Transport Optimal Computationnel

D'après Gabriel Peyré

19 novembre 2025

Table des matières

Résumé

`mailto:gabriel.peyre@ens.fr` Notes de cours sur `https://arxiv.org/abs/2505.06589` Syllabus : `https://docs.google.com/document/u/0/d/1JlDpcS0tkzX8CSgHlUf13ZHQRWu650EtNLrycT39dxk/mobilebasic`

Introduction

L'une des motivations principales du cours est de comparer, en apprentissage statistique, des données sous formes de distributions de probabilités, souvent discrètes (nuages de points). On peut penser au transport optimal comme de l'apprentissage non-supervisé : comment associer une distribution de probabilité paramétrique α_θ à une probabilité observée β sur un groupe de points. Dans ce cours, les lettres grecques sont réservées aux distributions de probabilités et les lettres latines aux points. Pour ce faire, on va faire une association de densité $\min_\theta D(\alpha_\theta, \beta)$ qui prend en compte une métrique d. L'idée étant d'utiliser la métrique d pour définir la métrique D, en utilisant la structure de l'espace sous-jacent pour les données. Il faut voir le transport optimal comme un mécanisme d'élévation de l'espace des données vers un espace de probabilité, de sorte que $D = d$ lorsqu'on considère des diracs.

1 Problème de Monge

1.1 Formulation de Monge

Le problème est a été défini pour le but militaire de construire des murs avec des sacs de sables déplacés par des soldats, par Gaspard Monge, dans un papier à l'académie des sciences.

Commençons par un exemple : si on part d'un ensemble de 3 points $x_1, \dots, x_3$ et qu'on veut atteindre $y_1, \dots, y_3$, quelle est la manière optimale de donner une bijection dont le coût du déplacement est minimal ? On cherche $T : x_i \to y_{\sigma(i)}$ pour $\sigma \in \mathfrak{S}_3$.

Définition 1.1 Le problème de Monge est le problème d'optimisation suivant :

$$M = \min_{\sigma \in \mathfrak{S}_n} \sum_{i=1}^{n} C_{i,\sigma(i)} \tag{Monge}$$

où $C \in \mathbb{R}^{n \times n}$. On dit que M est l'assignation optimale.

En général, on définit $C_{i,j} = c(x_i, y_j)$, et c sera généralement la distance géodésique sur une variété, ou une puissance de la norme p sur $\mathbb{R}^d$. Il n'y a pas besoin de supposer que x et y appartiennent aux mêmes espaces $\mathcal{X}$ et $\mathcal{Y}$.

Il est clair que le problème de Monge est combinatoirement complexe, puisque $|\mathfrak{S}_n| = n!$. Monge, historiquement, a proposé des liens avec l'optique. L'académie des sciences a proposé un prix a celui qui réussirait a proposer une solution (entendre de nos jours, algorithme polynomial pour la réponse). Aujourd'hui, nous avons un algorithme en $\mathcal{O}(n^3)$, qui est optimal dans le pire cas.

1.1.1 Zoologie de Sous-Problèmes

Dans cette section, on va s'intéresser à quelques sous-problèmes spécifiques, dont le résultat est calculable.

Cas 1-D Ici, on suppose que $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, et que $c(x, y) = h(x - y)$ pour h convexe. En général, on aura $c(x, y) = |x - y|^p$ avec $p \geq 1$. Monge avait étudié le cas $p = 1$ qui est de loin le plus difficile.

Dans ce cas, la solution est l'arrangement croissant : l'application T définit ci-dessus est croissante. On trie les points de gauche à droite et on assigne le plus petit x_i au plus petit y_i et ainsi de suite :

Théorème 1.1 Dans le cas 1-D, si M vérifie :

$$x_i < x_j \Rightarrow y_{M(i)} < y_{M(j)}$$

alors M est une solution de (Monge).

Démonstration. Si la propriété n'est pas vérifiée, alors il existe (i, i') tels que $(x_i - x_{i'})(y_{\sigma(i)} - y_{\sigma(i')}) < 0$ et en composant σ par la transposition $\tau_{i,i'}$ on obtient :

$$h(x_i - y_{\tau(\sigma(i))}) + h(x_{i'} - y_{\tau(\sigma(i'))}) < h(x_i - y_{(\sigma(i)}) + h(x_{i'} - y_{\sigma(i')})$$

par convexité de h. ■

Corollaire 1.1 Dans ce cas, on a une complexité en $\mathcal{O}(n \log n)$, en utilisant un algorithme de tri.

