ \mu \circ T^{-1} = \nu}} \int_X c(x, T(x)) \mu(dx) \end{aligned}$$

Les T vérifiant $\mu \circ T^{-1} = \nu$ sont appelés *plans de transport*. Néanmoins, le problème ainsi posé n'admet pas, a priori, de solution. En effet, il faut s'assurer qu'il existe des plans de transport, ce qui est faux en général : il suffit de μ soit une masse de Dirac

et ν une mesure atomique dont le support est de cardinal au moins 2 pour que de tels plans de transports n'existent pas. Ceci conduit à une autre formulation, distincte, du problème du transport optimal :

Problème de KANTOROVITCH : Déterminer γ une mesure sur $X \times Y$ réalisant :

$$\inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} cd\gamma$$

On dit alors que les γ sont des plans de transport. La différence avec le problème de MONGE est la suivante : les plans de transports considérés ne sont plus ici nécessairement déterministes, il se peut que la moitié de la masse située en un $x \in X$ soit envoyée sur $y_1 \in Y$ et l'autre moitié vers $y_2 \in Y$. En outre, il existe naturellement des plans de transport puisque $\mu \otimes \nu$ en est un. Dans la suite, on travaillera donc de préférence dans le cadre posé par le problème de KANTOROVITCH.

Existence de solutions au problème de KANTOROVITCH

Théorème 1. *Soit X et Y deux espaces métriques compacts. Soit μ et ν des mesures de probabilité respectivement sur $(X, \mathcal{B}(X))$ et $(Y, \mathcal{B}(Y))$. Soit $c : X \times Y \rightarrow \mathbb{R}$ continue. Alors le problème de KANTOROVITCH admet des solutions.*

Démonstration. On a déjà vu qu'il existe des plans de transport. Soit $(\gamma_n)_{n \in \mathbb{N}} \in \Pi(\mu, \nu)^{\mathbb{N}}$ telle que $\int_{X \times Y} cd\gamma_n$ tende vers l'infimum. $E = (C^0(X \times Y, \mathbb{R}); \|\cdot\|_{\infty})$ est un espace de Banach, donc d'après le théorème de Banach-Alaoglu, la boule unité de son dual est faible-* compacte. Or, l'intégration contre une mesure de probabilité γ_n est une forme linéaire continue dans cette boule unité. Il existe donc, après extraction, ϕ une forme linéaire sur E telle que pour toute fonction $f \in E$:

$$\int f d\gamma_n \rightarrow \phi(f)$$

Or, par le théorème de représentation de Riesz, on peut écrire ϕ comme l'intégration contre une certaine mesure γ sur $X \times Y$. On a donc $\gamma_n \rightarrow \gamma$ étroitement. Reste à vérifier que les mesures marginales de γ sont bien μ et ν , ce qui est évident en intégrant une fonction continue sur $X \times Y$ ne dépendant pas de la variable $y \in Y$ puis en passant à la limite, et en procédant de même pour une fonction de dépendant pas de la variable $x \in X$. \square

1.2 Formulation Duale du problème de KANTOROVITCH

Les applications que nous verrons dans la suite de la théorie du transport optimal concernent le transport entre mesures atomiques. Dans ce cadre les résultats de dualité s'énoncent naturellement en terme de Lagrangien (Cf. cours de modélisation de

B. Maury [6]). On se place ici dans la situation suivante : on se donne deux espaces compacts X et Y , ainsi que des mesures μ et ν respectivement sur X et Y . Nous développons ici les résultats de dualité, que nous utiliserons dans le cadre restreint des mesures atomiques à support fini. Soit γ une mesure positive sur $X \times Y$. Alors :

$$\sup_{(a,b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y)} \left[\int_X a \, d\mu + \int_Y b \, d\nu - \int_{X \times Y} a \oplus b \, d\gamma \right] = \begin{cases} 0 & \text{si } \gamma \in \Pi(\mu, \nu) \\ +\infty & \text{sinon} \end{cases}$$

En effet, si $\gamma \in \Pi(\mu, \nu)$, alors en intégrant sur chaque tranche pour un couple de fonctions mesurables (a,b) quelconque, on a l'égalité voulue. Autrement, on peut supposer, quitte à considérer les cas symétrique, qu'il existe $A \subset X$ mesurables tel que $\gamma(A \times Y) \geq \mu(A)$. On considère maintenant une fonction $a = \lambda \mathbf{1}_A$ et $b = 0$. On peut approcher cette fonction, en norme infinie, par une fonction continue. En faisant tendre $\lambda \rightarrow +\infty$ et $\epsilon \rightarrow 0$, on obtient l'égalité dans le second cas.

On peut donc écrire :

$$\begin{aligned} \min_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c \, d\gamma &= \inf_{\gamma \in \mathcal{M}_+(X \times Y)} \int_{X \times Y} c \, d\gamma + \sup_{(a,b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y)} \left[\int_X a \, d\mu + \int_Y b \, d\nu - \int_{X \times Y} a \oplus b \, d\gamma \right] \\ &= \inf_{\gamma \in \mathcal{M}_+(X \times Y)} \sup_{(a,b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y)} \left[\int_X a \, d\mu + \int_Y b \, d\nu + \int_{X \times Y} (c - a \oplus b) \, d\gamma \right] \end{aligned}$$

A ce stade du calcul, on souhaite justifier qu'il est licite d'inverser les opérateurs sup et inf dans l'expression ci-dessus. En effet, le cas échant, on aurait :

$$\begin{aligned} \min_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c \, d\gamma &= \sup_{(a,b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y)} \inf_{\gamma \in \mathcal{M}_+(X \times Y)} \left[\int_X a \, d\mu + \int_Y b \, d\nu + \int_{X \times Y} (c - a \oplus b) \, d\gamma \right] \\ &= \sup_{(a,b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y)} \left[\int_X a \, d\mu + \int_Y b \, d\nu + \inf_{\gamma \in \mathcal{M}_+(X \times Y)} \int_{X \times Y} (c - a \oplus b) \, d\gamma \right] \end{aligned}$$

Or si l'on impose désormais $\gamma \in \mathcal{M}_+(X \times Y)$ au lieu de $\gamma \in \Pi(\mu, \nu)$, alors, pour a et b fixées :

$$\inf_{\gamma \in \mathcal{M}_+(X \times Y)} \int_{X \times Y} (c - a \oplus b) \, d\gamma = \begin{cases} 0 & \text{si } a \oplus b \leq c \\ -\infty & \text{sinon} \end{cases}$$

Théorème 2 (dualité de KANTOROVITCH). *Sous les hypothèses du théorème précédent, et si on suppose en outre que $c \geq 0$ et qu'il existe $(a, b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y)$ tel que $c \geq a \oplus b$. La recherche de*

$$\sup_{\{(a,b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y) \mid a \oplus b \leq c\}} \int_X a \, d\mu + \int_Y b \, d\nu$$

est appelée appelée problème dual de KANTOROVITCH. On a :

$$\min_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c \, d\gamma = \sup_{\{(a,b) \in \mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y) \mid a \oplus b \leq c\}} \int_X a \, d\mu + \int_Y b \, d\nu$$

Remarque 1. *On peut démontrer que sous les hypothèses ci-dessus le problème dual admet un couple de fonction maximisantes, appelées potentiels de KANTOROVITCH. Nous verrons dans la dernière partie une esquisse de la preuve.*

1.3 Démonstration du théorème de dualité

Dans le calcul précédent, nous avons vu que le point essentiel de la preuve est la justification d'une inversion borne supérieure/ borne inférieure, autrement dit la possibilité d'appliquer un théorème du type minimax. Replaçons ce problème dans le cadre de l'optimisation convexe et posons :

$$\mathcal{L} : \begin{cases} \mathcal{M}_+(X \times Y) \times (\mathcal{C}_b^0(X) \times \mathcal{C}_b^0(Y)) \longrightarrow \mathbb{R} \\ (\gamma, (a, b)) \longmapsto \int_X a \, d\mu + \int_Y b \, d\nu + \int_{X \times Y} (c - a \oplus b) \, d\gamma \end{cases}$$

On reconnaît en \mathcal{L} un Lagrangien. On observe également que \mathcal{L} est affine en chacune de ses variables, donc a fortiori convexe par rapport à l'une et concave par rapport à l'autre. Le théorème de SION énonce que sous une hypothèse de compacité de l'un des deux espaces de départ et de régularité de \mathcal{L} , une telle inversion d'opérateurs borne supérieure/borne inférieure est possible, autrement dit que le Lagrangien \mathcal{L} dispose bien d'un point-selle. La non-compacité des espaces de départ pour \mathcal{L} rend impossible une application du théorème de SION et il semble impossible d'adapter les calculs qui précèdent, par exemple en restreignant les espaces de départ pour les rendre compacts, tout en conservant l'équivalence avec le problème dual. Nous allons donc proposer une preuve différente de l'inversion borne supérieure/borne inférieure.

La transformée de LEGENDRE-FENCHEL fournit l'existence d'un point selle pour les Lagrangiens convexes/concaves. C'est l'approche proposée dans [2].

Définition 1 (Transformée de LEGENDRE-FENCHEL). *Soit E un espace vectoriel normé, et f une fonctionnelle convexe sur E à valeurs dans $\mathbb{R} \cup \{+\infty\}$. La transformée de LEGENDRE-FENCHEL f^* de f est définie sur E^* par :*

$$f^*(u) = \sup_{x \in E} [\langle u|x \rangle - f(x)]$$

On vérifie facilement que f^ est convexe.*

Théorème 3 (Dualité de FENCHEL-ROCKAFELLAR). *Soit E un espace vectoriel normé, et f et g deux fonctionnelles convexes sur E à valeurs dans $\mathbb{R} \cup \{+\infty\}$. On suppose qu'il existe $x_0 \in E$ tel que :*

$$\begin{cases} f(x_0) < +\infty \\ g(x_0) < +\infty \\ f \text{ est continue en } x_0 \end{cases}$$

Alors :

$$\inf_{x \in E} [f(x) + g(x)] = \max_{u \in E^*} [-f^*(-u) - g^*(u)]$$

Démonstration. On reformule l'égalité à démontrer en :

$$m = \inf_{x \in E} [f(x) + g(x)] = \sup_{u \in E^*} \inf_{x, y \in E} [\langle u | x - y \rangle + f(x) + g(y)]$$

Pour un $u \in E^*$ fixé, on a naturellement

$$\inf_{x, y \in E} [\langle u | x - y \rangle + f(x) + g(y)] \leq m$$

en évaluant en un $x = y$ quelconque. On a donc d'office l'une des deux inégalités.

Définissons :

$$\begin{aligned} C(f) &= \{(x, \lambda) \in E \times \mathbb{R} | \lambda > f(x)\} \\ C(g) &= \{(x, \lambda) \in E \times \mathbb{R} | \lambda \leq m - g(x)\} \end{aligned}$$

Par convexité de f et g , il est immédiat que $\overset{\circ}{C}(f)$ et $C(g)$ sont également convexes, et non vide puisque f et g ne sont pas égales à $+\infty$. Par définition de m , $\overset{\circ}{C}(f)$ et $C(g)$ sont disjoints. En outre, puisque f est finie et continue en x_0 , elle est finie au voisinage de x_0 , donc $(x_0, f(x_0) + 1) \in \overset{\circ}{C}(f)$. D'après le théorème de Hahn-Banach, on dispose d'une forme linéaire non triviale sur $E \times \mathbb{R}$ telle que pour $z_f \in \overset{\circ}{C}(f)$ et tout $z_g \in C(g)$, $l(z_f) \geq l(z_g)$.

Ecrivons $l : (x, \lambda) \in E \times \mathbb{R} \mapsto \langle u | x \rangle + \alpha \lambda$ pour un $u \in E^*$ et un $\alpha \in \mathbb{R}$ tels que u et α ne sont pas simultanément nuls.

Vérifions que $\alpha > 0$. Par l'absurde, si $\alpha = 0$, alors, puisque f est continue en x_0 , il existe un voisinage U de x_0 et un $A > 0$ tel que $U \times [A; +\infty[\subset C(f)$. Soit $x \in U$, alors $\langle u | x - x_0 \rangle \geq 0$ puisque $g(x_0) < +\infty$. Or U est un voisinage de x_0 donc $u = 0$ et $l = 0$: contradiction. Donc $\alpha \neq 0$. En évaluant enfin en x_0 et en faisant tendre le paramètre $\lambda \rightarrow +\infty$ avec $\tilde{\lambda}$ fixé, on obtient enfin $\alpha > 0$.

Posons maintenant $v = \frac{1}{\alpha}u$. Pour $(x_f, \lambda) \in C(f)$ et $(x_g, \tilde{\lambda}) \in C(g)$:

$$\langle v | x_f \rangle + \lambda \geq \langle v | x_g \rangle + \tilde{\lambda}$$

En passant à la borne supérieure en $\tilde{\lambda}$ et inférieure en λ , il vient :

$$\langle v | x_f \rangle + f(x_f) \geq \langle v | x_g \rangle + m - g(x_g)$$

Et cette dernière inégalité demeure si $f(x_f)$ ou $g(x_g)$ n'est pas fini. Ceci permet de conclure directement quant à la seconde inégalité. \square

Démonstration du théorème de dualité de Kantorovitch : Soit donc X, Y des espaces compacts mesurables, et $c \in \mathcal{C}^0(X \times Y)$ positive. On pose $E = \mathcal{C}_b(X \times Y)$. D'après le théorème de Riesz, E^* n'est autre que l'espace des mesures de Radon sur $X \times Y$. On définit :

$$f : u \in E \mapsto \begin{cases} 0 & \text{si } u \geq -c \\ +\infty & \text{sinon} \end{cases}$$

$$g : u \in E \mapsto \begin{cases} \int_X a d\mu + \int_Y b d\nu & \text{si } u = a \oplus b \\ +\infty & \text{sinon} \end{cases}$$

Avec $x_0 = \mathbf{1}$, $f(x_0) < +\infty$, $g(x_0) < +\infty$ et f est évidemment continue en x_0 puisque c est positive par hypothèse.

On commence par calculer f^* et g^* . Soit $\gamma \in E^*$.

$$\begin{aligned} f^*(-\gamma) &= \sup_{\Phi \in E} \left[\int_{X \times Y} -\Phi d\gamma - f(\Phi) \right] \\ &= \sup_{\Phi \in E} \left[\int_{X \times Y} -\Phi d\gamma - \begin{cases} 0 & \text{si } \Phi \geq -c \\ +\infty & \text{sinon} \end{cases} \right] \\ &= \sup_{\{\Phi \in E \mid \Phi \geq -c\}} - \int_{X \times Y} \Phi d\gamma \\ &= \sup_{\{\Phi \in E \mid \Phi \leq c\}} \int_{X \times Y} \Phi d\gamma \\ &= \begin{cases} \int_{X \times Y} c d\gamma & \text{si } \gamma \in \mathcal{M}_+(X \times Y) \\ +\infty & \text{sinon} \end{cases} \end{aligned}$$

De même,

$$g^*(\gamma) = \begin{cases} 0 & \text{si } \forall a, b \int_{X \times Y} a \oplus b d\gamma = \int_X a d\mu + \int_Y b d\nu \\ +\infty & \text{sinon} \end{cases}$$

Il vient :

$$\begin{aligned} \sup_{\gamma \in E^*} [-f^*(-\gamma) - g^*(\gamma)] &= \sup_{\gamma \in E^*} \left[- \begin{cases} \int_{X \times Y} c d\gamma & \text{si } \gamma \in \mathcal{M}_+(X \times Y) \\ +\infty & \text{sinon} \end{cases} - \begin{cases} 0 & \text{si } \gamma \in \Pi(\mu, \nu) \\ +\infty & \text{sinon} \end{cases} \right] \\ &= \sup_{\gamma \in \Pi(\mu, \nu)} - \int_{X \times Y} c d\gamma \\ &= \inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c d\gamma \end{aligned}$$

En outre,

$$\begin{aligned}
\inf_{\Phi \in E} f(\Phi) + g(\Phi) &= \inf_{\Phi \in E} \left[\begin{array}{l} 0 \text{ si } \Phi \geq -c \\ + \infty \text{ sinon} \end{array} + \begin{array}{l} \int_X ad\mu + \int_Y bd\nu \text{ si } \Phi = a \oplus b \\ + \infty \text{ sinon} \end{array} \right] \\
&= \inf_{a \oplus b \geq -c} \int_X ad\mu + \int_Y bd\nu \\
&= \sup_{a \oplus b \leq c} \int_X ad\mu + \int_Y bd\nu
\end{aligned}$$

D'après le théorèmes de dualité de FENCHEL-ROCKAFELLAR, on a la dualité de KANTOROVITCH :

$$\inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} cd\gamma = \sup_{a \oplus b \leq c} \int_X ad\mu + \int_Y bd\nu$$

Remarque : Il existe d'autres preuves du théorème de dualité de Kantorovitch, reposant sur le concept de monotonie c -cyclique, voir par exemple [1] ou [3].

1.4 Cas particulier du transport optimal discret, introduction de problèmes linéaires

Nous nous intéressons désormais au cas des mesures atomiques sur des espaces vectoriels de dimension finie. Nous introduisons dans cette partie les notations adoptées dans [7]. Dans toute la suite, on fixe un $d \in \mathbb{N}$ et considère deux mesures μ et ν atomiques sur \mathbb{R}^d de supports respectifs $X = \{x_1; \dots; x_n\}$ et $Y = \{y_1; \dots; y_m\}$, et de même masse finie. On note a_i les masses affectées aux x_i et b_j les masses affectées aux y_j .

On se donne également la fonction de coût suivante que nous utiliserons pour les applications informatiques : $c(x, y) = \|x - y\|$. Pour $p \geq 1$, la p -ième distance de Wasserstein sur l'espace des mesures atomiques sur \mathbb{R}^d est définie par :

$$W_p^p(\mu, \nu) = \left(\inf_{\gamma \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c^p d\gamma \right)^{\frac{1}{p}}$$

En pratique, on utilisera $p = 2$.

Pour $p \geq 1$ on introduit la matrice de coûts :

$$M = (\|x_i - y_j\|^p)_{i \in [1; n], j \in [1; m]}$$

Une mesure positive $\gamma \in \Pi(\mu, \nu)$ n'est autre que la donnée des masses $\gamma_{i,j}$ affectées aux points (x_i, y_j) . Autrement dit, on peut la représenter comme une matrice $\Gamma = (\gamma_{i,j}) \in M_{n,m}(\mathbb{R}_+)$, et les contraintes marginales deviennent :

$$\begin{cases} \Gamma \mathbf{1}_m = (a_1, \dots, a_n)^t \\ \Gamma^t \mathbf{1}_n = (b_1, \dots, b_m) \end{cases}$$

On pose $a = (a_1, \dots, a_n)$ et $b = (b_1, \dots, b_m)$. Le théorème de dualité de KANTOROVITCH s'écrit ici :

$$W_2^2(\mu, \nu) = \sup_{\{\alpha, \beta \mid \forall i, j \ a_i + b_j \leq M_{i,j}\}} \langle \alpha \mid a \rangle + \langle \beta \mid b \rangle$$

La partie qui suit présente l'algorithme du simplexe qui permet de calculer, pour a et b fixés, un couple de potentiels de KANTOROVITCH (α, β) . Nous verrons plus loin que ces potentiels, ici sous forme de vecteur, permettent de calculer des barycentres au sens de la distance de Wasserstein par une descente de gradient.

Chapitre 2

Programmation linéaire

Dans cette partie, on s'intéresse à aux programmes linéaires, qui sont des problèmes d'optimisation sous contraintes de formes linéaires. On les utilise comme outil pour la recherche de gradients de fonctions convexes, en l'occurrence des coûts de transports. On développe également l'algorithme du simplexe qui permet de résoudre ces problèmes en temps fini.

2.1 Programmes linéaires

On appelle programme linéaire un problème posé sous la forme suivante :
Déterminer

$$\begin{cases} \min C^t X \\ \text{avec } AX = B ; X \geq 0 \end{cases}$$

où C , X et B sont des vecteurs colonnes, et A une matrice rectangle. Les vecteurs X vérifiant $AX = B ; X \geq 0$ sont dits *admissibles*. On appelle aussi programme linéaire un problème qui se pose sous une forme similaire sans la condition de positivité des coefficients de X . En pratique, on se ramène toujours à la forme précédente, dite *standard* par ajout de variables artificielles et/ou par changement de variables. Nous ne détaillerons pas les astuces de calcul pléthoriques qui permettent de s'y ramener. Il convient cependant de noter que la technique employée pour se ramener à la forme standard et les conventions choisies peuvent être déterminantes pour l'efficacité de l'algorithme de résolution : celui-ci nécessite a priori le calcul de l'inverse d'une matrice carrée de taille le nombre de contraintes. Cette opération est notoirement couteuse en temps d'exécution. Bien choisir les variables employées peut permettre d'exploiter une particularité des données et d'éviter ou accélérer ce calcul d'inverse.

Définition 2. Soit un programme linéaire sous forme standard, avec $A \in M_{m,n}(\mathbb{R})$ où $m \leq n$. On suppose A de rang m , ie que les contraintes sont linéairement indépendantes. Soit \tilde{A} une matrice extraite de A inversible, contenant les colonnes i_1, \dots, i_m de A . On se donne X l'unique solution du système $AX = B$ vérifiant $X_i = 0$ si $\forall j, i_j \neq i$. On dit

que X est une solution basique du programme linéaire, associée à la base constituée des colonnes i_1, \dots, i_m de A .

Lemme 1. Soit $A_1, \dots, A_p \in \mathbb{R}^n$, et $X, B \in \mathbb{R}^n$ tels que $\sum_{i=1}^p x_i A_i = B$ et $X \geq 0$. Si les A_i ne sont pas linéairement indépendants, alors il existe $\hat{x}_1, \dots, \hat{x}_{p-1} \in \mathbb{R}_+$ et i_1, \dots, i_{p-1} tels que :

$$\sum_{j=1}^{p-1} \hat{x}_j A_{i_j} = B$$

Démonstration. : On considère la matrice dont les colonnes sont les A_i . Par hypothèse, on peut trouver $T = (t_1, \dots, t_p)^t \in \text{Ker}(A) \setminus \{0\}$. Quitte à multiplier par -1 , on dispose d'au moins un i tel que $t_i > 0$. Soit $\epsilon > 0$.

$$\sum_{i=1}^p (x_i - \epsilon t_i) A_i = B$$

On fait tendre $\epsilon \rightarrow 0$ positivement, jusqu'à ce que l'un des termes de la somme s'annule. On prend pour $i_1; \dots; i_{p-1}$ les indices des termes restants, complétés si nécessaire d'indices de termes s'étant annulés. Enfin on prend $\hat{x}_{i_j} = x_{i_j} - \epsilon t_{i_j}$. \square

Théorème 4. Soit un programme linéaire donné sous la forme précédente. On suppose $A \in M_{m,n}(\mathbb{R})$, avec $m \leq n$ et A de rang m . Alors :

1. S'il existe un vecteur admissible, il existe une solution basique du programme.
2. S'il existe un vecteur admissible réalisant le minimum voulu, alors il existe une solution basique réalisant ce minimum.

Démonstration. Soit $X = (x_1, \dots, x_n)^t$ un vecteur admissible. Notons A_1, \dots, A_n les colonnes de A . Alors :

$$\sum_{i=1}^n x_i A_i = B$$

D'après le lemme précédent, si x n'est pas une solution basique, on dispose de \hat{X} admissible vérifiant $X \geq 0$, et tel que $\#\text{supp}(\hat{X}) \leq n - 1$, en conservant le rang de la famille des colonnes de A qui ne sont pas multipliées par 0 dans la somme. On itère ce procédé jusqu'à obtenir m colonnes de A formant une famille libre, et le vecteur \hat{X} obtenu in fine correspond bien à une solution basique.

Pour le second point du théorème, on utilise exactement le même procédé : on part d'une solution optimale cette fois, puis on effectue une succession d'opérations qui permettent d'aboutir à une solution basique. Il suffit donc de montrer que cette succession d'opération laisse invariante la valeur de $C^t X$.

On reprend les conventions du Lemme 1, et on considère la fonction :

$$\Phi : \epsilon \longmapsto C^t(X - \epsilon T)$$

Où X est une solution optimale du programme linéaire.

$\Phi \geq 0$ au voisinage de 0 et atteint un minimum local en 0 par hypothèse. Or Φ est linéaire, donc Φ est constante. Il en résulte qu'à chaque itération dans la méthode itérative précédente, $C^t \hat{X} = C^t X$, ce qui achève la preuve. \square

2.2 L'algorithme du simplexe

Le principe de l'algorithme du simplexe est donc le suivant : sous l'hypothèse où le programme linéaire admet une solution optimale, on peut la chercher sous la forme d'une solution basique optimale. On commence par se donner une solution basique, et on écrit le système dans cette base de façon à avoir les m premières colonnes de A formant la matrice identité. On prend comme première solution basique celle donnée par les m premières colonnes de A , et on suppose qu'elle n'est pas optimale. On fait également l'hypothèse que A n'est pas *dégénérée*, ie que m colonnes distinctes de A sont toujours linéairement indépendantes. On note Y_1, \dots, Y_n les colonnes de A .

Etape 1 : On détermine quelle colonne doit intégrer la base de \mathbb{R}^m . Soit $X = (x_1, \dots, x_n)$ une solution de $AX = B$. Alors :

$$x_1 Y_1 + x_2 Y_2 + \dots + x_m Y_m = B - x_{m+1} Y_{m+1} - \dots - x_n Y_n$$

On effectue le produit scalaire à gauche par (c_1, \dots, c_m) . Il vient :

$$\sum_{i=1}^m x_i c_i = z_0 - \sum_{i=m+1}^n x_i z_i$$

En posant $z_0 = \sum_{i=1}^m c_i B_i$ et pour $i \geq m+1$: $z_i = \sum_{j=1}^m c_j Y_{j,i}$. Puis on ajoute $\sum_{i=m+1}^n x_i c_i$ de part et d'autre. Il vient :

$$C^t X = z_0 + \sum_{i=m+1}^n x_i (c_j - z_j)$$

z_0 est ici la valeur de $C^t X_0$ où X_0 est la solution basique initiale. On a supposé que cette dernière n'était pas optimale, ce qui permet d'affirmer qu'il existe au moins un $j \in m+1, \dots, n$ tel que $c_j - z_j < 0$. En effet, une solution basique autre a tous ses coefficients positifs. Un tel j peut être introduit dans la base afin de diminuer $C^t X$. En général, on choisit d'introduire la colonne j_0 pour laquelle $c_j - z_j$ est minimal.

Etape 2 : On doit déterminer quelle colonne doit quitter la base afin de conserver une solution basique. On s'inspire de la démonstration du lemme. Y_{j_0} s'écrit :

$$Y_{j_0} = \sum_{i=1}^m y_{j_0,i} Y_i$$

puisqu'on s'est ramené au cas où les premières colonnes de A forment la matrice identité. On prend $\epsilon \geq 0$ et on considère $B - \epsilon Y_{j_0}$. On regarde la première valeur de ϵ qui annule un coefficient, à la ligne l . La colonne numéro l quitte alors la base pour être remplacée par Y_{j_0} .

Etape 3 : On réécrit le système dans la nouvelle base donnée, et on itère le procédé jusqu'à obtenir une solution basique optimale.

Terminaison de l'algorithme : A chaque passage dans une boucle, la grandeur $C^t X$ diminue strictement. Or il existe un nombre fini de solutions basiques et une seule valeur de $C^t X$ est associée à chacune d'entre elles. Ceci impose que l'algorithme se termine. On repère que l'algorithme cesse lorsque dans l'étape 1 aucun des facteurs $c_j - z_j$ n'est strictement inférieur à 0. Il faut alors justifier que l'on a trouvé une solution basique optimale. C'est le cas puisque toute autre solution basique X doit avoir ses coefficients x_i positifs, aussi $C^t X \geq z_0$, z_0 étant la valeur de $C^t X_0$ où X_0 est la dernière solution basique rencontrée. X_0 est donc optimale et l'algorithme renvoie un bon résultat.

Pour plus de détails sur cet algorithme, on pourra se référer à [4] ou bien à [5].

Chapitre 3

Application au traitement de l'image : barycentres de Wasserstein par descente de gradient

On reprend ici les notations introduites dans le chapitre 1 pour le problème de transport optimal entre deux mesures atomiques sur \mathbb{R}^n .

Soit $N \in \mathbb{N}^*$. On se donne une famille des mesures de probabilité atomiques $(\nu_i)_{1 \leq i \leq N}$ sur un espace Ω , telles que pour tout i , ν_i est portée par un ensemble Y_i de cardinal m_i et est représentée par un vecteur de masses b_i . On dit qu'une mesure de probabilité atomique $\mu_{a,X}$ portée par un ensemble fini X et représentée par un vecteur de masse a est un barycentre de Wasserstein des (ν_i) si :

$$(a, X) \text{ minimise } \frac{1}{N} \sum_i W_p^p(\mu_{a,X}, \nu_i)$$

En pratique, nous allons prendre $\Omega = [1, 256]^{n^2}$ par exemple, pour qu'une image de n^2 pixels puisse être représentée comme une mesure atomique sur Ω .

3.1 Potentiels de Kantorovitch comme sous gradients

Posons $f : (a, X) \mapsto \frac{1}{N} \sum_i W_p^p(\mu_{a,X}, \nu_i)$. On souhaite réaliser une descente de gradient pour atteindre le minimum de f .

Définition 3. Soit C une partie convexe de \mathbb{R}^d pour un certain $d \geq 1$. Soit Ξ une fonctionnelle convexe sur C . On dit qu'un vecteur ω est un sous-gradient pour Ξ en $x \in C$ si :

$$\forall y \in C, \Xi(x) + \langle \omega | y - x \rangle \leq \Xi(y)$$

Théorème 5 (Potentiels de KANTOROVITCH comme sous-gradients). Soit $n, m \in \mathbb{N}^*$. Soit $b \in \mathbb{R}_+^m$ un vecteur de masses et $M = (m_{i,j}) \in M_{n,m}(\mathbb{R}_+)$ (matrice de coûts). La

fonctionnelle :

$$\Phi : a \in \mathbb{R}^n \longmapsto \sup_{\{\alpha, \beta | \forall i, j \alpha_i + \beta_j \leq m_{i,j}\}} \alpha^t a + \beta^t b$$

est convexe. On identifie $\Phi(a)$ comme un coût de transport minimal, et pour tout $a \in \mathbb{R}^n$ et tout couple de potentiels de KANTOROVITCH optimaux (α^*, β^*) , α^* est un sous-gradient pour Φ en a .

Démonstration. Pour $a \in \mathbb{R}^n$, $\Phi(a)$ est la solution du problème dual associé au programme linéaire :

$$\begin{cases} \min \sum_{i,j} m_{i,j} \gamma_{i,j} \\ \gamma \in \Pi(a, b) \end{cases}$$

Et on sait donc par dualité que $\Phi(a)$ est la solution de ce problème primal pour un b fixé. On écrit ce problème sous forme canonique :

$$\begin{cases} \min \langle c | u \rangle \\ Au = v \end{cases}$$

Pour une certaine matrice A et un certain vecteur c . Soit $v \in \mathbb{R}^n$ fixé. On pose $P(v) = \{u | Au = v, u \geq 0\}$, et $F(v) = \min_{u \in P(v)} \langle c | u \rangle$. F est une fonction convexe. Prenons en effet $\theta \in [0, 1]$ et v_1, v_2 des vecteurs. Posons $v = (1 - \theta)v_1 + \theta v_2$. Remarquons qu'en se restreignant dans un premier temps aux $(1 - \theta)u_1 + \theta u_2$ avec $u_1 \in P(v_1)$, $u_2 \in P(v_2)$ il vient :

$$\inf_{u=(1-\theta)u_1+\theta u_2; u_1 \in P(v_1), u_2 \in P(v_2)} \langle c | u \rangle = (1 - \theta)F(v_1) + \theta F(v_2)$$

D'où il résulte que :

$$F((1 - \theta)v_1 + \theta v_2) \leq (1 - \theta)F(v_1) + \theta F(v_2)$$

Soit maintenant p une solution optimale au problème dual pour v . On sait par dualité que $\langle p | v \rangle = F(v)$. Soit maintenant $u = Ax_0$ pour un certain $x_0 \geq 0$. Pour x une solution admissible au problème primal associé à u , on a $\langle p | u \rangle \leq \langle c | x \rangle$, donc par passage à la borne inférieure, $\langle p | u \rangle \leq F(u)$. D'où enfin :

$$F(v) + \langle p | u - v \rangle \leq F(u)$$

Revenons à Φ . Si on considère plus largement

$$\tilde{\Phi}(a, b) = \sup_{\{\alpha, \beta | \forall i, j \alpha_i + \beta_j \leq m_{i,j}\}} \alpha^t a + \beta^t b$$

les considérations précédentes permettent d'affirmer que pour un (a, b) fixé, tout couple (α^*, β^*) solution du problème dual est un sous-gradient pour $\tilde{\Phi}$. A fortiori, α^* est donc un sous-gradient pour Φ . Ceci achève la preuve du théorème. \square

Revenons à notre fonctionnelle

$$f : (a, X) \in \Xi \times \Omega^n \mapsto \frac{1}{N} \sum_{i=1}^N W_p^p(\mu_{a,X}, \nu_i)$$

où Ξ est un convexe inclus dans Σ_n que l'on suppose ouvert.

Soit $X \in \Omega^n$ fixé et $a \in \Xi$. Si on se donne pour tout i compris entre 1 et N un potentiel de KANTOROVICH optimal α_i^* associé au problème $W_p^p(\mu_{a,X}, \nu_i)$, alors naturellement,

$$\frac{1}{N} \sum_{i=1}^N \alpha_i^*$$

est un sous-gradient de l'application partielle $f(\cdot, X)$ au point a .

En effet, pour $1 \leq i \leq N$, on a par dualité de KANTOROVICH

$$W_p^p(\mu_{X,a}, \nu_i) = \sup_{\{\alpha, \beta \mid \forall i, j \ \alpha_i + \beta_j \leq \|x_i - y_j\|^p\}} \alpha^t a + \beta^t b$$

qui admet donc comme sous gradient à X fixé α_i^* .

L'application $a \mapsto f(a, X)$ étant convexe, l'algorithme de descente de gradient va donc converger vers un minimiseur de la fonctionnelle partielle (à X fixé). Nous pouvons procéder ainsi en calculant les α_i^* avec l'algorithme du simplexe (cf chapitre précédent). Mais nous voulons minimiser cette fonctionnelle par rapport à a et à X , ce qui conduit à nous intéresser à ce qui se passe en X .

3.2 Différentiabilité en le support des images

Dans cette partie, nous conserverons les notations de la partie précédente (à la définition de f près). Le but de ce chapitre est d'approcher un minimiseur de la fonctionnelle $f : (a, X) \mapsto \frac{1}{N} \sum_{i=1}^N W_p^p(\mu_{a,X}, \nu_i)$ par une descente de gradient.

La différentiabilité ou du moins sous-différentiabilité en le support X est beaucoup moins évidente que celle en a . Nous allons ici nous borner à l'étude des cas $p = 2$ et $\Omega = \mathbb{R}^d$ comme c'est le cas dans l'article de Cuturi [7].

Nous allons représenter le support de $\mu_{a,X}$ X comme une famille de n vecteurs $(x^i)_{1 \leq i \leq n}$ dans \mathbb{R}^d et donc comme une matrice de taille $d \times n$. On notera également (x_{ki}) cette matrice : x_{ki} désigne alors la k -ème coordonnée du i -ème vecteur x^i de la famille. De même, si Y représente le support d'une mesure ν de taille m , c'est une famille de m vecteurs et on peut donc la représenter comme une matrice $d \times m$. On note b_j la pondération de y^j pour la mesure ν . Puisque $p = 2$, les coûts qui nous intéressent sont les $m_{i,j} = \|x^i - y^j\|^2 = \sum_{k=1}^d |x_{ik} - y_{jk}|^2 = \sum_{k=1}^d |x_{ik}|^2 - 2x_{ik}y_{jk} + |y_{jk}|^2$. Si on note $M_{XY} = m_{i,j}$ la matrice des coûts associée, le coût d'un plan de transport $T = (t_{ij})$ est donné par

$$C(T) = \langle T, M_{XY} \rangle = \sum_{i,j} m_{ij} t_{ij}$$

On remarque alors que pour $T \in U(a, b)$

$$\langle T, M_{XY} \rangle = \sum_{i,j,k} |x_{ik}|^2 t_{ij} - 2x_{ik}y_{jk}t_{ij} + |y_{jk}|^2 t_{ij} = \sum_{i,k} a_i |x_{ik}|^2 + \sum_{j,k} b_j |y_{jk}|^2 - 2\langle T, {}^tXY \rangle$$

et donc si on note $\mathbf{x} = \text{diag}({}^tXX)$ et $\mathbf{y} = \text{diag}({}^tYY)$, on a alors

$$\langle T, M_{XY} \rangle = {}^t\mathbf{x}\mathbf{a} + {}^t\mathbf{y}\mathbf{b} - 2\langle T, {}^tXY \rangle$$

Ainsi,

$$W_2^2(\mu_{a,X}, \nu) = \min_{T \in U(a,b)} {}^t\mathbf{x}\mathbf{a} + {}^t\mathbf{y}\mathbf{b} - 2\langle T, {}^tXY \rangle = {}^t\mathbf{x}\mathbf{a} + {}^t\mathbf{y}\mathbf{b} + 2 \min_{T \in U(a,b)} \langle X, -Y^tT \rangle$$

Nous allons maintenant étudier, pour $a \in \Theta$ et ν fixés, la fonction

$$g : X \in M_{d,n}(\mathbb{R}) \mapsto W_2^2(\mu_{a,X}, \nu) = {}^t\mathbf{x}\mathbf{a} + {}^t\mathbf{y}\mathbf{b} + 2 \min_{T \in U(a,b)} \langle X, -Y^tT \rangle$$

On a alors la remarque évidente suivante :

Remarque 2. g est un minimum de fonctions convexes, et n'est donc pas a priori convexe.