R Dans le cas où h est strictement convexe, tous les arrangements optimaux sont croissants, et donc on a l'unicité de la solution M dans le cas où tous les points sont distincts.

Ce cas intervient notamment dans le cas ou par exemple on veut comparer deux groupes de niveaux, ou l'égalisation en niveaux de gris de deux histogrammes de luminance (balance des blancs).

L'algorithme ne se généralise pas en deux dimensions, puisque les trajectoires ne peuvent pas se recouper pour un arrangement optimal (par l'inégalité du parallélogramme) mais que cette propriété n'implique pas l'optimalité de la solution. Pire, il peut exister un nombre exponentiel de solution non-optimales dans ce cas.

Couplage Dans le cas où on peut modéliser le problème par un problème de couplage de graphes, on peut utiliser l'algorithme hongrois pour obtenir une solution en $\mathcal{O}(n^3)$.

1.2 Formulation de Monge Continue

Ici, on va s'intéresser à des mesures boréliennes $\alpha \in \mathcal{M}(\mathcal{X}), \beta \in \mathcal{M}(\mathcal{Y})$, pour $\mathcal{X}, \mathcal{Y}$ des espaces métriques. Pour des questions de facilité, on supposera $\mathcal{X}, \mathcal{Y}$ compacts, et pour l'optimisation on les supposera de plus complets, généralement des parties de $\mathbb{R}^d$. Lorsqu'on aura besoin de retirer l'hypothèse de compacité, on ne demandera de $\mathcal{X}$ et $\mathcal{Y}$ que d'être des espaces polonais. On suppose de plus que α et β sont des mesures de probabilité (positives et de somme 1). On notera l'espace de mesures de probabilité $\mathcal{M}_+^1(\mathcal{X}) = \mathcal{P}(\mathcal{X})$.

Théorème 1.2 — Riesz-Markov Si $\mathcal{X}$ est un espace polonais, et $\mathcal{C}_c(\mathcal{X})$ est l'ensemble des fonctions continues à support compact sur $\mathcal{X}$ muni de la norme $\|\cdot\|_\infty$:

$$\mathcal{M}(\mathcal{X}) = (\mathcal{C}_c(\mathcal{X}))^\circ$$

Démonstration. On utilise pour identification le produit scalaire : $\langle f, \alpha \rangle = \int f \, d\alpha$. ■

C'est une généralisation du Théorème de Riesz sur un espace de Hilbert. Quelques exemples :

Discrète $\alpha = \sum a_i \delta_{x_i}$ où $a \in \Delta_n$. On appellera souvent le vecteur a l'histogramme, et le vecteur (δ) le nuage de points. Par linéarité, on a :

$$\int f \, \mathrm{d}\left[\sum a_i \delta_{x_i}\right] = \sum a_i f(x_i)$$

Continue/À densité Si $\alpha = \rho_\alpha \beta$ a densité ρ_α par rapport à β :

$$\int f(x) \, \mathrm{d}\alpha(x) = \int f(x) \rho_\alpha(x) \, \mathrm{d}\beta(x)$$

C'est-à-dire : $\langle \cdot, \alpha \rangle = \langle \cdot \times \rho_\alpha, \beta \rangle$. On notera ceci $\alpha << \beta$.