Un calcul simple donne par ailleurs :

$${}^t\mathbf{x}\mathbf{a} + {}^t\mathbf{y}\mathbf{b} - 2\langle T, {}^tXY \rangle = {}^t\mathbf{y}\mathbf{b} + \|X \text{diag}(a^{1/2}) - Y^tT \text{diag}(a^{-1/2})\|^2 - \|Y^tT \text{diag}(a^{-1/2})\|^2.$$

On peut dans un premier temps faire la remarque suivante pour mieux comprendre la nature de g :

Remarque 3. $U(a, b)$ est un polytope convexe car c'est un convexe compact défini par un nombre fini d'inégalités affines. Notons T_1, \dots, T_M ses points extrémaux. Alors les points extrémaux de $-Y^tU(a, b)$ sont les $(-Y^tT_i)_{1 \leq i \leq M}$ et donc puisque $T \mapsto \langle X, -Y^tT \rangle$ est linéaire, elle atteint son minimum en un des points extrémaux. On peut donc écrire

$$g(X) = {}^t\mathbf{x}\mathbf{a} + {}^t\mathbf{y}\mathbf{b} + 2 \min_{1 \leq i \leq M} \langle X, -Y^tT_i \rangle$$

Si on note

$$\Gamma_i = \{X \in M_{d,n}(\mathbb{R}) : \forall i' \neq i, \langle X, -Y^tT_i \rangle \leq \langle X, -Y^tT_{i'} \rangle\}$$

Alors Γ_i est convexe car défini par des inégalités affines, et sur chaque Γ_i , on a

$$g(X) = g_i(X) = {}^t\mathbf{x}\mathbf{a} + {}^t\mathbf{y}\mathbf{b} + 2\langle X, -Y^tT_i \rangle = {}^t\mathbf{y}\mathbf{b} + \|X \text{diag}(a^{1/2}) - Y^tT_i \text{diag}(a^{-1/2})\|^2 - \|Y^tT_i \text{diag}(a^{-1/2})\|^2$$

qui est une fonction convexe (une sorte de parabole).

Remarquons que chaque g_i est de classe \mathcal{C}^1 et que l'on a

$$\forall X \in M_{d,n}(\mathbb{R}), \nabla g_i(X) = 2(X \text{diag}(a^{1/2}) - Y^t T_i \text{diag}(a^{-1/2}))$$

On peut alors établir la proposition suivante.

Proposition 1. *On a*

1. $g(X) \xrightarrow{\|X\| \rightarrow \infty} +\infty$ et donc en particulier g possède un minimum ;
2. si X est un minimiseur de g , alors il existe $1 \leq i \leq M$ tel que

$$X = Y^t T_i \text{diag}(a^{-1})$$

Il existe donc $1 \leq i \leq M$ tel que g atteint son minimum en $Y^t T_i \text{diag}(a^{-1})$;

3. *il est possible de trouver algorithmiquement un minimiseur de g .*

Démonstration. Prouvons tous les points.

1. En écrivant $g_i(X) = {}^t \mathbf{y} b + \|X \text{diag}(a^{1/2}) - Y^t T_i \text{diag}(a^{-1/2})\|^2 - \|Y^t T_i \text{diag}(a^{-1/2})\|^2$, il est évident que pour tout $1 \leq i \leq M$, on a

$$g_i(X) \xrightarrow{\|X\| \rightarrow \infty} +\infty$$

et donc que $g(X) = \min_{1 \leq i \leq M} g_i(X) \xrightarrow{\|X\| \rightarrow \infty} +\infty$ comme minimum d'un nombre fini de fonctions tendant vers l'infini. Cette même écriture $g = \min_i g_i$ montre que g est continue et donc que g possède un minimum sur $M_{d,n}(\mathbb{R})$.

2. Soit X un minimiseur de g . Il existe alors $1 \leq i \leq M$ tel que $g(X) = g_i(X) \geq g_i(Y^t T_i \text{diag}(a^{-1})) \geq g(Y^t T_i \text{diag}(a^{-1}))$. En particulier si X est un minimiseur de g , toutes les inégalités sont des égalités et donc $g_i(X) = g_i(Y^t T_i \text{diag}(a^{-1}))$ d'où $X = Y^t T_i \text{diag}(a^{-1})$. Par existence d'un minimiseur d'après le point précédent, on a bien l'existence de $1 \leq i \leq M$ tel que g atteint son minimum en $Y^t T_i \text{diag}(a^{-1})$.
3. Expliquons rapidement le principe d'un tel algorithme. On peut calculer les T_i comme point extrémaux d'un domaine défini par des contraintes affines. On peut donc calculer les $Y^t T_i \text{diag}(a^{-1})$. Puis on trouve les i tels que $Y^t T_i \text{diag}(a^{-1}) \in \Gamma_i$ en testant les inégalités ; il en existe au moins un (le minimiseur de 2.). Puis pour ces i restant, on trouve celui qui correspond à la valeur de $g(Y^t T_i \text{diag}(a^{-1}))$ minimale : $Y^t T_i \text{diag}(a^{-1})$ est alors le minimiseur cherché. □

L'algorithme présenté ci-dessus a l'énorme défaut de nécessiter le calcul de tous les points extrémaux de $U(a, b)$ ce qui n'est pas possible en un temps raisonnable. De plus, il permet de minimiser g , mais notre but est de minimiser f en a et en X . Nous allons expliquer un autre algorithme qui ressemble plutôt à une méthode de NEWTON par rapport à X . Cuturi prétend dans [7] que cet algorithme converge vers un minimiseur de f en a et en X .

Si on reprend la définition, il vient

$$W_2^2(\mu_{a,X}, \nu) = {}^t \mathbf{y}b + \|X \text{diag}(a^{1/2}) - Y^t T^* \text{diag}(a^{-1/2})\|^2 - \|Y^t T^* \text{diag}(a^{-1/2})\|^2$$

où T^* est un plan de transport optimal. Remarquons que ce plan n'est pas forcément un des T_i évoqués précédemment si il n'y a pas unicité du plan de transport optimal. Si on dispose d'un tel plan de transport, et si à a fixé on remplace X par un élément de la forme $\theta X + (1 - \theta)Y^t T^* \text{diag}(a^{-1})$ (ce qui est l'idée d'une méthode de Newton : on approche quadratiquement notre fonction partielle), alors on a $g(\theta X + (1 - \theta)Y^t T^* \text{diag}(a^{-1})) < g(X)$ dès que $\theta \in [0, 1[$. La raison pour laquelle on a intérêt à calculer T^* pour faire cette étape alors que c'est très coûteux est que puisque l'on calcule une solution du problème dual α^* de toutes les manières, calculer parallèlement une solution T^* du problème primal n'est pas très compliqué. L'algorithme que nous allons mettre en place pour tenter de calculer un minimiseur de f se déroule de la façon suivante :

- On se donne en entrée les mesures $(\nu_i)_{1 \leq i \leq N}$ avec leur support $Y_i \in M_{d, m_i}(\mathbb{R})$ et les pondérations $b_i \in \Sigma_{m_i}$ associées, et on initialise l'algorithme avec un $X \in M_{d, n}(\mathbb{R})$ et un $a \in \Sigma_n$.
- Puis tant que notre (X, a) n'a pas convergé (ce qu'il faudrait prouver), on calcule d'abord un nouveau a par un pas de descente de gradient à X fixé (comme dans la partie précédente). Puis pour tout $1 \leq i \leq N$, on calcule un plan de transport optimal T_i^* associé au coût $W_2^2(\mu_{X, a}, \nu_i)$. On applique enfin la méthode de NEWTON proposée ci-dessus pour changer X en

$$(1 - \theta)X + \theta \left(\frac{1}{N} \sum_{i=1}^N Y_i^t T_i^* \right) \text{diag}(a^{-1})$$

Remarque 4. On peut également prescrire le cardinal X , ce qui revient à ajouter une contrainte supplémentaire.

3.3 Un exemple simple

On considère maintenant des images codées sur 80x80 pixels en niveaux de gris. On les représente dans $[1, 80]^2$ comme des masses de dirac en les (i, j) avec i et j entiers, avec une masse proportionnelle à la teinte du pixel situé en (i, j) . Prenons les 4 images suivantes (un cercle désigne un pixel noir en son centre).

La moyenne géométrique de ces 4 images donne un nuage de point. Il devient délicat de remarquer qu'il s'agit d'une succession de point alignés verticalement. Si on considère un nombre plus important d'images, disons dès le seuil de 10, un observateur n'est plus capable de reconnaître cette propriété dans le nuage de point : la moyenne que l'on a produite a perdu le caractère géométrique sous-jacent à toutes ces images. Sans mauvais jeu de mot, on en a perdu le sens commun. C'est l'intérêt du barycentre de Wasserstein.

Montrons que la barycentre de Wasserstein des 4 images précédentes a bien la forme voulu. Par souci de concision, on impose au barycentre d'être un mesure dont le support

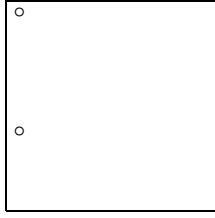


FIGURE 3.1 – a

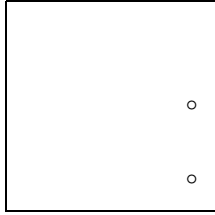


FIGURE 3.3 – c

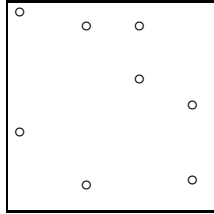


FIGURE 3.5 – Moyenne Empirique

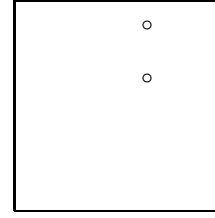


FIGURE 3.2 – b

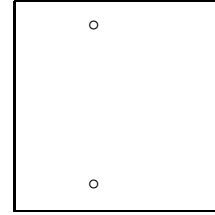


FIGURE 3.4 – d

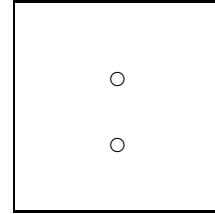


FIGURE 3.6 – Barycentre de Wasserstein

est de cardinal 2. Nous justifierons cette restriction un peu plus loin.

Etant données deux mesures $\{\delta_{(x_1, y_1)}; \delta_{(x_1, y_1 + \epsilon)}\}$ et $\{\delta_{(x_2, y_2)}; \delta_{(x_2, y_2 + \epsilon')}\}$, leur barycentre de Wasserstein pour W_2 est $\left\{ \delta_{\left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)}; \delta_{\left(\frac{x_1+x_2}{2}, \frac{y_1+y_2+\epsilon+\epsilon'}{2}\right)} \right\}$ dès que $\epsilon, \epsilon' > 0$. Autrement dit, il est obtenu par translation/dilatation entre les deux images/mesures de départ.

On écrit toutes les images qui précèdent comme des mesures de la forme $\mu_i = \left\{ \delta_{(x_{a,i}, y_{a,i})}; \delta_{(x_{b,i}, y_{b,i})} \right\}$ avec $i \in [1, 4]$. On cherche donc le barycentre de Wasserstein sous la forme $\rho = \left\{ \delta_{(x_a, y_a)}; \delta_{(x_b, y_b)} \right\}$.

Soit $i \in [1, 4]$.

$$W_2^2(\mu_1, \rho) = (x_{a,i} - x_a)^2 + (y_{a,i} - y_a)^2 + (x_{b,i} - x_b)^2 + (y_{b,i} - y_b)^2$$

En minimisant la somme des W_2^2 , on trouve :

$$\left\{ \begin{array}{l} x_a = \frac{x_{a,1} + x_{a,2} + x_{a,3} + x_{a,4}}{4} \\ y_a = \frac{y_{a,1} + y_{a,2} + y_{a,3} + y_{a,4}}{4} \\ x_b = \frac{x_{b,1} + x_{b,2} + x_{b,3} + x_{b,4}}{4} \\ y_b = \frac{y_{b,1} + y_{b,2} + y_{b,3} + y_{b,4}}{4} \end{array} \right.$$

Ce qui correspond bien à un barycentre de Wasserstein de nos 4 images de la forme avancée précédemment. Il convient pour terminer de justifier notre recherche d'un barycentre dont le support est de cardinal 2 et de discuter le résultat sans cette contrainte. D'une part on peut restreindre l'étude au cas où l'image finale est de support fini dans la mesure où c'est toujours ce que l'on fait : quand on fixe une image de taille 80x80 pixels par exemple, ce n'est jamais qu'une image 1000x1000 pixels dont le support est de cardinal $80^2 = 6400$. D'autre part, sans cette contrainte, on aurait obtenu un résultat similaire (à observer sur une simulation numérique, car le calcul sur papier est fastidieux). Les points chargés de masse sous l'hypothèse "le support est de cardinal " sont toujours les plus chargés dans le cas général, mais certains points dans leur voisinage le sont aussi, formant une tache colorée dont l'intensité est décroissante à mesure que l'on s'éloigne du centre de la tache.

Ce calcul modeste peut sembler parfaitement trivial. Il illustre cependant le principal intérêt de la distance de Wasserstein pour le traitement d'image : les barycentres produits conservent relativement bien les formes géométriques des images de départ. Une application plus frappante proposée par Cuturi dans [7] est de calculer la moyenne d'écritures humaines. Si on demande à un panel d'individu d'écrire le chiffre 2 sur un écran tactile, tous l'écrivent avec une forme différente, une taille différente et une position différente sur l'écran. La moyenne empirique des écritures est une tache, leur barycentre de Wasserstein est un 2 lisible.

Chapitre 4

Calcul par régularisation entropique

Dans tout ce chapitre, on fixe $a \in \Sigma_n$ et $b \in \Sigma_m$ représentant deux mesures de probabilités discrètes où toutes les pondérations sont strictement positives, $M = (m_{ij})$ une matrice des coûts (où (m_{ij}) est le coût de transport de i vers j). On rappelle que l'on note

$$U(a, b) = \left\{ T \in M_{n,m}(\mathbb{R}_+) : \sum_{j=1}^m t_{ij} = a_i, 1 \leq i \leq n; \sum_{i=1}^n t_{ij} = b_j, 1 \leq j \leq m \right\}$$

l'ensemble des plans de transports de marginales a et b . Nous cherchons à minimiser, pour $T \in U(a, b)$ le coût de transport

$$C(T) = \langle T, M \rangle = \sum_{i=1}^n \sum_{j=1}^m m_{ij} t_{ij}$$

Le but de cette partie est de trouver un moyen efficace d'approcher un plan de transport optimal T^* , ainsi qu'une solution α^*, β^* du problème dual

$$\max_{\alpha \oplus \beta \leq M} {}^t \alpha a + {}^t \beta b$$

4.1 Régularisée entropique et approximation de la solution

Dans cette section, nous allons définir une suite de nouveaux problèmes dits "régularisées" qui vont approcher notre problème initial. Dans un premier temps nous allons définir ces nouveaux problèmes, et dans un second temps nous allons montrer que leurs solutions convergent en un certain sens vers les solutions de notre problème initial. Nous nous inspirons ici principalement du cours de B. Maury [6] pour les démonstrations, et reprenons en grande partie les notations de l'article [7].

4.1.1 Définitions, formulation des problèmes primal et dual

Pour trouver une bonne estimation du plan de transport, on va quelque peu modifier la fonction de coût en lui ajoutant un facteur entropique.

Définition 4 (Régularisée entropique). Soit $n, m \in \mathbb{N}$, et $T = (t_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ un plan de transport. On définit son entropie

$$h(T) = \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq m} t_{ij} \log t_{ij}$$

Pour $\lambda > 0$, on définit la régularisée entropique de C comme

$$C_\lambda(T) = C(T) + \frac{1}{\lambda} h(T) = \sum m_{ij} t_{ij} + \frac{1}{\lambda} \sum t_{ij} \log t_{ij}$$

Le problème de minimisation de la fonction de coût C_λ admet une formulation duale.

Proposition 2. On a la formulation duale :

$$\min_{T \in U(a,b)} C_\lambda(T) = \max_{(\alpha, \beta) \in \mathbb{R}^{n+m}} t \alpha a + t \beta b - \frac{1}{e\lambda} \sum_{i,j} e^{-\lambda(m_{ij} - \alpha_i - \beta_j)}$$

On utilise d'abord le lemme

Lemme 1. Puisque a et b ont toutes leur coordonnées strictement positives (hypothèse sur les pondérations faite au début), pour tout $\lambda > 0$, C_λ admet un unique minimiseur T^λ sur $U(a, b)$ avec $t_{ij}^\lambda > 0$ pour tout i, j .

Démonstration. Puisque $U(a, b)$ est un convexe compact et que C_λ est continue et strictement convexe sur $U(a, b)$, C_λ admet bien un unique minimiseur T^λ sur $U(a, b)$.

On montre que $t_{ij}^\lambda > 0$ en remarquant que la dérivée en 0 de $x \log x$ est $-\infty$ ce qui force à éviter les bords. □

Formulons le problème en termes de Lagrangien. Tout découlera ensuite de l'existence et "quasi-unicité" d'un point-selle.

Lemme 2. Soit $U = M_{n,m}(\mathbb{R}_{++})$, on veut minimiser C_λ sous les contraintes affines

$$\varphi_i(T) = a_i - \sum_j t_{ij} = 0, \quad \psi_j(T) = b_j - \sum_i t_{ij} = 0$$

On introduit donc le Lagrangien

$$\begin{aligned} \mathcal{L}_\lambda & : U \times \mathbb{R}^{n+m} \longrightarrow \mathbb{R} \\ T, (\alpha, \beta) & \longmapsto C_\lambda(T) + \sum_i \alpha_i \varphi_i(T) + \sum_j \beta_j \psi_j(T) \end{aligned}$$

Alors il admet un point selle $(T^\lambda, \alpha^\lambda, \beta^\lambda)$

Démonstration. Notons $K = \{T \in M_{n,m}(\mathbb{R}) : \varphi_i(T) = 0, \psi_j(T) = 0 \forall i, j\}$

On sait que la fonctionnelle C_λ admet un minimiseur sur $U \cap K$ en T^λ . Puisque les contraintes d'égalité qui définissent K sont affines et que C_λ est une fonctionnelle \mathcal{C}^1 sur U , on a l'existence de multiplicateurs de Lagrange : il existe $\alpha^\lambda, \beta^\lambda$ tels que

$$(\alpha^\lambda, \beta^\lambda) \cdot (\varphi(T^\lambda), \psi(T^\lambda)) = 0 \text{ et } \nabla C_\lambda(T^\lambda) + \sum_i \alpha_i^\lambda \nabla \varphi_i(T^\lambda) + \sum_j \beta_j^\lambda \nabla \psi_j(T^\lambda) = 0$$

La première condition est réalisée puisque $T^\lambda \in U(a, b)$, la condition d'optimalité est donc

$$\forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, m \rrbracket, m_{ij} + \frac{1}{\lambda}(1 + \log t_{ij}^\lambda) - \alpha_i^\lambda - \beta_j^\lambda = 0 \quad (*)$$

Réciproquement, puisque la fonctionnelle C_λ est convexe et que les contraintes le sont également, un tel triplet $T^\lambda, \alpha^\lambda, \beta^\lambda$ est un point-selle du lagrangien. \square

Nous pouvons maintenant prouver la formulation duale du problème :

Démonstration. Puisque le lagrangien admet comme point-selle $T^\lambda, \alpha^\lambda, \beta^\lambda$, on a :

$$\min_{T \in U(a,b)} C_\lambda(T) = \max_{(\alpha, \beta) \in \mathbb{R}^{n+m}} G_\lambda(\alpha, \beta) \text{ avec } G_\lambda(\alpha, \beta) = \inf_{T \in U} L_\lambda(T, \alpha, \beta)$$

Or pour α, β fixés :

$$\mathcal{L}_\lambda(T, \alpha, \beta) = \sum_{i=1}^n \alpha_i a_i + \sum_{j=1}^m \beta_j b_j + \left(\sum_{i,j} (m_{ij} - \alpha_i - \beta_j + \frac{1}{\lambda} \log t_{ij}) t_{ij} \right)$$

où le terme entre parenthèses est une somme de termes strictement convexes en les t_{ij} . En minimisant la valeur de chaque terme, on retrouve comme dans (*) que le minimiseur est caractérisé par

$$t_{ij} = \frac{1}{e} e^{-\lambda(m_{ij} - \alpha_i - \beta_j)}$$

et donc finalement on a bien :

$$G_\lambda(\alpha, \beta) = \alpha a + \beta b - \frac{1}{e\lambda} \sum_{i,j} e^{-\lambda(m_{ij} - \alpha_i - \beta_j)}$$

\square

Remarquons que l'on a l'unicité de la solution du problème primal car le minimiseur de C_λ est unique. En revanche, nous n'avons pas a priori unicité de la solution au problème dual et donc de α^λ . La proposition suivante montre qu'il y a presque unicité :

Proposition 3. *Le problème dual admet un unique maximiseur $(\alpha^\lambda, \beta^\lambda)$ tel que la moyenne de α^λ est nulle.*

Démonstration. Calculons la matrice Hessienne de G_λ ce qui revient essentiellement à calculer la matrice Hessienne de

$$F : (\alpha, \beta) \mapsto \sum_{i,j} e^{\lambda(\alpha_i + \beta_j - m_{ij})}$$

On a que

$$\frac{\partial F(\alpha, \beta)}{\partial \alpha_{i_1} \partial \alpha_{i_2}} = \begin{cases} 0 & \text{si } i_1 \neq i_2 \\ \lambda^2 e^{\lambda \alpha_i} \sum_j e^{\lambda(\beta_j - m_{ij})} & \text{si } i = i_1 = i_2 \end{cases}$$

$$\frac{\partial F(\alpha, \beta)}{\partial \beta_{j_1} \partial \beta_{j_2}} = \begin{cases} 0 & \text{si } j_1 \neq j_2 \\ \lambda^2 e^{\lambda \beta_j} \sum_i e^{\lambda(\alpha_i - m_{ij})} & \text{si } j = j_1 = j_2 \end{cases}$$

$$\frac{\partial F(\alpha, \beta)}{\partial \alpha_i \partial \beta_j} = \lambda^2 e^{\lambda(\alpha_i + \beta_j - m_{ij})}$$

On a donc, pour $(\alpha, \beta) \in \mathbb{R}^{n+m}$ fixé, pour tout $(x, y) \in \mathbb{R}^{n+m}$, $H(\alpha, \beta)[(x, y)] \cdot (x, y)$ qui vaut

$$\lambda^2 \sum_i e^{\lambda \alpha_i} x_i^2 \sum_j e^{\lambda(\beta_j - m_{ij})} + \lambda^2 \sum_j e^{\lambda \beta_j} y_j^2 \sum_i e^{\lambda(\alpha_i - m_{ij})} + 2\lambda^2 \sum_{i,j} x_i y_j e^{\lambda(\alpha_i + \beta_j - m_{ij})}$$

ce qui s'écrit également

$$\lambda^2 \sum_{i,j} (x_i + y_j)^2 e^{\lambda(\alpha_i + \beta_j - m_{ij})}$$

Cette quantité est strictement positive si, et seulement si $(x, y) \notin \text{Vect}(-1, 1)$ et est nulle si et seulement si $(x, y) \in \text{Vect}(-1, 1)$.

La matrice Hessienne de G_λ est donc négative et de noyau engendré par $(-1, 1)$ pour tout couple (α, β) .

On remarque ensuite que $\forall v \in \text{Vect}(-1, 1)$, $G_\lambda((\alpha, \beta) + v) = G_\lambda(\alpha, \beta)$ ce qui permet de restreindre G_λ à l'orthogonal de $(-1, 1)$. On sait alors que sur cet ensemble, toutes les Hessiennes de G_λ sont définies négatives, et donc que G_λ admet un unique maximiseur sur l'orthogonal de $(-1, 1)$ qui est aussi le maximiseur sur tout \mathbb{R}^{n+m} . Ainsi, l'ensemble des maximiseur est une droite affine de la forme $(\alpha, \beta) + \mathbb{R}(-1, 1)$ et il existe bien un unique maximiseur $(\alpha^\lambda, \beta^\lambda)$ tel que la moyenne de α^λ soit nulle. \square

Dorénavant, nous désignerons par $(T^\lambda, \alpha^\lambda, \beta^\lambda)$ l'unique point-selle de L_λ tel que la moyenne de α^λ soit nulle.

Remarque 5. On peut réécrire C_λ comme une divergence de KULLBACK-LEIBLER :

$$C_\lambda(T) = \frac{1}{\lambda} \sum_{i,j} t_{ij} \log \left(\frac{t_{ij}}{\eta_{ij}} \right) \text{ avec } \eta_{ij} = e^{-\lambda m_{ij}}$$

Les conditions d'optimalité (*) s'écrivent alors

$$1 + \log \left(\frac{t_{ij}}{\eta_{ij}} \right) + \lambda \alpha_i + \lambda \beta_j = 0$$

4.1.2 Convergence des solutions du problème régularisé vers les solution du problème initial

Nous avons donc, pour $\lambda > 0$, un point selle du Lagrangien $(T^\lambda, \alpha^\lambda, \beta^\lambda)$ tel que la moyenne de α^λ soit nulle. La question qui se pose est maintenant de savoir si ce triplet converge, lorsque λ tend vers $+\infty$ vers un point selle du lagrangien du problème non régularisé.

On peut montrer (cf [6]) que T^λ converge vers un plan de transport optimal du problème non régularisé qui minimise l'entropie parmi les plans de transports optimaux. Pour ce qui est des $\alpha^\lambda, \beta^\lambda$, on peut montrer qu'elles sont bornées et que toute suite extraite convergente converge vers une solution du problème dual; le problème est que cette solution n'est pas toujours unique. Mais algorithmiquement, si on se donne λ suffisamment grand, alors le triplet sera proche d'un point-selle, quand bien même celui-ci ne serait pas unique.

4.2 Calcul effectif des solutions au problèmes primal et dual pour la régularisée entropique

D'après la section précédente, nous pouvons donc avoir une bonne approximation de la solution de notre problème si on résoud le problème régularisé pour λ suffisamment grand. Dans cette partie, nous proposons une façon efficace de calculer ces solutions, en s'appuyant principalement sur l'algorithme de SINKHORN [9] conformément à ce qui est proposé dans l'article de Cuturi [7].

4.2.1 Reformulation du problème

Ici nous reformulons juste le problème pour se ramener à l'algorithme de SINKHORN.

Proposition 4. *Posons $K = (\eta_{ij}) = (e^{-\lambda m_{ij}}) \in M_{n,m}(\mathbb{R}_{++})$ Alors il existe deux vecteurs $u \in \mathbb{R}_{++}^n$ et $v \in \mathbb{R}_{++}^m$ tels que*

$$T^\lambda = \text{diag}(u)K \text{diag}(v) \text{ et } \alpha^\lambda = -\frac{\log u}{\lambda} + \frac{\sum_i \log u_i}{\lambda n} (1, \dots, 1)$$

Démonstration. Soit $(T^\lambda, \alpha^\lambda, \beta^\lambda)$ le point-selle de L_λ . On pose

$$u_i := e^{-\lambda \alpha_i^\lambda}, \quad 1 \leq i \leq n \text{ et } v_j = e^{-\lambda \beta_j^\lambda - 1}, \quad 1 \leq j \leq m$$

Les vecteurs u et v vérifient évidemment la deuxième condition. La première est conséquence directe de l'écriture des conditions d'optimalité. □

Énonçons maintenant une forme de réciproque

Proposition 5. Soit $T \in U(a, b)$. Supposons qu'il existe deux vecteurs $u \in \mathbb{R}_{++}^n$ et $v \in \mathbb{R}_{++}^m$ tels que

$$T = \text{diag}(u)K\text{diag}(v) \Leftrightarrow \forall (i, j) \in \llbracket 1, n \rrbracket \times \llbracket 1, m \rrbracket, t_{ij} = u_i v_j \eta_{ij}$$

Alors $T^\lambda = T$ et $\alpha^\lambda = -\frac{\log u}{\lambda} + \frac{\sum_i \log u_i}{\lambda n} (1, \dots, 1)$

Démonstration. Posons

$$\alpha_i := -\frac{\log u_i}{\lambda} \text{ et } \beta_j = -\frac{\log v_j}{\lambda} + 1$$

Nous avons alors immédiatement la relation (*) qui est la relation sur les gradients dans le théorème de KUHN et TUCKER. Or puisque $T \in U(a, b)$ il vérifie les contraintes et donc par le théorème de KUHN et TUCKER, (T, α, β) est un point-selle de L_λ . Il s'en suit que $T = T^\lambda$ par unicité du minimiseur de C_λ et que α est de la forme $\alpha^\lambda + x$. En regardant la moyenne, on trouve le résultat. \square

4.2.2 Résolution : l'algorithme de SINKHORN

Théorème 6 (SINKHORN(1967)). Soit $m, n \in \mathbb{N}^*$.

Fixons des entiers strictement positifs r_1, \dots, r_n et c_1, \dots, c_m (r pour row et c pour column). On pose $\mu = \frac{\sum_j c_j}{\sum_i r_i}$.

Soit $A \in \mathcal{M}_{n,m}(\mathbb{R}_{++})$ une matrice à coefficients strictement positifs. Alors :

1. il existe une unique matrice B telle que :
 - $\forall i \in \llbracket 1, n \rrbracket, \sum_{j=1}^m B_{ij} = \mu r_i$ et $\forall j \in \llbracket 1, m \rrbracket, \sum_{i=1}^n B_{ij} = c_j$ (on impose les sommes des lignes et des colonnes) ;
 - Il existe deux matrices $D_1 \in D_n(\mathbb{R}_{++})$ et $D_2 \in D_m(\mathbb{R}_{++})$ telles que

$$D_1 A D_2 = B$$

2. les matrices D_1 et D_2 sont unique à une multiplication par un scalaire strictement positif près ;
3. Le procédé qui consiste à renormaliser successivement les lignes et les colonnes de A (pour que la somme vaille bien ce que l'on veut sur chaque ligne/colonne) converge. Plus précisément si on pose
 - * $A_0 = A$;
 - * pour $k \geq 0$

i on note $r^{(k)} = A_k e$ et donc $r_i^{(k)}$ est la valeur de la somme de la i -ème ligne de la matrice A_k . On définit $r/r^{(k)}$ le vecteur de coordonnées $r_i/r_i^{(k)}$.
On pose alors $S_k = \text{diag}(r/r^{(k)}) \in D_n(\mathbb{R}_{++})$ et $A'_k = S_k A_k$ la renormalisation des lignes de A^k .

ii on pose de même $c^{(k)} = {}^t e A'_k$; la somme des termes de la colonne j de la matrice A'_k vaut donc $c_j^{(k)}$. On pose alors $T_k = \text{diag}(c/c^{(k)}) \in D_m(\mathbb{R}_{++})$ et $A_{k+1} = A'_k T_k$ qui est donc une renormalisation des colonnes.

Alors la suite A_k (où les colonnes sont renormalisées) converge vers B et la suite A'_k converge vers $\frac{1}{\mu}B$.

Commençons par les preuves d'unicité.

Démonstration. Supposons que l'on ait

$$D_1 = \text{diag}(x), \quad D_2 = \text{diag}(y), \quad D'_1 = \text{diag}(x'), \quad D'_2 = \text{diag}(y')$$

telles que D_1AD_2 et $D'_1AD'_2$ satisfassent les conditions voulues. Posons

$$\forall i \in \llbracket 1, n \rrbracket, \quad p_i = \frac{x'_i}{x_i} \quad \text{et} \quad \forall j \in \llbracket 1, m \rrbracket, \quad q_j = \frac{y'_j}{y_j}$$

On a que

$$\mu r_i = \sum_{j=1}^m x'_i a_{ij} y'_j = p_i \sum_{j=1}^m x_i a_{ij} q_j y_j$$

Donc

$$\min_{1 \leq i \leq n} p_i = p_{i_0} = \mu r_{i_0} \left(\sum_{j=1}^m x_{i_0} a_{i_0 j} q_j y_j \right)^{-1} \geq \left(\max_{1 \leq j \leq m} q_j \right)^{-1} \mu r_{i_0} \left(\sum_{j=1}^m x_{i_0} a_{i_0 j} y_j \right) = \left(\max_{1 \leq j \leq m} q_j \right)^{-1} \quad (1)$$

De même on montre que

$$\max_{1 \leq j \leq m} q_j \leq \left(\min_{1 \leq i \leq n} p_i \right)^{-1}$$

On a donc

$$\max_{1 \leq j \leq m} q_j = \left(\min_{1 \leq i \leq n} p_i \right)^{-1}$$

Mais si les q_j ne sont pas tous égaux, l'inégalité est stricte dans (1). Donc tous les q_j sont égaux à un certain q . De même, tous les p_i sont égaux à un certain p et on a $pq = 1$. On a ainsi montré que $D'_1 = pD_1$, $D'_2 = qD_2$ et donc $D'_1AD'_2 = pqD_1AD_2 = D_1AD_2$ ce qui montre à la fois l'unicité dans le point 1. et dans le point 2.. \square

Pour prouver l'existence et la convergence de l'algorithme de renormalisation, nous allons reprendre la procédure du théorème en posant également deux nouvelles suites auxiliaires $x^{(k)} \in \mathbb{R}_{++}^n$ et $y^{(k)} \in \mathbb{R}_{++}^m$.

Définition 5 (Suites auxiliaires). *On note $a = \min_{ij} a_{ij}$ et $\delta = \max_{i,j} c_j, r_i$. On définit alors*

- ★ $x^{(0)} = \delta^{1/2} a^{-1/2} e \in \mathbb{R}_{++}^n$ et $y_j^{(0)} = \delta^{-1/2} a^{1/2} c_j (\sum_{i=1}^n a_{ij})^{-1}$ et $A_0 = \text{diag}(x^{(0)}) A \text{diag}(y^{(0)})$
la renormalisation des colonnes de A ;
- ★ pour $k \geq 0$

i on note $r^{(k)} = A_k e$ et donc $r_i^{(k)}$ est la valeur de la somme de la i -ème ligne de la matrice A_k . On définit $r/r^{(k)}$ le vecteur de coordonnées $r_i/r_i^{(k)}$.

On pose alors $S_k = \text{diag}(r/r^{(k)}) \in D_n(\mathbb{R}_{++})$ et $A'_k = S_k A_k$ la renormalisation des lignes de A^k .

On note $M_k = \|(r_i x_i^{(k)}/r_i^{(k)})_{1 \leq i \leq n}\|$. On pose alors $x_i^{(k+1)} = \delta^{1/2} a^{-1/2} M_k^{-1} (r/r^{(k)})_i x_i^{(k)}$.

ii on pose de même $c^{(k)} = {}^t e A'_k$; la somme des termes de la colonne j de la matrice A'_k vaut donc $c_j^{(k)}$. On pose alors $T_k = \text{diag}(c/c^{(k)}) \in D_m(\mathbb{R}_{++})$ et $A_{k+1} = A'_k T_k$ qui est donc une renormalisation des colonnes.

On pose enfin $y_j^{(k+1)} = \delta^{-1/2} a^{1/2} M_k (c/c^{(k)})_j y_j^{(k)}$

Remarque 6. Pour que les suites A_k coïncident dans la définition que nous en faisons ici et dans l'énoncé du théorème précédent, il faut que les colonnes de A soit déjà renormalisée. Il est évident qu'on peut s'y ramener quitte à effectuer une étape d'algorithme avant.

Remarque 7. Au vu des définitions de x^k et $y^{(k)}$, on a bien que pour tout $k \in \mathbb{N}$, $A_k = \text{diag}(x^{(k)}) \text{Adiag}(y^{(k)})$. En particulier, puisque les A_k sont toutes renormalisée en colonnes, on a que pour tout $k \in \mathbb{N}$:

$$\forall j \in \llbracket 1, m \rrbracket, \sum_{i=1}^n x_i^{(k)} a_{ij} y_j^{(k)} = c_j$$

Nous allons prouver le lemme suivant :

Lemme 3. Posons

$$V := \left\{ (x, y) \in \mathbb{R}^{n+m} : \forall j \in \llbracket 1, m \rrbracket, \sum_{i=1}^n x_i a_{ij} y_j = c_j \text{ et } \|x\| \leq \delta^{1/2} a^{-1/2} \text{ et } \|y\| \leq \delta^{1/2} a^{-1/2} \right\}$$

On définit sur V la fonction

$$\begin{aligned} \varphi : V &\longrightarrow \mathbb{R}_+ \\ (x, y) &\longmapsto \max_{1 \leq i \leq n} \left(r_i^{-1} \sum_{j=1}^m x_i a_{ij} y_j \right) - \min_{1 \leq i \leq n} \left(r_i^{-1} \sum_{j=1}^m x_i a_{ij} y_j \right) \end{aligned}$$

Alors :

1. pour tout $k \in \mathbb{N}$, $(x^{(k)}, y^{(k)}) \in V$;
2. $\varphi(x^{(k)}, y^{(k)}) \rightarrow 0$;
3. toute valeur d'adhérence de la suite $(x^{(k)}, y^{(k)})$ est à coordonnées strictement positives.

Avant de le prouver, expliquons juste pourquoi ce lemme prouve tout ce qu'il nous manquait.

Le premier point du lemme montre que $x^{(k)}$ et $y^{(k)}$ sont bornées ce qui montre que la suite A_k que nous avons défini dans la deuxième procédure est bornée puisque $A_k = \text{diag}(x^{(k)})\text{Adiag}(y^{(k)})$. Soit $B' = \lim_k A_{\psi(k)}$ une valeur d'adhérence de la suite (A_k) . Puisque la suite $(x^{(\psi(k))}, y^{(\psi(k))})$ est bornée, elle a une valeur d'adhérence (x, y) qui d'après le lemme a toutes ses coordonnées strictement positives : on peut donc écrire $B' = \text{diag}(x)\text{Adiag}(y)$ avec x, y strictement positifs. Par propriété des A_k , pour tout j , $\sum_i B'_{ij} = c_j$ en faisant passer cette condition à la limite. Puisque $\varphi(x^{(k)}, y^{(k)}) \rightarrow 0$, on a également l'existence d'une constante K strictement positive telle que pour tout i , $\sum_j B'_{ij} = Kr_i$. Or $K \sum_i r_i = \sum_j c_j$ et donc $K = \mu$, ce qui montre que B' est la matrice B du théorème dont nous avons prouvé l'unicité (cela prouve donc l'existence de B). On a ainsi que la suite A_k de la deuxième procédure admet une unique valeur d'adhérence qui est B : elle converge donc vers B .

Montrons à présent le lemme.

Démonstration. 1. Tout d'abord, une remarque précédente montre que pour tout $k \in \mathbb{N}$, le couple $(x^{(k)}, y^{(k)})$ vérifie la condition sur la somme des colonnes. Montrons donc les bornes voulues.