Le produit scalaire défini ci-dessus induit une norme duale dite norme de variation totale :

$$\|\alpha\|_\star = \|\alpha\|_{TV} = \sup_{\|f\|_\infty \leq 1} \langle f, \alpha \rangle$$

Proposition 1.1 Pour $\alpha, \beta \in \mathcal{M}(\mathcal{X})$:

$$\|\alpha - \beta\|_{TV} = |\alpha - \beta|(\mathcal{X}) = \int_{\mathcal{X}} d(|\alpha - \beta|)(x)$$

Quelques exemples :

Discrète Si $\alpha = \sum a_i \delta_{x_i}$ et $\beta = \sum b_i \delta_{x_i}$ où $a, b \in \Delta_n$ (ici, on suppose que les nuages de points sont égaux, mais pas que les a_i, b_i sont non nuls) :

$$\|\alpha - \beta\|_{TV} = \sum_i |x_i - y_i|$$

Continue/Denses Si $\alpha = \rho_\alpha \mathcal{L}$ et $\beta = \rho_\beta \mathcal{L}$:

$$\|\alpha - \beta\|_{TV} = \int |\rho_\alpha - \rho_\beta| \, \mathrm{d}\mathcal{L}$$

Mesures Singulières Si $\alpha = 0$ quand $\beta \neq 0$ et réciproquement, $\|\alpha - \beta\|_{TV} = 2$

Définition 1.2 — Push-Forward. Si on a une application de transport $T : \mathcal{X} \to \mathcal{Y}$, on peut définir une application $T_\sharp : \mathbb{P}(\mathcal{X}) \to \mathbb{P}(\mathcal{Y})$ définie par :

1. $T_\sharp \delta_x = \delta_{T(x)}$;
2. $T_\sharp$ est linéaire.

De manière équivalente : $\beta = T_\sharp \alpha$ qui vérifie $\beta(B) = \alpha(T^{-1}(B))$.

Proposition 1.2 Si $\beta = T_\sharp \alpha$, on a :

$$\int g \, \mathrm{d}\beta = \int g \circ T \, \mathrm{d}\alpha, \text{i.e. } \mathbb{E}(g(Y)) = \mathbb{E}(g(T(X)))$$

pour $Y = T(X)\tilde{\beta}$ un vecteur aléatoire.

On peut alors réécrire le problème de Monge dans le cas continu :

Définition 1.3 Si $\alpha \in \mathbb{P}(\mathcal{X}), \beta \in \mathbb{P}(\mathcal{Y})$, et c est une fonction de coût :

$$M = \inf_{T : \mathcal{X} \to \mathcal{Y}} \left\{ \int c(x, T(x)) \, \mathrm{d}\alpha(x) \,\middle|\, T_\sharp \alpha = \beta \right\} \tag{Monge}$$

Proposition 1.3 Dans le cas discret : $T_\sharp\alpha = \beta \Leftrightarrow \exists\sigma \in \mathfrak{S}_n, T(x_i) = y_{\sigma(i)}$.

Théorème 1.3 Si α a une densité par rapport à β, alors il existe T telle que $T_\sharp\alpha = \beta$.

R Dans le cas discret, cela revient à dire que $|supp(\alpha)| \geq |supp(\beta)|$. Attention, cette application T n'est pas unique !

Théorème 1.4 — Brenier, 1991 Si $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ et $c(x, y) = \|x - y\|^2$ (généralisé à une distance géodésique mis à la puissance $1 < p < +\infty$) et si α a une densité par rapport à la mesure de Lebesgue (ou au moins ne donne pas de mesure aux espaces de dimension $< d$) alors il y a un unique transport optimal T solution de (Monge) et T est caractérisé par $T = \nabla\varphi$ où φ est convexe (et ∇ est le gradient riemannien). Il y a un unique $\nabla\varphi$ où φ est convexe et $(\nabla\varphi)_\sharp\alpha = \beta$, c'est le transport optimal.

Corollaire 1.2 Les gradients des fonctions convexes $\mathcal{G}$ sont des transports optimaux.

Proposition 1.4 Pour $T \in \mathcal{G}$, $\langle Tx - Ty, x - y\rangle \geq 0$, les gradients de fonctions convexes sont dans l'ensemble $\mathcal{H}$ des fonctions croissantes.

Dans le cas $\mathcal{X} = \mathcal{Y} = \mathbb{R}$, on a même $\mathcal{G} = \mathcal{H}$. En plus grande dimension, $T(x) = Rx$ où R est une petite rotation n'est pas un gradient mais est croissante. Il est possible que φ ne soit pas différentiable, mais on peut montrer que φ va être différentiable presque partout, et donc nos équations sont à entendre comme des égalités à ensemble de mesure nulle près.