- On a $\|x^{(0)}\| = \|\delta^{1/2}a^{-1/2}e\| = \delta^{1/2}a^{-1/2}$ et $\|y^{(0)}\| = \|\delta^{-1/2}a^{1/2}c_j (\sum_{i=1}^n a_{ij})^{-1}\| \leq \delta^{1/2}a^{-1/2}$. Ils sont donc bien dans V
- Soit $k \in \mathbb{N}$. Par définition, $|x_i^{(k+1)}| = \delta^{1/2}a^{-1/2}M_k^{-1}(r_i/r_i^{(k)})x_i^{(k)} \leq \delta^{1/2}a^{-1/2}M_k^{-1}M_k$ par définition de M_k et donc

$$\|x^{(k)}\| \leq \delta^{1/2}a^{-1/2}$$

- Soit $1 \leq j \leq m$. Alors

$$c_j^{(k)} = \sum_{i=1}^n (A'_k)_{ij} = \sum_{i=1}^n (r/r^{(k)})_i x_i^{(k)} a_{ij} y_j^{(k)}$$

Et donc

$$(c/c^{(k)})_j y_j^{(k)} = c_j \left(\sum_{i=1}^n (r/r^{(k)})_i x_i^{(k)} a_{ij} \right)^{-1}$$

Ainsi, pour tout $i \in \llbracket 1, n \rrbracket$, on a :

$$(c/c^{(k)})_j y_j^{(k)} \leq c_j \left((r/r^{(k)})_i x_i^{(k)} a_{ij} \right)^{-1} \leq \left((r/r^{(k)})_i x_i^{(k)} \right)^{-1} \delta a^{-1}$$

et donc puisque ceci est vrai pour tout i par définition de M_k il vient

$$(c/c^{(k)})_j y_j^{(k)} \leq M_k^{-1} \delta a^{-1}$$

Ainsi, par définition

$$|y_j^{(k+1)}| = \delta^{-1/2} a^{1/2} M_k(c/c^{(k)})_j y_j^{(k)} \leq \delta^{1/2} a^{-1/2}$$

et donc

$$\|y^{(k+1)}\| \leq \delta^{1/2} a^{-1/2}$$

La suite $(x^{(k)}, y^{(k)})$ est donc bien à valeurs dans V .

2. On note ici $\alpha_i^{(k)} = \frac{r_i^{(k)}}{r_i}$ et $\beta_j^{(k)} = \frac{c_j^{(k)}}{c_j}$ pour tout $k \in \mathbb{N}$, $1 \leq i \leq n$, $1 \leq j \leq m$. Soit $k \in \mathbb{N}$, on fixe $1 \leq i_1^k, i_2^k \leq n$ et $1 \leq j_1^k, j_2^k \leq m$ tels que $\alpha_{i_2^k}^{(k)} = \max_{1 \leq i \leq n} \alpha_i^{(k)}$, $\alpha_{i_1^k}^{(k)} = \min_{1 \leq i \leq n} \alpha_i^{(k)}$ et de même pour β . On a alors

$$\begin{aligned} \varphi(x^{(k+1)}, y^{(k+1)}) &= \alpha_{i_2^{k+1}}^{(k+1)} - \alpha_{i_1^{k+1}}^{(k+1)} \\ &= r_{i_2^{k+1}}^{-1} \sum_{j=1}^m x_{i_2^{k+1}}^{(k+1)} a_{i_2^{k+1}j}^{(k+1)} y_j^{(k+1)} - r_{i_1^{k+1}}^{-1} \sum_{j=1}^m x_{i_1^{k+1}}^{(k+1)} a_{i_1^{k+1}j}^{(k+1)} y_j^{(k+1)} \\ &= r_{i_2^{k+1}}^{-1} \sum_{j=1}^m \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1}j}^{(k)} \left(\beta_j^{(k)} \right)^{-1} y_j^{(k)} \\ &\quad - r_{i_1^{k+1}}^{-1} \sum_{j=1}^m \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1}j}^{(k)} \left(\beta_j^{(k)} \right)^{-1} y_j^{(k)} \\ &= r_{i_2^{k+1}}^{-1} \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1}j_2^k}^{(k)} \left(\beta_{j_2^k}^{(k)} \right)^{-1} y_{j_2^k}^{(k)} \\ &\quad + r_{i_2^{k+1}}^{-1} \sum_{j \neq j_2^k} \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1}j}^{(k)} \left(\beta_j^{(k)} \right)^{-1} y_j^{(k)} \\ &\quad - r_{i_1^{k+1}}^{-1} \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1}j_1^k}^{(k)} \left(\beta_{j_1^k}^{(k)} \right)^{-1} y_{j_1^k}^{(k)} \\ &\quad - r_{i_1^{k+1}}^{-1} \sum_{j \neq j_1^k} \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1}j}^{(k)} \left(\beta_j^{(k)} \right)^{-1} y_j^{(k)} \end{aligned}$$

Par les définitions précédentes, on a donc

$$\begin{aligned} \varphi(x^{(k+1)}, y^{(k+1)}) &\leq r_{i_2^{k+1}}^{-1} \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1}j_2^k}^{(k)} \left(\beta_{j_2^k}^{(k)} \right)^{-1} y_{j_2^k}^{(k)} \\ &\quad + r_{i_2^{k+1}}^{-1} \left(\beta_{j_1^k}^{(k)} \right)^{-1} \sum_{j \neq j_2^k} \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1}j}^{(k)} y_j^{(k)} \\ &\quad - r_{i_1^{k+1}}^{-1} \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1}j_1^k}^{(k)} \left(\beta_{j_1^k}^{(k)} \right)^{-1} y_{j_1^k}^{(k)} \\ &\quad - r_{i_1^{k+1}}^{-1} \left(\beta_{j_2^k}^{(k)} \right)^{-1} \sum_{j \neq j_1^k} \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1}j}^{(k)} y_j^{(k)} \end{aligned}$$

En remarquant qu'avec les définitions

$$\sum_j \left(\alpha_i^{(k)} \right)^{-1} x_i^{(k)} a_{ij} y_j^{(k)} = \sum_j (A'_k)_{ij} = r_i^{(k)}$$

on a

$$\begin{aligned} \varphi(x^{(k+1)}, y^{(k+1)}) &= \left(\beta_{j_2^k}^{(k)} \right)^{-1} r_{i_2^{k+1}}^{-1} \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1} j_2^k} y_{j_2^k}^{(k)} \\ &+ \left(\beta_{j_1^k}^{(k)} \right)^{-1} \left(1 - r_{i_2^{k+1}}^{-1} \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1} j_2^k} y_{j_2^k}^{(k)} \right) \\ &- \left(\beta_{j_1^k}^{(k)} \right)^{-1} r_{i_1^{k+1}}^{-1} \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1} j_1^k} y_{j_1^k}^{(k)} \\ &- \left(\beta_{j_2^k}^{(k)} \right)^{-1} \left(1 - r_{i_1^{k+1}}^{-1} \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1} j_1^k} y_{j_1^k}^{(k)} \right) \\ &= \left(\left(\beta_{j_1^k}^{(k)} \right)^{-1} - \left(\beta_{j_2^k}^{(k)} \right)^{-1} \right) \\ &\quad \left(1 - r_{i_1^{k+1}}^{-1} \left(\alpha_{i_1^{k+1}}^{(k)} \right)^{-1} x_{i_1^{k+1}}^{(k)} a_{i_1^{k+1} j_1^k} y_{j_1^k}^{(k)} - r_{i_2^{k+1}}^{-1} \left(\alpha_{i_2^{k+1}}^{(k)} \right)^{-1} x_{i_2^{k+1}}^{(k)} a_{i_2^{k+1} j_2^k} y_{j_2^k}^{(k)} \right) \end{aligned}$$

Si on pose $d_k = \min_{i,j} r_i^{-1} \alpha_i^{(k)-1} x_i^{(k)} a_{ij} y_j^{(k)}$ on a alors la majoration

$$\varphi(x^{(k+1)}, y^{(k+1)}) \leq \left(\left(\beta_{j_1^k}^{(k)} \right)^{-1} - \left(\beta_{j_2^k}^{(k)} \right)^{-1} \right) (1 - 2d_k)$$

Posons $\rho = \min_{i,j} (r_i, c_j)$ et $R = \sum_{i,j} a_{ij}$.

– On a pour $k \in \mathbb{N}$ et $1 \leq i \leq n$:

$$\alpha_i^{(k)-1} x_i^{(k)} = x_{ik} \frac{r_i}{\sum_{j=1}^m x_i^{(k)} a_{ij} y_j^{(k)}} = \left(\sum_{j=1}^m a_{ij} y_j^{(k)} \right)^{-1} r_i$$

En utilisant le fait que $y_j^{(k)} \leq \delta^{1/2} a^{-1/2}$ et les définitions, il vient

$$\alpha_i^{(k)-1} x_i^{(k)} \geq \left(\sum_{i,j} a_{ij} \right)^{-1} r_i a^{1/2} \delta^{-1/2} \geq \rho R^{-1} a^{1/2} \delta^{-1/2}$$

– De même pour $k \in \mathbb{N}$ et $1 \leq j \leq m$:

$$y_j^{(k)} = c_j \left(\sum_{j=1}^m a_{ij} x_i^{(k)} \right)^{-1} \geq \left(\sum_{j=1}^m a_{ij} \right)^{-1} c_j a^{1/2} \delta^{-1/2} \geq \rho R^{-1} a^{1/2} \delta^{-1/2}$$

– Ainsi il vient :

$$d_k \geq \rho^2 R^{-2} a^2 \delta^{-2}$$

– En remarquant ensuite que pour tout $k \in \mathbb{N}$ et tout $1 \leq j \leq m$, on a

$$\beta_j^{(k)} = c_j^{-1} \sum_{i=1}^n (A'_k)_{ij} = c_j^{-1} \sum_{i=1}^m \alpha_i^{(k)-1} x_i^{(k)} a_{ij} y_j^{(k)}$$

Puisque $\sum_{i=1}^m x_i^{(k)} a_{ij} y_j^{(k)} = c_j$ par définition de nos suites, on a donc

$$\left(\alpha_{i_2^k}^{(k)} \right)^{-1} \leq \beta_j^{(k)} \leq \left(\alpha_{i_1^k}^{(k)} \right)^{-1}$$

Ceci étant valable pour tout j , on a donc en particulier :

$$\left(\left(\beta_{j_1^k}^{(k)} \right)^{-1} - \left(\beta_{j_2^k}^{(k)} \right)^{-1} \right) \leq \left(\alpha_{i_2^k}^{(k)} - \alpha_{i_1^k}^{(k)} \right)$$

Finalement, on a donc

$$\varphi(x^{(k+1)}, y^{(k+1)}) \leq (1 - 2\rho^2 R^{-2} a^2 \delta^{-2}) \left(\alpha_{i_2^k}^{(k)} - \alpha_{i_1^k}^{(k)} \right) = (1 - 2\rho^2 R^{-2} a^2 \delta^{-2}) \varphi(x^{(k)}, y^{(k)})$$

où $0 \leq (1 - 2\rho^2 R^{-2} a^2 \delta^{-2}) < 1$ et donc

$$\boxed{\varphi(x^{(k)}, y^{(k)}) \rightarrow 0}$$

3. Pour montrer cela, il suffit de montrer que toutes les coordonnées sont bornées par des constantes strictement positives indépendantes de k .

- Par un calcul précédent (deux premiers développements de $\alpha_{i_1^{k+1}}^{(k+1)}$), on a que pour tout $k \in \mathbb{N}$ et pour tout $1 \leq i \leq m$

$$\alpha_i^{(k+1)} = r_i^{-1} \sum_{j=1}^m \left(\alpha_i^{(k)} \right)^{-1} x_i^{(k)} a_{ij} \left(\beta_j^{(k)} \right)^{-1} y_j^{(k)}$$

Puisque

$$\sum_{j=1}^m \left(\alpha_i^{(k)} \right)^{-1} x_i^{(k)} a_{ij} y_j^{(k)} = r_i$$

on trouve l'encadrement

$$\left(\beta_{j_2^k}^{(k)} \right)^{-1} \leq \alpha_i^{(k+1)} \leq \left(\beta_{j_1^k}^{(k)} \right)^{-1}$$

En le combinant à l'encadrement

$$\left(\alpha_{i_2^k}^{(k)} \right)^{-1} \leq \beta_j^{(k)} \leq \left(\alpha_{i_1^k}^{(k)} \right)^{-1}$$

on trouve que pour tout $k \in \mathbb{N}$ et $1 \leq i \leq n$

$$\alpha_i^{(k)} \geq \alpha_{i_1^0}^{(0)}$$

Puisque précédemment on a montré la majoration

$$\alpha_i^{(k)-1} x_i^{(k)} \geq \left(\sum_{i,j} a_{ij} \right)^{-1} r_i a^{1/2} \delta^{-1/2} \geq \rho R^{-1} a^{1/2} \delta^{-1/2}$$

on a donc que pour tout $k \in \mathbb{N}$ et $1 \leq i \leq n$,

$$x_i^{(k)} \geq \alpha_{i_1}^{(0)} \rho R^{-1} a^{1/2} \delta^{-1/2} > 0$$

ce qui montre que $\boxed{\inf_i |x_i^{(k)}| \geq \alpha_{i_1}^{(0)} \rho R^{-1} a^{1/2} \delta^{-1/2} > 0}$.

- On avait montré avant que pour tout $k \in \mathbb{N}$ et $1 \leq j \leq m$

$$y_j^{(k)} \geq \rho R^{-1} a^{1/2} \delta^{-1/2}$$

ce qui montre que

$$\boxed{\forall k \in \mathbb{N}, \inf_j |y_j^{(k)}| \geq \rho R^{-1} a^{1/2} \delta^{-1/2} > 0}$$

□

Chapitre 5

Application à la diffusion de mesures de probabilité sur les sommets d'un graphe

Cette partie s'inspire fortement de l'article de Solomon [8]. Avant de considérer des applications informatiques plus fines que les précédentes, nous allons devoir revisiter la dualité de KANTOROVITCH pour esquisser la preuve résultat suivant (voir [2], thm. 2.18) :

Théorème 7. *Soit μ, ν deux mesures de probabilité sur X et Y des compacts de \mathbb{R} . On note F et G leurs fonctions de répartition. Soit γ une mesure de probabilité sur \mathbb{R}^2 telle que pour tout x, y :*

$$\gamma(] -\infty, x] \times] -\infty, y]) = \min(F(x), G(y))$$

Alors $\gamma \in \Pi(\mu, \nu)$ et γ est un plan de transport optimal pour la fonction de coût $c(x, y) = |x - y|^2$. Par conséquent :

$$W_2^2(\mu, \nu) = \int_{[0,1]} |F^{-1}(t) - G^{-1}(t)|^2 dt.$$

5.1 Monotonie c-cyclique

On commence par regarder quelques résultats intermédiaires présentés dans [1].

Définition 6 (Ensembles c-cycliquement monotones). *Soit X, Y des ensembles et $c : X \times Y \mapsto \mathbb{R} \cup \{+\infty\}$. Soit $\Gamma \subset X \times Y$. On dit que Γ est c-cycliquement monotone si pour toute famille finie $(x_1, y_1), \dots, (x_n, y_n) \in \Gamma$ et pour tout σ permutation de $[1, n]$:*

$$\sum_i c(x_i, y_i) \leq \sum_i c(x_i, y_{\sigma(i)})$$

On dit qu'un plan de transport est c-cyclique si son support est c-cycliquement monotone.

Définition 7 (Fonctionnelle c -cylique). Avec les notations de la définition précédente, pour une fonctionnelle $\Phi : X \mapsto \mathbb{R} \cup \{+\infty\}$, on appelle c -transformée de Φ la fonctionnelle donnée par :

$$\forall y \in Y, \Phi^c(y) = \inf_{x \in X} [c(x, y) - \Phi(x)]$$

Réciproquement, pour $\Psi : Y \mapsto \mathbb{R} \cup \{+\infty\}$, on appelle \bar{c} -transformée de Ψ la fonctionnelle donnée par :

$$\forall x \in X, \Phi^{\bar{c}}(x) = \inf_{y \in Y} [c(x, y) - \Psi(y)]$$

On dit que Φ est c -convexe s'il existe $\bar{\Phi}$ telle que $\Phi = \bar{\Phi}^c$, et on dit que Ψ est \bar{c} -concave s'il existe $\bar{\Psi}$ telle que $\Psi = \bar{\Psi}^{\bar{c}}$.

Nous énonçons maintenant deux propositions dont nous ne donnerons volontairement pas la preuve afin de ne pas alourdir notre propos. On trouvera des preuves détaillées dans [1], thm 1.36 et 1.37.

Proposition 6. Si $\emptyset \neq \Gamma \subset X \times Y$ est c -cycliquement monotone, et si $c : X \times Y \rightarrow \mathbb{R}$, alors on dispose d'une fonction c -concave $\Phi \neq -\infty : X \rightarrow \mathbb{R} \cup \{-\infty\}$ telle que :

$$\Gamma \subset \{(x, y) \in X \times Y \mid \Phi \oplus \Phi^c(x, y) = c(x, y)\}.$$

Proposition 7. Si γ est un plan de transport optimal et si c est continue sur $X \times Y$, alors Γ est c -cycliquement monotone.

Il en résulte un raffinement du théorème de dualité de KANTOROVITCH :

Théorème 8 (Retour sur la dualité de KANTOROVITCH). Si X et Y sont compacts :

$$\min_{\gamma \in \Pi(\mu, \nu)} = \max_{\{\Phi \text{ } c\text{-convexe} \mid \Phi \oplus \Phi^c \leq c\}} \left[\int_X \Phi d\mu + \int_Y \Phi^c d\nu \right]$$

5.2 Idée de la preuve de l'expression de la distance de Wasserstein avec les fonctions de répartition

Il faut commencer par démontrer le résultat suivant :

Proposition 8 (Critère d'optimalité de KNOTT-SMITH). Soit μ et ν des mesures de probabilité sur X et Y des compacts de \mathbb{R}^d , que l'on étend naturellement en mesures de probabilité sur \mathbb{R}^d . On considère la fonction de coût $c(x, y) = |x - y|^2$. Si un plan de transport $\gamma \in \Pi(\mu, \nu)$ de support Γ est optimal alors il existe une fonction convexe semi-continue inférieurement Φ telle que $\Gamma \subset \text{Graph}(\partial\Phi)$.

Démonstration. On sait d'après ce qui précède que l'on dispose d'un plan γ de transport optimal. On sait également que l'on dispose d'une fonction ϕ sur X telle que (ϕ, ϕ^c) réalise le sup dans le théorème de dualité de KANTOROVITCH :

$$\int_{X \times Y} (\phi \oplus \phi^* - c) d\gamma = 0$$

Or l'intégrande est positif, donc nul γ -presque partout. Le résultat découle alors d'une caractérisation du sous-différentiel (cf [2] prop. 2.4 p. 69, 2.12 p.80). \square

Preuve du théorème principal : On reprend les notations du théorème. Montrons que

$$\Gamma \subset \{(x, y) | F(x^-) \leq G(y) \text{ et } G(y^-) \leq F(x)\}.$$

En effet, si on se donne un couple (x_0, y_0) tel que $F(x_0^-) > G(y_0)$. F est croissante donc on dispose de U un voisinage de x tel que $F(x) \geq G(y_0)$ pour tout $x \in U$. Or G est croissante et continue à droite donc on dispose de V un voisinage de y_0 tel que pour tout $(x, y) \in U \times V$:

$$\min F(x), G(y) = G(y).$$

En particulier $\gamma(U \times V) = 0$ donc $(x_0, y_0) \notin \Gamma$. On peut vérifier que le membre de droite correspond à un ensemble monotone, donc Γ est inclus dans le sous-différentiel d'une fonction semi-continue inférieurement convexe. D'après le critère de KNOTT-SMITH, γ est un plan de transport optimal.

Remarquons maintenant que si λ est la mesure de Lebesgue sur \mathbb{R}^2 , $\gamma = \lambda \circ (F^{-1}, G^{-1})$. Cette égalité apparaît naturellement en regardant les boréliens de la forme $] -\infty, x[\times] -\infty, y[$.

Il en résulte que :

$$\int_{X \times Y} c d\gamma = \int_{[0,1]} c(F^{-1}(t), G^{-1}(t)) dt.$$

L'idée qui motive ce chapitre est relativement simple. Donnons-nous un graphe non orienté $G = (V, E)$ où V est l'ensemble des sommets et E l'ensemble des arêtes (qui peut représenter par exemple des villes d'un pays reliées par des routes).

Supposons que sur un sous-ensemble V_0 de V , nous ayons des informations qui peuvent s'exprimer comme distributions (par exemple, on peut prendre V_0 l'ensemble des ports et pour chaque port de V_0 , on donnerait les arrivages de blé sur l'année).

L'objectif est de proposer une approche mathématique et algorithmique à la résolution des deux problèmes suivants :

- comment déduire des informations sur les sommets de V_0 des informations sur tous les sommets, et quel modèle adopter ?
- comment faire diffuser l'information, en partant des sommets de V_0 sur tous les sommets de V ?

5.3 Éléments de calcul différentiel sur les graphes

Le but est de présenter un cadre assez général permettant de faire du calcul différentiel sur les graphes : on voit assez bien le lien avec le problème initial puisque l'un des objectifs est de faire de la "diffusion".

Pour cette partie, nous nous inspirons fortement des articles de Y.-S. Chung [10], [11] et [12], en précisant quelques définitions.

5.3.1 Premières définitions sur un graphe

Soit $G = (V, E)$ un graphe non orienté, où V est l'ensemble des sommets et $E \subset \mathcal{P}_2(V) = \{\{x, y\} : x, y \in V\}$ est l'ensemble des arêtes. On note également $V(G)$ l'ensemble des sommets de G et $E(G)$ l'ensemble de ses arêtes. On va supposer dans tout ce qui suit que G est un graphe fini ce qui simplifie notamment l'étude des fonctions sur G . Il est possible de faire cela sur des graphes dénombrables, à condition d'imposer des conditions du type L^1 sur ces fonctions.

Pour tout $(x, y) \in V^2$, on note $x \sim y$ le fait que $\{x, y\} \in E$, c'est à dire que x et y soient reliés par une arête. On se permettra également dans la suite le léger abus de notation $x \in G$ pour désigner $x \in V$. De plus, on notera indifféremment $f : G \rightarrow \mathbb{R}$ ou $f : V \rightarrow \mathbb{R}$ une application de l'ensemble des sommets de G dans \mathbb{R} .

Nous allons également nous donner une application $\omega : V \times V \rightarrow \mathbb{R}_+$ qui vérifie

- $\omega(x, y) > 0 \Leftrightarrow \{x, y\} \in E$
- $\forall (x, y) \in V^2, \omega(x, y) = \omega(y, x)$

Ainsi, $\omega(x, y)$ représente en quelque sorte la capacité à conduire de l'information entre le sommet x et le sommet y ; n'avoir pas d'arête entre un sommet x et y est équivalent à ce que l'information ne puisse pas se propager directement entre x et y . Nous appellerons ω la fonction de pondération des arêtes.

Donnons maintenant les définitions principales permettant de faire du calcul différentiel sur les graphes.

Définition 8 (poids et probabilités de passage). *Étant donné un sommet $x \in V$, on pose $d_\omega x$ le poids total du sommet x , c'est à dire*

$$d_\omega x := \sum_{y \in V} \omega(x, y) = \sum_{y \sim x} \omega(x, y)$$

Étant donné $x, y \in G$, on pose

$$p(x, y) = \begin{cases} \frac{\omega(x, y)}{d_\omega x} & \text{si } d_\omega x > 0 \\ 1 & \text{si } d_\omega x = 0 \text{ et } x = y \\ 0 & \text{sinon} \end{cases}$$

qui représente la probabilité de passage du sommet x au sommet y .

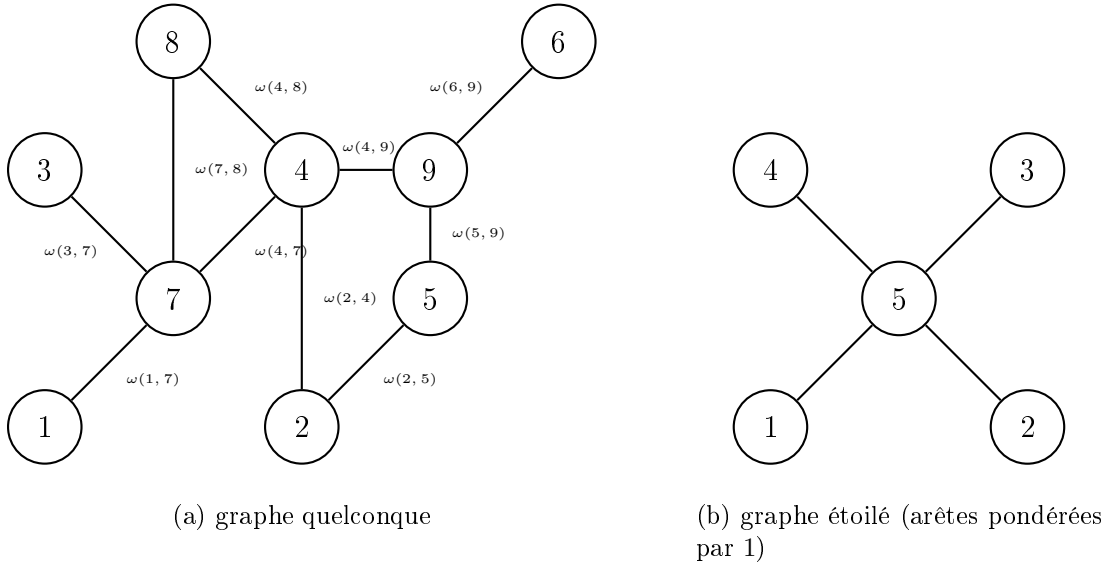


FIGURE 5.1 – Exemples de graphes G

Définition 9 (intégrale, différentielle). Soit $f : G \rightarrow \mathbb{R}$. Alors :

- on définit l'intégrale de f sur G

$$\int_G f d_\omega = \int_G f(x) d_\omega x := \sum_{x \in V} f(x) d_\omega x$$

Naturellement, on définit le volume du graphe G :

$$\text{Vol}(G) = \int_G 1 d_\omega = \sum_{x \in V} d_\omega x = \sum_{(x,y) \in V^2} \omega(x,y)$$

- on définit la dérivée directionnelle de f en $x \in G$ selon $y \in G$:

$$D_{\omega,y} f(x) = [f(y) - f(x)] \sqrt{p(x,y)}$$

Naturellement, on définit le gradient de f en $x \in V$:

$$\nabla_\omega f(x) := (D_{\omega,y}(x))_{y \in V}$$

L'expression du gradient que nous avons ci-dessus permet de définir le laplacien de

$f : G \rightarrow \mathbb{R}$ en posant

$$\begin{aligned}
\forall x \in V, \Delta_\omega f(x) &:= - \sum_{y \in V} D_{\omega,y} \circ D_{\omega,y} f(x) \\
&= \sum_{y \in V} D_{\omega,y} f(x) \sqrt{p(x,y)} \\
&= \sum_{y \in V} [f(y) - f(x)] p(x,y) \\
&= \sum_{y \sim x} [f(y) - f(x)] p(x,y)
\end{aligned}$$

5.3.2 Calcul différentiel sur un sous-graphe

Dans cette partie, nous allons fixer G un graphe qui sera en quelque sorte notre espace de départ. Nous allons maintenant faire du calcul différentiel sur des sous-graphes de G .

Rappelons que (S, i) est un sous-graphe de G si $i : V(S) \hookrightarrow V(G)$ et si $i(E(S)) = \{\{i(x), i(y)\} : \{x, y\} \in E(S)\} \subset E(G) \cap \mathcal{P}_2(V(S))$. On dit que S est un sous-graphe induit de G si $i(E(S)) = E(G) \cap \mathcal{P}_2(V(S))$. Pour simplifier les notations, on considèrera à partir de maintenant que $V(S)$ est inclus dans $V(G)$ et on omettra i .

Étant donné un sous-graphe de G , il est naturellement muni d'une fonction de pondération héritée de ω :

$$\begin{cases} \omega_S(x, y) = \omega(x, y) \text{ si } \{x, y\} \in E(S) \\ \omega_S(x, y) = 0 \text{ sinon} \end{cases}$$

D'après ce qui précède, nous disposons d'un calcul différentiel sur le graphe S . Mais étant donné $f : G \rightarrow \mathbb{R}$, $\tilde{f} : S \rightarrow \mathbb{R}$ sa restriction à S , on a

$$\int_G f \mathbf{1}_{V(S)} \frac{d_{\omega_S} x}{d_\omega x} d_\omega = \int_S \tilde{f} d_{\omega_S}$$

En général, nous n'avons donc pas

$$\int_G f \mathbf{1}_S d_\omega = \int_S \tilde{f} d_{\omega_S}$$

car la mesure point par point dépend du graphe.

La définition de l'intégrale sur un sous-graphe S de G n'est donc pas évidente ; nous allons distinguer deux cas :

– étant donné $f : G \rightarrow \mathbb{R}$, alors

$$\int_S f d_\omega := \int_G f \mathbf{1}_S d_\omega$$

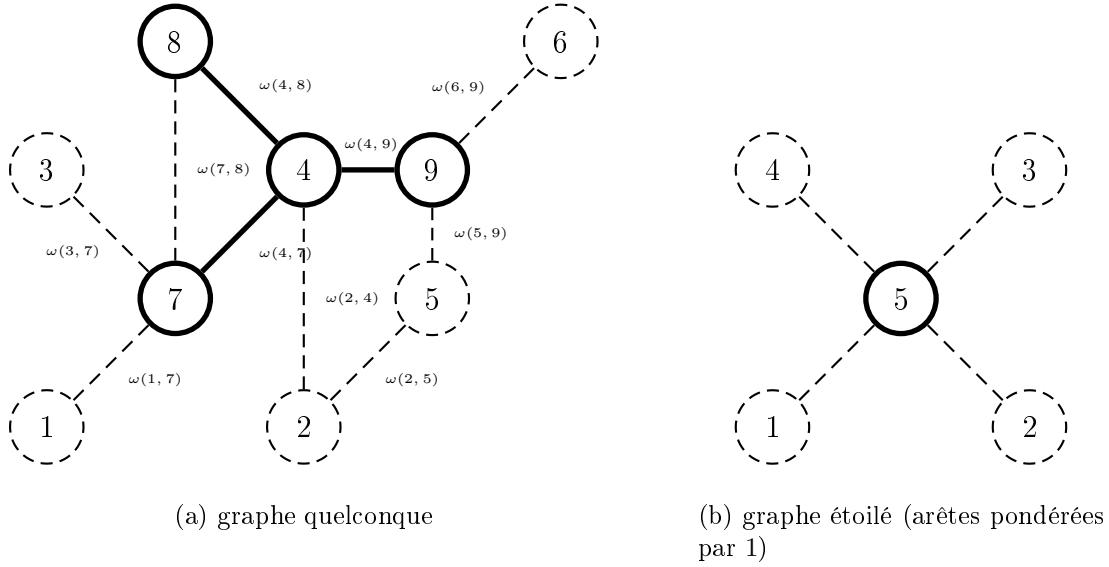


FIGURE 5.2 – Exemples de sous-graphes S (en trait plein) de G (en pointillé)

– étant donné $f : S \rightarrow \mathbb{R}$, alors

$$\int_S f d_\omega := \int_S f d_{\omega_S}$$

et plus généralement, on fera du calcul différentiel sur f en la considérant sur le graphe S muni de ω_S que l'on renote ω pour simplifier (sauf dans quelques démonstrations où on mèlera les deux).

Ainsi, préciser le graphe de départ de nos fonctions sera primordial car notre calcul différentiel en dépend (on donne le graphe de départ et non le sommet de départ car les arêtes présentes ou non jouent un rôle dans la pondération et donc le calcul différentiel).

Si nous devons faire une analogie avec le calcul différentiel classique, notre graphe initial G serait en quelque sorte notre espace de départ \mathbb{R}^d . Étant donné un sous-graphe S de G , nous allons par la suite le voir comme un ouvert Ω : l'étude d'équations sur Ω ne peut se faire sans des conditions sur son bord, que nous allons définir.

Définition 10 (bord, adhérence). *Soit S un sous-graphe de G . On définit*

– le bord de S :

$$\partial S := \{z \in V(G) \setminus V(S) : \exists x \in V(S), z \sim x\}$$

Par la suite, si on considèrera ∂S à la fois comme un sous-ensemble de $V(G)$ et comme le graphe $(\partial S, \emptyset)$.

– l'adhérence de S :

$$\bar{S} := (S \cup \partial S, E(S) \cup \{\{x, y\} \in E(G) : x \in \partial S \text{ et } y \in S\})$$

c'est à dire le graphe dont les sommets sont la réunion des sommets de S et de son bord, et dont les arêtes sont celles de S plus celles qui relient S à son bord. On omet celles qui auraient pu relier deux sommets de ∂S dans G .

Illustrons ce que représente l'adhérence : nous représentons ici \overline{S} en trait plein, S en gras et le reste du graphe G en pointillés.

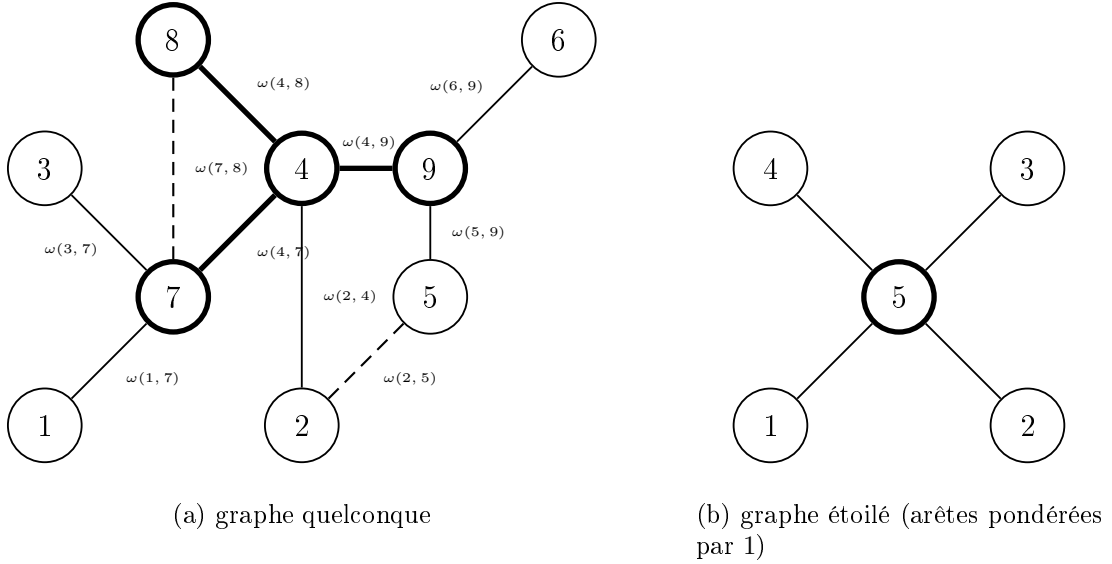


FIGURE 5.3 – Illustration des adhérences

Remarque 8. S est un sous-graphe de G mais aussi de \overline{S} . Il est même un sous-graphe induit de \overline{S} .

Le bord dépend évidemment du graphe de départ G . Dorénavant, nous fixerons G , S désignera toujours un sous-graphe de G , et \overline{S} désignera l'adhérence de S dans G .

Supposons que S soit un sous-graphe d'un graphe H , et soit $f : S \rightarrow \mathbb{R}$. On appelle extension de f à H l'application de H dans \mathbb{R} qui coïncide avec f sur S et qui est nulle ailleurs. On la note encore f .

Lemme 4. *Supposons que S soit un sous-graphe induit de G . Alors*

$$\forall x \in S, d_{\omega_{\overline{S}}}x = d_{\omega}x$$

Démonstration. Soit $x \in S$ et $y \in G$. Supposons que $\{x, y\} \in E(G)$, alors $y \in S$ ou $y \in \partial S$ et donc en particulier $y \in \overline{S}$. Si $y \in \partial S$, alors $\{x, y\} \in E(\overline{S})$ par définition. Si $y \in S$, alors puisque S est induit, $\{x, y\} \in E(S) \subset E(\overline{S})$. Dans tous les cas, $\{x, y\} \in E(\overline{S})$. Ainsi,

$$\forall y \in G, \{x, y\} \in E(\overline{S}) \Leftrightarrow \{x, y\} \in E(G)$$

et le lemme en découle □

On déduit de ce lemme les conséquences directes suivantes :

Proposition 9. *Si S est un sous-graphe induit de G et $f : \bar{S} \rightarrow \mathbb{R}$ et $g : G \rightarrow \mathbb{R}$ sont deux fonctions, alors*

- *si f et g coïncident sur \bar{S}*

$$\forall x \in S, \forall y \in \bar{S}, D_{\omega,y}f(x) = D_{\omega,y}g(x);$$

- *si f et g coïncident sur S*

$$\int_S f(x)d_\omega x = \int_S g(x)d_\omega x$$

Définissons enfin les dérivées normales sortantes, ce qui nous permettra d'obtenir des formules type formule de Green.

Définition 11 (dérivée normale sortante). *Soit G un graphe, S un sous-graphe de G , $f : G \rightarrow \mathbb{R}$ et $z \in \partial S$. On définit la dérivée normale sortante de f en z :*

$$\frac{\partial f}{\partial_\omega n}(z) = \sum_{y \in S} [f(z) - f(y)] \frac{\omega(z, y)}{d'_\omega z} = f(z) - \sum_{y \in S} f(y) \frac{\omega(y, z)}{d'_\omega z} \text{ avec } d'_\omega z = \sum_{y \in S} \omega(y, z)$$

Remarquons que si $G = \bar{S}$, alors $d'_\omega x = d_\omega x$ et donc $\Delta f(z) = -\frac{\partial f(z)}{\partial_\omega n}$. De manière générale, $d'_\omega z = d_{\omega \bar{S}} z$.

5.3.3 Formules

De ces définitions, nous déduisons des formules qui ressemblent énormément à celles que nous connaissons en analyse et qui montrent la conservations de nombreuses preuves et propriétés de ces opérateurs.