Proposition 1.5 $T(x) = Ax$ est un gradient si et seulement si A est symétrique. $T \in \mathcal{G}$ si et seulement si A est définie positive.

Définition 1.4 La 2-distance de Wasserstein est définie par :

$$W_2^2(\alpha, \beta) = \inf_{T_\sharp\alpha=\beta} \int \|x - Tx\|^2 \, \mathrm{d}\alpha(x)$$

Quelques exemples avec les gaussiennes :

Cas 1D Ici, on suppose $\alpha = \mathcal{N}(m_\alpha, \sigma_\alpha^2)$, $\beta = \mathcal{N}(m_\beta, \sigma_\beta^2)$. En prenant $T(x) = \frac{\sigma_\beta}{\sigma_\alpha}(x - m_\alpha) + m_\beta$, on vérifie bien que $T_\sharp\alpha = \beta$ et en primitivant T, on vérifie bien que c'est le gradient (la dérivée) d'une fonction convexe et c'est donc bien un transport optimal. Le coût du transport est :

$$W_2^2(\alpha, \beta) = \|m_\alpha - m_\beta\|^2 + \|\sigma_\alpha - \sigma_\beta\|$$

Cas d-Dimensionnel Ici, $\alpha = \mathcal{N}(m_\alpha, \Sigma_\alpha)$ et $\beta = \mathcal{N}(m_\beta, \Sigma_\beta)$ où $\Sigma_\alpha = \mathbb{E}((X - m_\alpha)^{\mathrm{t}}(X - m_\alpha))$ est une matrice définie positive. On suppose que $\Sigma_\alpha > 0$. On prend alors $T(x) = A(x - m_\alpha) + m_\beta$ où $A\Sigma_\alpha{}^{\mathrm{t}}A = \Sigma_\beta$ pour trouver le transport optimal. Cette équation (dite de Ricatti) a une solution symétrique définie positive. On a de plus :

$$W_2^2(\alpha, \beta) = \|m_\alpha - m_\beta\|^2 + \mathcal{B}^2(\Sigma_\alpha, \Sigma_\beta)$$

où $\mathcal{B}$ est la distance de Bures définie par :

$$\mathcal{B}(A, B) = \mathrm{Tr}(A + B - 2(A^{1/2}BA^{1/2})^{1/2})$$

2 Formulation de Kantorovitch

2.1 Définition Discrète

La formulation de Kantorovitch est une relaxation convexe de la formulation de Monge. Il a obtenu un prix Nobel d'économie pour ceci. Ici, on se limite au cas discret $\alpha = \sum_n \alpha_i \delta_{x_i}$ et $\beta = \sum_m b_j \delta_{y_j}$.

Définition 2.1 Un *couplage* ou un *plan* est une matrice $M \in \mathbb{R}_+^{n\times m}$ qui représente le coût de transport de x_i à y_j et telle que :

$$\sum_j P_{i,j} = a_i \wedge \sum_i P_{i,j} = b_j$$

R C'est la même notion que celle de couplage en probabilité : un vecteur aléatoire sur l'espace produit.

On remarquera que les équations définissant un plan peuvent se mettre sous la forme :

$$P\mathbb{1}_m = a \text{ et } {}^{\mathsf{t}}P\mathbb{1}_n = b$$

On autorise ainsi la division de masse, les problèmes de Kantorovitch devenant des problèmes sur des graphes bipartis.

Définition 2.2 Le polytope des couplages entre α et β est l'ensemble des plans entre α et β :

$$\mathrm{Couplages}(\alpha, \beta) = \left\{ P \in \mathbb{R}_+^{n\times m} \,\middle|\, P_{i,j} \geq 0, P\mathbb{1}_m = a \text{ et } {}^{\mathsf{t}}P\mathbb{1}_n = b \right\}$$

Kantorovitch avait fait l'hypothèse très forte que l'économie est linéaire.

Définition 2.3 Le *problème de Kantorovitch* est le problème d'optimisation suivant :

$$P = \mathrm{argmin}_P \left\{ \langle C, P\rangle \,|\, P \in \mathrm{Couplages}(\alpha, \beta) \right\} \tag{Kantorovitch}$$

où $C \in \mathbb{R}^{n\times m}$ est une matrice de coût. On dit que P est le plan optimal.