Théorème 9. *Soit G un graphe, f et g deux fonctions de G dans \mathbb{R} . Alors*

$$-2 \int_G g (\Delta_\omega f) d_\omega = \int_G \nabla_\omega g \cdot \nabla_\omega f d_\omega \quad (5.1)$$

Démonstration. Puisque pour tout $x, y \in G$, $p(x, y)d_\omega x = p(y, x)d_\omega y$, on a

$$\begin{aligned} \int_G \nabla_\omega f \cdot \nabla_\omega g d_\omega &= \sum_{x, y \in G} (f(x) - f(y))(g(x) - g(y)) \sqrt{p(x, y)} \sqrt{p(x, y)} d_\omega x \\ &= \sum_{x \in G} g(x) \sum_{y \in G} (f(x) - f(y)) p(x, y) d_\omega x + \sum_{y \in G} g(y) \sum_{x \in G} (f(y) - f(x)) p(y, x) d_\omega y \\ &= -2 \int_G g \Delta_\omega f d_\omega \end{aligned}$$

□

Ce théorème montre en particulier qu'étant donné deux fonctions $f, g : G \rightarrow \mathbb{R}$, on a

$$\int_G g \Delta_\omega f d_\omega = \int_G f \Delta_\omega g d_\omega \quad (5.2)$$

De plus, si on définit l'énergie de Dirichlet de $f : G \rightarrow \mathbb{R}$:

$$\mathcal{E}_D(f) := \frac{1}{2} \int_G |\nabla_\omega f|^2 d_\omega$$

alors d'après la formule 5.1, on a également

$$\mathcal{E}_D(f) = \int_G f (-\Delta_\omega f) d_\omega$$

Théorème 10 (laplacien du produit). *Soit G un graphe, f et g deux fonctions de G dans \mathbb{R} . Alors*

$$\Delta_\omega(fg) = f \Delta_\omega g + \nabla_\omega f \cdot \nabla_\omega g + g \Delta_\omega f \quad (5.3)$$

Démonstration. Soit $x \in G$. Alors

$$\begin{aligned} \Delta_\omega(fg)(x) &= \sum_{y \in G} [f(y)g(y) - f(x)g(x)] p(x, y) \\ &= \sum_{y \in G} [f(x)g(y) - f(x)g(x)] p(x, y) + [f(y)g(x) - f(x)g(x)] p(x, y) \\ &\quad + \sum_{y \in G} [f(y)g(y) + f(x)g(x) - f(y)g(x) - g(y)f(x)] p(x, y) \\ &= f(x) \Delta_\omega g(x) + g(x) \Delta_\omega f(x) + \sum_{y \in G} (f(y) - f(x))(g(y) - g(x)) \sqrt{p(x, y)} \sqrt{p(x, y)} \\ &= f(x) \Delta_\omega g(x) + g(x) \Delta_\omega f(x) + \nabla_\omega f(x) \cdot \nabla_\omega g(x) \end{aligned}$$

□

On en particulier, pour $f : G \rightarrow \mathbb{R}$

$$\Delta_\omega(f^2) = 2f \Delta_\omega f + |\nabla_\omega f|^2 \quad (5.4)$$

Théorème 11 (formule de Green). *Soit $f, g : \bar{S} \rightarrow \mathbb{R}$ deux fonctions. Alors on a la formule de Green :*

$$\int_S (g \Delta_\omega f - f \Delta_\omega g) d_\omega = \int_{\partial S} \left(g \frac{\partial f}{\partial_\omega n} - f \frac{\partial g}{\partial_\omega n} \right) d_\omega \quad (5.5)$$

Démonstration. D'après la formule 5.2, on a

$$\int_S (g \Delta_\omega f - f \Delta_\omega g) d_\omega + \int_{\partial S} (g \Delta_\omega f - f \Delta_\omega g) d_\omega = \int_{\bar{S}} (g \Delta_\omega f - f \Delta_\omega g) d_\omega = 0$$

Or d'après une remarque dans la définition de la dérivée normale sortante,

$$\forall z \in \partial S, \Delta_\omega f(z) = -\frac{\partial f(z)}{\partial_\omega n}$$

ce qui nous donne le résultat voulu. \square

Soit $f, g : \overline{S} \rightarrow \mathbb{R}$. En appliquant la formule de Green à fg et 1, et la formule 5.3 pour calculer $\Delta_\omega(fg)$, on trouve

$$\int_S \nabla_\omega f \cdot \nabla_\omega g + \int_S [f\Delta_\omega g + g\Delta_\omega f] = \int_{\partial S} \frac{\partial(fg)}{\partial_\omega n} \quad (5.6)$$

5.3.4 Laplacien, propriétés des fonctions harmoniques

Dans cette partie, nous allons nous intéresser à l'opérateur Laplacien et à l'existence de solutions $f : \overline{S} \rightarrow \mathbb{R}$ à un problème du type :

$$\begin{cases} \Delta_\omega f = 0 \text{ sur } S \\ f \text{ est prescrite sur } \partial S \end{cases}$$

Laplacien

Soit G un graphe, on peut voir le laplacien de G comme une matrice :

$$\Delta_\omega(x, y) := \begin{cases} p(x, y) \text{ si } x \neq y \\ p(x, x) - 1 \text{ si } x = y \end{cases}$$

en remarquant que Δ_ω peut se voir comme un endomorphisme linéaire de $\mathbb{R}^{V(G)}$. Étant donné $f : G \rightarrow \mathbb{R}$, on a donc avec ces notations

$$\forall x \in G, \Delta_\omega f(x) = \sum_{y \in G} \Delta_\omega(x, y)f(y)$$

Dans cette partie, nous allons étudier cette matrice et fixer une fois pour toute les notations liées au laplacien.

Définition 12 (matrice symétrique associée au laplacien). *Soit G un graphe tel que*

$$\forall x \in G, d_\omega x > 0$$

Posons $D = \text{diag}((\sqrt{d_\omega x})_{x \in G})$. On définit la matrice symétrique associée au laplacien :

$$\forall x, y \in G, \mathcal{L}_\omega(x, y) := \begin{cases} \frac{\omega(x, y)}{\sqrt{d_\omega x d_\omega y}} \text{ si } x \neq y \\ \frac{\omega(x, x)}{d_\omega x} - 1 \text{ sinon} \end{cases}$$

On a directement

$$\mathcal{L}_\omega = D\Delta_\omega D^{-1}$$

Nous allons maintenant nous intéresser aux valeurs propres et aux vecteurs propres de la matrice \mathcal{L}_ω . Dans la suite, étant donné deux fonctions f, g de G dans \mathbb{R} (que l'on identifie à $\mathbb{R}^{V(G)}$), on note $\langle \cdot, \cdot \rangle$ le produit scalaire canonique, c'est à dire

$$\langle f, g \rangle = \sum_{x \in G} f(x)g(x)$$

Proposition 10. *Les valeurs propres de \mathcal{L}_ω sont comprises entre -2 et 0 .*

Démonstration. Soit donc $f : G \rightarrow \mathbb{R}$. On a

$$\langle f, \mathcal{L}_\omega f \rangle = \sum_{x, y \in G} \left[f(x) \frac{\omega(x, y)}{\sqrt{d_\omega x} \sqrt{d_\omega y}} f(y) \right] - \sum_{x \in G} f(x)^2$$

Or

$$\begin{aligned} \left| \sum_{x, y \in G} \left[f(x) \frac{\omega(x, y)}{\sqrt{d_\omega x} \sqrt{d_\omega y}} f(y) \right] \right| &= \left| \sum_{x, y \in G} \left[f(x) \sqrt{\frac{\omega(x, y)}{d_\omega x}} \right] \left[f(y) \sqrt{\frac{\omega(y, x)}{d_\omega y}} \right] \right| \\ \text{(Cauchy-Schwarz)} \quad &\leq \left[\sum_{x, y \in G} f(x)^2 \frac{\omega(x, y)}{d_\omega x} \right]^{1/2} \left[\sum_{x, y \in G} f(y)^2 \frac{\omega(y, x)}{d_\omega y} \right]^{1/2} \\ &= \sum_{x \in G} f(x)^2 \end{aligned}$$

Ce qui montre que pour tout $f : G \rightarrow \mathbb{R}$, on a

$$-2\langle f, f \rangle \leq \langle \mathcal{L}_\omega f, f \rangle \leq 0$$

et donc que les valeurs propres de \mathcal{L}_ω sont comprises entre -2 et 0 . En particulier, \mathcal{L}_ω est une matrice symétrique négative. \square

En se servant de la preuve précédente, on peut décrire le noyau de \mathcal{L}_ω . En effet, par le cas d'égalité de Cauchy-Schwarz, on a

$$\mathcal{L}_\omega f = 0 \Leftrightarrow \exists k \geq 0, \forall x, y \in G, f(x) \sqrt{\frac{\omega(x, y)}{d_\omega x}} = k f(y) \sqrt{\frac{\omega(x, y)}{d_\omega y}}$$

et donc

$$\mathcal{L}_\omega f = 0 \Leftrightarrow \exists k \geq 0, \forall \{x, y\} \in E(G), \frac{f(x)}{\sqrt{d_\omega x}} = k \frac{f(y)}{\sqrt{d_\omega y}} \quad (5.7)$$

Soit f une fonction de laplacien nulle sur G . Notons G_1, \dots, G_N les composantes connexes du graphe G , et prenons une telle composante G_i .

Soit x_0 un sommet fixé de G_i , et écrivons $f(x_0) = \lambda_i \sqrt{d_\omega x_0}$.

Puisque $d_\omega x_0 > 0$, il existe $y \in G_i$ tel que $y \sim x_0$. En appliquant 5.7 à x_0 et y , et en échangeant leur rôles, on trouve que nécessairement $k = 1$ (sauf si f est nulle en x_0 auquel cas on montre facilement que f est nulle sur toute la composante connexe).

Puis pour tout voisin y de x_0 , on a d'après 5.7 que $\frac{f(y)}{d_\omega y} = \lambda_i$. Par connexité de G_i , on en déduit

$$\exists \lambda_i \in \mathbb{R}, \forall x \in G_i, f(x) = \lambda_i f_i(x)$$

où

$$\begin{aligned} f_i & : G_i & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt{d_\omega x} \end{aligned}$$

Ainsi, on trouve que

$$\text{Ker}(\mathcal{L}_\omega) \subset \bigoplus_{i=1}^N \mathbb{R} f_i$$

Il est facile de montrer que cette inclusion est en fait une égalité. On a donc

$$\text{Ker}(\mathcal{L}_\omega) = \bigoplus_{i=1}^N \mathbb{R} f_i$$

La dimension du noyau du laplacien est donc le nombre de composantes connexes du graphe G . On a donc vérifié le résultat intuitif suivant : les éléments du noyau du laplacien sont exactement les fonctions localement constantes.

Fonctions harmoniques

Nous allons maintenant énoncer quelques propriétés des fonctions harmoniques et montrer l'existence de solutions au problèmes du type

$$\begin{cases} \Delta_\omega f = 0 \text{ sur } S \\ \Delta_\omega f \text{ est prescrite sur } \partial S \end{cases} \quad (5.8)$$

Définition 13 (fonction harmonique). *Soit $f : \bar{S} \rightarrow \mathbb{R}$. On dit que f est harmonique si*

$$\forall x \in S, \Delta_\omega f(x) = 0$$

Proposition 11. *Soit S un sous-graphe connexe de G et $f : \bar{S} \rightarrow \mathbb{R}$ telle que*

$$\forall x \in S, \Delta_\omega f(x) \geq 0$$

Si f atteint son maximum sur S , alors f est constante.

Démonstration. Supposons que f atteigne son maximum en $x \in S$. On a

$$\Delta_\omega f(x) \geq 0 \Leftrightarrow \sum_{y \sim x} [f(y) - f(x)] p(x, y) \geq 0$$

Tous les termes de la somme de droite sont négatif car le maximum de f est atteint en x . Ils sont donc tous nuls, et puisque pour tout $y \sim x$, on a que $p(x, y) > 0$, pour tout $y \sim x$, $f(x) = f(y)$ et donc le maximum de f est aussi atteint en tous les voisins de x . Par connexité de S et donc de \overline{S} , la fonction est constante. \square

Remarquons que si S n'est pas connexe, on a tout de même prouvé que si $f : \overline{S} \rightarrow \mathbb{R}$ atteint son maximum en $x \in S$, alors f est constante égale à $f(x)$ sur la composante connexe de x dans \overline{S} .

Corollaire 1. *Soit S un sous-graphe connexe de G tel que toute composante connexe de S est de bord non vide. Soit $f : \overline{S} \rightarrow \mathbb{R}$ telle que*

$$\forall x \in S, \Delta_\omega f(x) \geq 0$$

Alors f atteint son maximum sur ∂S , ie

$$\forall x \in \overline{S}, f(x) \leq \max_{y \in \partial S} f(y)$$

Démonstration. Il suffit d'appliquer le maximum : si le maximum est atteint sur S , alors il est atteint sur toute une composante connexe dans \overline{S} et donc par hypothèse en au moins un point du bord. \square

Un cas particulier du corollaire précédent est le cas où on se donne G connexe, $V_0 \subset V(G)$ non vide et S le sous-graphe induit par les sommets $V(G) \setminus V_0$.

En effet, fixons C une composante connexe de S , $c \in C$ et $v_0 \in V_0$ que nous avons supposé non vide. Puisque G est connexe, il existe un chemin reliant c à v_0 ; en prenant le premier élément de ce chemin qui n'est pas dans C , celui-ci sera dans le bord de C et donc de S .

Comme conséquence de la proposition et du corollaire précédent, nous avons les deux propositions suivantes.

Proposition 12 (principe du minimum et du maximum). *Soit S un sous-graphe connexe de G et $f : \overline{S} \rightarrow \mathbb{R}$. Alors*

- i si $\forall x \in S, \Delta_\omega f(x) \geq 0$ et si f atteint son maximum sur S , f est constante ;*
- ii si $\forall x \in S, \Delta_\omega f(x) \leq 0$ et si f atteint son minimum sur S , f est constante ;*
- iii si f est harmonique et si f atteint son maximum ou son minimum sur S , f est constante.*

Proposition 13 (régularité des fonctions harmoniques). *Soit S un sous-graphe de G dont le bord de toutes les composantes connexes est non vide et $f : \overline{S} \rightarrow \mathbb{R}$. Alors*

- i si $\forall x \in S, \Delta_\omega f(x) \geq 0$,*

$$\forall x \in \overline{S}, f(x) \leq \max_{y \in \partial S} f(y)$$

ii si $\forall x \in S, \Delta_\omega f(x) \leq 0$,

$$\forall x \in \overline{S}, f(x) \geq \min_{y \in \partial S} f(y)$$

iii si f est harmonique,

$$\forall x \in \overline{S}, \min_{y \in \partial S} f(y) \leq f(x) \leq \max_{y \in \partial S} f(y)$$

Nous allons maintenant voir que la formulation variationnelle fonctionne aussi pour le calcul différentiel sur les graphes.

Théorème 12. Soit G un graphe, V_0 un sous-ensemble de $V(G)$, et $f_0 : V_0 \rightarrow \mathbb{R}$. Notons S le sous-graphe de G induit par $V(G) \setminus V_0$. Posons

$$E = \{g : G \rightarrow \mathbb{R} : g|_{V_0} = f_0\}$$

l'ensemble des fonctions sur le graphe à valeurs prescrites sur V_0 . Si f minimise l'énergie de Dirichlet parmi les fonctions dont la valeur au bord est f_0

$$\mathcal{E}_D(f) = \frac{1}{2} \int_G |\nabla_\omega f|^2 d_\omega = \min_{g \in E} \mathcal{E}_D(g)$$

alors f est solution au problème

$$\begin{cases} \Delta_\omega f = 0 \text{ sur } S \\ \forall v_0 \in V_0, f(v_0) = f_0(v_0) \end{cases}$$

Démonstration. L'application

$$\begin{aligned} \mathcal{E}_D & : \mathbb{R}^{V(G)} \longrightarrow \mathbb{R} \\ g & \longmapsto \frac{1}{2} \int_G |\nabla_\omega g|^2 d_\omega \end{aligned}$$

est une application \mathcal{C}^1 sur $\mathbb{R}^{V(G)}$. De plus, pour $x \in V(G)$, on a

$$\frac{\partial \mathcal{E}_D(f)}{\partial x} = \frac{1}{2} \sum_{y \in G} 2 [f(x) - f(y)] p(x, y) d_\omega x = -d_\omega x \Delta_\omega f(x)$$

On suppose qu'elle atteint son minimum en f sur le sous-espace vectoriel affine $(f_0 \oplus 0) + \text{Ker}(P_{\mathbb{R}^{V_0}})$. On a alors l'existence de multiplicateurs de Lagrange :

$$\exists (\lambda_x) \in \mathbb{R}^{V(G)}, \forall x \in V(G), \frac{\partial \mathcal{E}_D(f)}{\partial x} + P_{\mathbb{R}^{V_0}}(\lambda_x) = 0$$

et donc en particulier

$$\forall x \in V(G) \setminus V_0, \frac{\partial \mathcal{E}_D(f)}{\partial x} = 0$$

Pour $x \in V(G) \setminus V_0$, si $d_\omega x > 0$, cela montre directement que $\Delta_\omega f(x) = 0$ et sinon, x n'a pas de voisins et donc on a également $\Delta_\omega f(x) = 0$. \square

Ceci est une généralisation de l'énoncé initialement souhaité : si on prend S un sous-graphe de G et que l'on applique le théorème avec $G \leftarrow \overline{S}$ et $V_0 \leftarrow \partial S$, on montre que si f minimise l'énergie de Dirichlet, alors f est solution du problème 5.8. En fait, la seule information qui compte vraiment est celle que l'on a sur ∂S .

Proposition 14. *Reprenons les mêmes notations que dans le théorème précédent.*

i Si S est connexe et si $\partial S \neq \emptyset$, alors il existe une unique solution au problème

$$\begin{cases} \Delta_\omega f = 0 \text{ sur } S \\ f = f_0 \text{ sur } V_0 \end{cases} \quad (5.9)$$

ii Si S est connexe et si $\partial S = \emptyset$, alors l'ensemble des solutions au problème 5.9 est l'ensemble des fonctions constantes.

iii L'ensemble des solutions au problème 5.9 est un espace vectoriel affine de dimension N , où N est le nombre de composantes connexes à bord vide de S . En particulier, si toutes les composantes connexes de S sont à bord non vide, la solution est unique.

Démonstration. • Pour l'existence d'un minimiseur de l'énergie de Dirichlet, il suffit de remarquer que celle-ci est strictement convexe si G est connexe et a comme limite $+\infty$ en $+\infty$. Sinon, on décompose G en composantes connexes, on remarque que l'énergie de Dirichlet de G est la somme des énergies de Dirichlet de ses composantes connexes, et on minimise chacune d'entre elles.

• L'unicité est une application triviale du principe du maximum : pour le premier point, on se donne deux solutions et on s'intéresse à leur différence, qui est harmonique à l'intérieure et nulle sur au bord donc nulle ; pour le deuxième, on a déjà montré que l'ensemble des fonctions de laplacien nul sur un graphe connexe est l'ensemble des fonctions constantes. □

Nous avons ici tout pour mettre en place un algorithme qui, étant donné un graphe G , un sous ensemble V_0 non vide de $V(G)$, une fonction $f_0 \in \mathbb{R}^{V_0}$, nous fournit l'ensemble des solutions au problème 5.9.

Pour simplifier, supposons que les sommets sont numérotés sous la forme $\underbrace{1, \dots, m}_{V_0}, \underbrace{m+1, \dots, n}_{V(G) \setminus V_0}$.

On peut alors voir f_0 comme un vecteur $Y_0 \in \mathbb{R}^m$, et un élément f de E (l'ensemble des fonctions dont la restriction à V_0 vaut f_0) comme un vecteur de la forme

$$f = \begin{pmatrix} Y_0 \\ X \end{pmatrix} \text{ avec } X \in \mathbb{R}^{n-m}$$

Si on décompose la matrice du laplacien sur G sous la forme

$$\Delta_\omega = \begin{pmatrix} \Delta_\omega^{V_0} & B \\ C & \Delta_\omega^{V \setminus V_0} \end{pmatrix}$$

Ainsi, on a

$$\Delta_\omega f = \begin{pmatrix} \Delta_\omega Y_0 \\ CY_0 \end{pmatrix} + \begin{pmatrix} BX \\ \Delta_\omega^{V \setminus V_0} X \end{pmatrix}$$

Trouver f de laplacien nul sur S (défini dans les théorèmes précédents) revient donc à résoudre le système

$$\Delta_\omega^{V \setminus V_0} X = -CY_0$$

5.4 Harmonisation de distributions de probabilités dans un graphe

Dans cette partie, nous allons fixer une bonne fois pour toute un graphe pondéré $G = (V, E, \omega)$, un sous-ensemble de sommets non vide $V_0 \subset V(G)$. Nous allons encore noter S le sous-graphe induit par les sommets $V(G) \setminus V_0$, et nous allons de plus supposer que toute composante connexe de S a un bord non vide (rappelons que c'est directement le cas si G est lui même connexe).

On note $\text{Prob}(\mathbb{R})$ l'ensemble des mesures de probabilité sur \mathbb{R} . Au lieu d'associer un réel à un sommet du graphe, on va dans toute la suite lui associer une mesure de probabilité. Ainsi, le problème devient : étant donné des mesures de probabilité μ_v pour $v \in V_0$, comment définir des mesures μ_x pour tout $x \in G$ de manière "harmonique" ?

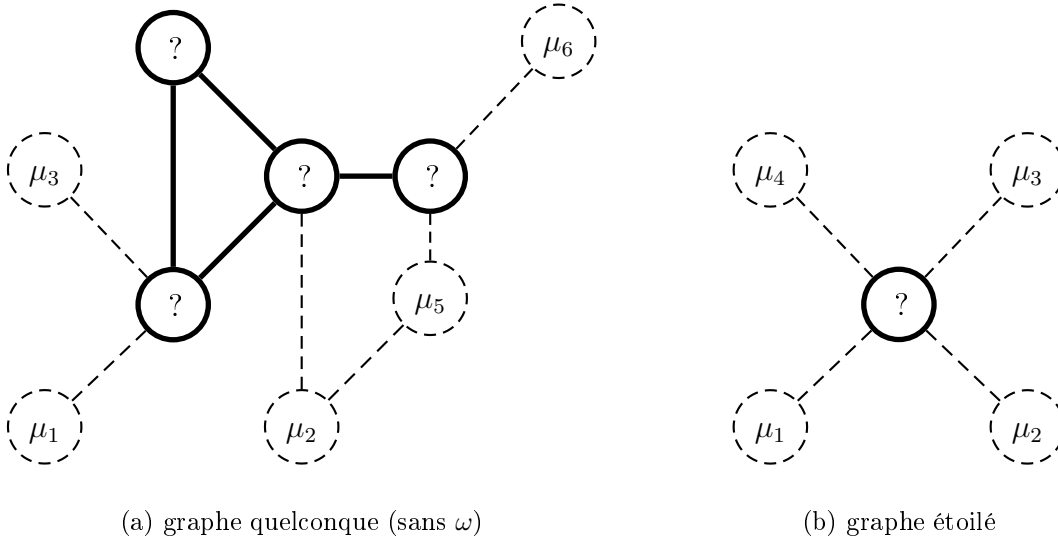


FIGURE 5.4 – Illustration de l'objectif : trouver les ? harmoniquement

La première approche intuitive est d'essayer de se ramener au problème précédent. Supposons que nos mesures μ_v pour $v \in V_0$ soient des densités ρ_v par rapport à la mesure de LEBESGUE. On pourrait, pour tout $x \in \mathbb{R}$, définir $(\rho_v(x))$ comme solution au problème de Dirichlet dans avec $(\rho_v(x))_{v \in V_0}$ comme conditions au bord (problème réel de la partie précédent), ce qui définirait les densités ρ_v point par point. Le problème

de cette approche est le même que celui de l'interpolation eulérienne des mesures : si on l'applique à un graphe où l'inconnu se situe entre deux mesures δ_0 et δ_1 , alors cette interpolation va mener à une mesure de la forme $t(\delta_0 + \delta_1)$ alors que l'on voudrait plutôt, comme dans l'interpolation de MCCANN une mesure de la forme $t\delta_{0.5}$.

C'est pour cela que l'on va plutôt utiliser des distances de WASSERSTEIN pour comparer nos mesures. Plus précisément, nous allons définir l'équivalent de l'énergie de Dirichlet et essayer de la minimiser.

Définition 14 (énergie de Dirichlet). *Soit $\mu : G \rightarrow \text{Prob}(\mathbb{R})$ une application qui à chaque sommet x de G associe une mesure de probabilité μ_x . On définit l'énergie de Dirichlet de μ :*

$$\mathcal{E}_D(\mu) = \frac{1}{2} \sum_{x,y \in G} W_2^2(\mu_x, \mu_y) p(x, y) d_\omega x = \frac{1}{2} \sum_{x,y \in G} \omega(x, y) W_2^2(\mu_x, \mu_y)$$

Rappelons quelques définitions et propriétés propres aux distributions de probabilité sur \mathbb{R} .

Étant donné une mesure de probabilité $\mu : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$, on appelle fonction de répartition de μ l'application

$$\begin{aligned} F &: \mathbb{R} \longrightarrow [0, 1] \\ x &\longmapsto \mu([-\infty, x]) \end{aligned}$$

Elle est croissante et continue à droite. On appelle pseudo-inverse continu à droite l'application

$$\begin{aligned} F^{-1} &:]0, 1[\longrightarrow \mathbb{R} \\ s &\longmapsto \inf \{x \in \mathbb{R}, F(x) > s\} \end{aligned}$$

qui est comme son nom l'indique continue à droite, et croissante. De plus, μ est la mesure image de la mesure de lebesgue sur $]0, 1[$ par la fonction F^{-1} . Il y a donc correspondance entre

- les mesures de probabilités sur \mathbb{R} ;
- les fonctions de $]0, 1[$ dans \mathbb{R} continues à droite et croissantes.

Par exemple, donnons-nous une probabilité atomique à support fini $\mu = \sum_{i=1}^n p_i \delta_{x_i}$ avec $x_1 < \dots < x_n$, $p_i > 0$ et $\sum_{i=1}^n p_i = 1$. Si on note, pour tout $0 \leq i \leq n$, $s_i = \sum_{j=1}^i p_j$, alors $F^{-1}(s) = x_i$ pour tout $s \in [s_{i-1}, s_i[$.

Réciproquement, si on se donne une fonction en escalier croissante continue à droite de $]0, 1[$ dans \mathbb{R} , il est facile de retrouver la mesure atomique dont elle est le pseudo-inverse continu à droite.

Rappelons enfin le résultat montré dans une section précédente.

Théorème 13 (VILLANI). *Soit $\mu_0, \mu_1 \in \text{Prob}(\mathbb{R})$ et soit F_0^{-1} et F_1^{-1} leur pseudo-inverses. Alors*

$$W_2^2(\mu_0, \mu_1) = \int_0^1 (F_1^{-1}(s) - F_0^{-1}(s))^2 ds$$

Nous allons donc maintenant essayer de reformuler notre problème. On a donc toujours une mesure μ_v pour tout $v \in V_0$. Pour tout $v \in V_0$, on note F_v^{-1} le pseudo-inverse continu à droite de μ_v . Pour tout $s \in]0, 1[$, on note g_s la solution au problème de Dirichlet classique 5.9 :

$$\begin{cases} \Delta_\omega g_s = 0 & \text{sur } S \\ g_s(v) = F_v^{-1}(s) & \text{sur } V_0 \end{cases}$$

c'est le minimiseur de l'énergie de Dirichlet associée.

On a alors le lemme suivant :

Lemme 5. *Pour tout $x \in G$, $s \mapsto g_s(x)$ est croissante et continue à droite.*

Démonstration. Soit $0 < s < s' < 1$ et soit $x \in G$. Puisque $\Delta_\omega(g_{s'} - g_s) = 0$ sur S , par la proposition 13 (propriété des fonctions harmoniques)

$$\min_{v \in V_0} (g_{s'}(v) - g_s(v)) \leq (g_{s'} - g_s)(x) \leq \max_{v \in V_0} (g_{s'}(v) - g_s(v))$$

et donc

$$\min_{v \in V_0} (F_v^{-1}(s') - F_v^{-1}(s)) \leq (g_{s'} - g_s)(x) \leq \max_{v \in V_0} (F_v^{-1}(s') - F_v^{-1}(s))$$

Cela montre directement à la fois la croissance et la continuité à droite des g_s . □

On peut donc définir, pour tout $x \in G$, μ_x comme la mesure image par $s \mapsto g_s(x)$ de la mesure uniforme sur $]0, 1[$ (i.e. la mesure de LEBESGUE) : $s \mapsto g_s(x)$ est alors le pseudo-inverse de cette mesure.

On a alors le théorème fondamental suivant :

Théorème 14. *L'application $\mu : G \rightarrow \text{Prob}(\mathbb{R})$ que nous venons de définir minimise l'énergie de Dirichlet.*

Démonstration. On note χ l'ensemble des fonctions $g : G \times]0, 1[\rightarrow \mathbb{R}$ telles que pour tout $x \in G$, $s \rightarrow g(x, s)$ soit $\mathcal{B}(]0, 1[)$ -mesurable et telles que pour tout $s \in]0, 1[$ et pour tout $v \in V_0$, $g(v, s) = F_v^{-1}(s)$.

Considérons le problème de minimisation

$$\min_{g \in \chi} \hat{\mathcal{E}}_D(g) := \frac{1}{2} \sum_{x, y \in G} \omega(x, y) \int_0^1 (g(x, s) - g(y, s))^2 ds$$

On sait que, par définition de nos $g_s(x)$ comme solution du problème de Dirichlet à s fixé (et donc comme minimiseur de l'énergie de Dirichlet associée), que pour tout $s \in]0, 1[$ et pour tout $g \in \chi$

$$\frac{1}{2} \sum_{x, y \in E} \omega(x, y) (g_s(x) - g_s(y))^2 \leq \frac{1}{2} \sum_{x, y \in G} \omega(x, y) (g(x, s) - g(y, s))^2$$

Si bien qu'en intégrant par rapport à s , on trouve que pour tout $g \in \chi$,

$$\frac{1}{2} \sum_{x,y \in G} \omega(x,y) \int_0^1 (g_s(x) - g_s(y))^2 ds \leq \frac{1}{2} \sum_{x,y \in G} \omega(x,y) \int_0^1 (g(x,s) - g(y,s))^2 ds$$

et donc

$$\frac{1}{2} \sum_{x,y \in G} \omega(x,y) \int_0^1 (g_s(x) - g_s(y))^2 ds \leq \min_{g \in \chi} \hat{\mathcal{E}}_D(g)$$

Or le théorème de VILLANI énoncé susdite ment montre d'une part que

$$\frac{1}{2} \sum_{x,y \in G} \omega(x,y) \int_0^1 (g_s(x) - g_s(y))^2 ds = \frac{1}{2} \sum_{x,y \in G} \omega(x,y) W_2^2(\mu_x, \mu_y)$$

et d'autre part que

$$\min_{g \in \chi} \hat{\mathcal{E}}_D(g) \leq \min_{\mu} \mathcal{E}_D(\mu)$$

où μ vérifie les conditions sur le bord. On a donc montré que μ que nous avons construit est un minimiseur de l'énergie de Dirichlet. \square

Nous avons donc mis au point une méthode pour harmoniser les mesures inconnues en fonction de celles au bord. De plus, on voit que la mise en place algorithmique est évidente : puisque l'on sait résoudre le problème de Dirichlet classique, il suffit de s'y ramener en calculant les pseudo-inverses, ce qui d'après une remarque précédente est facile pour des mesures atomiques, et qui l'est également pour les mesures à densité connue.

Mais cette harmonisation en est-elle vraiment une ? La proposition suivante montre que notre méthode ne déforme pas trop les mesures ; ni les espérances, ni les variances n'explorent.

Proposition 15. *On a :*

- si pour tout $v \in V_0$, $t \mapsto t$ est μ_v -intégrable, alors pour tout $x \in G$, $t \mapsto t$ est μ_x -intégrable et si $m(x)$ désigne l'espérance de μ_x pour tout $x \in G$, on a

$$\inf_{v \in V_0} m(v) \leq m(x) \leq \sup_{v \in V_0} m(v)$$

- de même, si pour tout $v \in V_0$, μ_v admet un moment d'ordre deux, alors pour tout $x \in G$, μ_x admet un moment d'ordre deux et si $\sigma^2(x)$ désigne la variance de μ_x pour tout $x \in G$, on a

$$0 \leq \sigma^2(x) \leq \max_{v \in V_0} \sigma^2(v)$$

Démonstration. Rappelons tout d'abord le fait suivant : soit μ une mesure sur les boréliens de \mathbb{R} et F^{-1} son pseudo inverse. Alors pour toute fonction φ positive ou μ intégrable, on a

$$\int_{\mathbb{R}} \varphi(t) d\mu = \int_0^1 \varphi(F^{-1}(s)) ds \quad (5.10)$$

- On a

$$\forall s \in]0, 1[, \forall x \in G, \min_{v \in V_0} g_s(v) \leq g_s(x) \leq \max_{v \in V_0} g_s(v)$$

ce qui montre que

$$\forall s \in]0, 1[, \forall x \in G, |g_s(x)| \leq \max_{v \in V_0} |g_s(v)| \leq \sum_{v \in V_0} |g_s(v)|$$

et donc

$$\forall x \in G, \int_0^1 |g_s(x)| ds \leq \sum_{v \in V_0} \int_0^1 |g_s(v)| ds$$

Or d'après 5.10, cela signifie exactement

$$\forall x \in G, \int_{\mathbb{R}} |t| d\mu_x \leq \sum_{v \in V_0} \int_{\mathbb{R}} |t| d\mu_v$$

donc $t \mapsto t$ est bien μ_x intégrable pour tout $x \in G$.

Puis si on pose

$$\begin{aligned} m &: G \longrightarrow \mathbb{R} \\ x &\longmapsto \int_{\mathbb{R}} t d\mu_x = \int_0^1 g_s(x) ds \end{aligned}$$

on a

$$\forall x \in S, \Delta_\omega m(x) = \int_0^1 \underbrace{\Delta_\omega g_s(x)}_{=0} ds = 0$$

Ce qui montre bien, d'après la proposition 13

$$\forall x \in G, \min_{v \in V_0} m(v) \leq m(x) \leq \max_{v \in V_0} m(v)$$

- Toujours d'après 5.10, on a

$$\forall x \in G, \sigma^2(x) = \int_0^1 [g_s(x) - m(x)]^2 ds$$

Or d'après la formule 5.4, on a, pour tout $s \in]0, 1[$,

$$\forall x \in S, \Delta_\omega([g_s(x) - m(x)]^2) = 2(g_s(x) - m(x)) \underbrace{\Delta_\omega(g_s(x) - m(x))}_{=0} + |\nabla_\omega(g_s(x) - m(x))|^2$$

et donc

$$\forall x \in S, \Delta_\omega([g_s(x) - m(x)]^2) \geq 0$$

En intégrant par rapport à s , on trouve donc

$$\forall x \in S, \Delta_\omega \sigma^2(x) \geq 0$$

et donc d'après la proposition 13, on a

$$\forall x \in G, \sigma^2(x) \leq \max_{v \in V_0} \sigma^2(v)$$

ce qui montre à la fois l'existence des moments d'ordre deux et l'encadrement souhaité. □

Remarque 9. *Si les mesures sont atomiques, alors tous les barycentres calculés seront également des mesures atomiques. Il est facile de vérifier que l'on a généralisé la notion de barycentre de WASSERSTEIN dans \mathbb{R} (il suffit de prendre un graphe étoilé).*

Remarque 10. *On voit facilement comment réaliser un algorithme permettant de calculer les mesures manquantes : puisque l'on a exposé dans la partie précédente un algorithme permettant de calculer les solutions au problème de Dirichlet, il suffit de pouvoir calculer les pseudo-inverses, ce qui est facile dans les deux cas suivants :*

- pour les mesures atomiques comme expliqué plus haut ;
- de manière approchée pour les mesures à densité.

5.5 Diffusion sur les graphes

Avec le calcul différentiel que nous avons mis en place, l'article [12] prouve l'existence et l'unicité de solutions à des problèmes de diffusion du type suivant :

$$\begin{cases} \partial_t F(x, t) - \Delta_\omega F(x, t) = H(x, t) \text{ pour } x \in S \\ F(x, 0) = f(x) \\ F(v, t) = g(v, t) \text{ pour } v \in \partial S \end{cases}$$

pourvu que H , f et g aient une certaine régularité.

En utilisant le même type de méthodes que dans la section précédente, il est possible de faire cela avec des mesures : on peut faire diffuser des mesures sur un graphe en ajoutant une composante temporelle. La principale difficulté de formalisation vient que la régularité en temps imposée sur les pseudo-inverses continus à droite est trop forte : il est cependant possible d'affaiblir un peu ces hypothèses pour, à défaut d'avoir unicité des solutions au problème précédent, on ait existence.

Conclusion

Nous avons donc vu ici deux applications de la théorie du transport optimal et de son premier résultat fondamental, la dualité de KANTOROVICH. Ces deux applications tirent profit du fait que le transport optimal, et plus précisément la distance de Wasserstein conserve de nombreuses propriétés géométriques : dans le premier cas celles des images, et dans le second celles des distributions de probabilité. Mais le gros inconvénient de l'utilisation de notions de transport optimal à des fins algorithmiques est le calcul des plans de transport optimal ainsi que des potentiels de KANTOROVICH : dans nos deux problèmes, il a fallu trouver un moyen d'accélérer la convergence des algorithmes, que ce soit par la régularisation entropique et l'algorithme de SINKHORN dans le premier cas où la restriction à des distributions de probabilité sur \mathbb{R} et l'astuce du pseudo-inverse dans le second.

Bibliographie

- [1] Filippo Santambrogio, *Optimal Transport for Applied Mathematicians*, Birkhauser 2015.
- [2] Cédric Villani, *Topics in Optimal Transportation*, Graduate studies in Mathematics vol. 58, American Mathematical Society, 2000.
- [3] Cédric Villani, *Optimal Transport, old and new*, Springer, 2009.
- [4] David G. Luenberger, Yinyu Ye, *Linear and Nonlinear Programming*, Springer, third edition, 2008.
- [5] Bertsimas et Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, 1997.
- [6] Bertrand Maury, *Modélisation*, polycopié du cours de modélisation dispensé à l'ENS, 2016.
- [7] Marco Cuturi, Arnaud Doucet, *Fast Computation of Wasserstein Barycenters*, arxiv 2014.
- [8] J. Solomon, R. Rostamov, L. Guibas, A. Butscher, *Wasserstein propagation for semi-supervised learning*, arxiv, 2014.
- [9] Richard Sinkhorn, *Diagonal equivalence to matrices with prescribed row and column sums*, The American Mathematical Monthly, 74(4) : 402-405, 1967.
- [10] S.-Y. Chung and C. A. Berenstein, *ω -Harmonic functions and inverse conductivity problems on networks*, SIAM J. Appl. Math. 65 (2005), no. 4, 1200–1226 (electronic).
- [11] F. Chung and S.-T. Yau, *Discrete Green's functions*, J. Combin. Theory Ser. A 91 (2000), no. 1-2, 191–214.
- [12] S.-Y. Chung, y.-S. Chung and J.-H. Kim, *Diffusion and Elastic Equations on Networks*, RIMS, Kyoto Univ. 43 (2007), 699–726.