C'est un problème de programmation linéaire. En général la méthode du simplexe n'est pas polynomial, mais il existe un type de simplexes pour lequel elle l'est, et est en $\mathcal{O}\left(\left(n^3 m + m^3 n\right) \log\left(mn\right)\right)$

Proposition 2.1 Il existe toujours une solution, et il existe toujours une solution dite *éparse*, telle que :

$$|\{(i,j) \,|\, P_{i,j} \neq 0\}| \leq n + m - 1$$

Démonstration. La preuve d'existence vient du fait que l'ensemble des couplages est un compact non vide (car $P = a^{\mathsf{t}}b$ est un couplage dit *indépendant*). ■

Le cas générique est de plus unique, c'est-à-dire que si C, α, β n'a pas une unique solution, en ajoutant du bruit on retrouve une solution optimale.

2.2 Équivalence à Monge

On s'intéresse ensuite aux matrices de permutation P_n, dans le cas $n = m$. On cherche à résoudre le problème non-convexe $\min\limits_{P\in P_n} \langle C, P\rangle$. Clairement, si $\mathcal{B}_n$ est l'ensemble convexe des matrices bistochastiques (l'ensemble des couplages !) :

$$\min_{P\in\mathcal{B}_n} \langle C, P\rangle \leq \min_{P\in P_n} \langle C, P\rangle$$

Définition 2.4 L'ensemble des points extrême d'un convexe C est :

$$\mathrm{Extr}(C) = \left\{ P \,\middle|\, \forall (Q,R) \in C^2, P = \frac{Q+R}{2} \Rightarrow Q = R \right\}$$

Théorème 2.1 Si C est compact, $\mathrm{Extr}(C) \neq \varnothing$.

Théorème 2.2 — Krain-Millman Si C est un compact convexe, alors $C = \mathrm{Hull}\left(\mathrm{Extr}\left(C\right)\right)$.

Proposition 2.2 Si C est compact :

$$\mathrm{Extr}(C) \cap \left(\mathrm{argmin}_{p \in C} \langle C, P \rangle\right) \neq \varnothing$$

Démonstration. En notant que l'ensemble de droite est convexe et compact, on trouve $\mathrm{Extr}(S) \neq \varnothing$ et que $\mathrm{Extr}(S) \subseteq \mathrm{Extr}(C)$. ■

Théorème 2.3 — Birkhoff - von Neumann On a $\mathrm{Extr}(\mathcal{B}_n) = P_n$

Démonstration. On montre d'abord $P_n \subset \mathrm{Extr}(\mathcal{B}_n)$. Cela découle de $\mathrm{Extr}([0,1]) = \{0,1\}$. Si $P \in P_n$, est telle que $P = (Q+R)/2$ avec $Q_{i,j}, R_{i,j} \in [0,1]$, puisque $P_{i,j} \in \{0,1\}$ alors nécessairement $Q_{i,j} = R_{i,j} \in \{0,1\}$.

Montrons maintenant $\mathrm{Extr}(\mathcal{B}_n) \subset P_n$ en montrant que $P_n^c \subset \mathrm{Extr}(\mathcal{B}_n)^c$ avec le complémentaire considéré dans $\mathcal{B}_n$. Choisir $P \in \mathcal{B}_n \setminus P_n$ revient à choisir $\mathbb{P} = (Q+R)/2$ où Q, R sont des matrices bistochastiques distinctes. P définit un graphe biparti de taille $2n$. Le graphe est composé d'arêtes isolées quand $P_{i,j} = 1$ et d'arêtes connectées quand $0 < P_{i,j} < 1$. Si i est un tel sommet à gauche (j à droite), puisque $\sum_j P_{i,j} = 1$, il y a deux arêtes (i, j_1) et (i, j_2) en sortant (de même, (i_1, j) et (i_2, j) entrant en j). On peut donc toujours extraire un cycle par récurrence de la forme :

$$(i_1, j_1, i_2, j_2, \ldots, i_p, j_p), \quad \text{i.e.} \quad i_{p+1} = i_1.$$

On suppose que ce cycle est le plus court de l'ensemble fini de cycle. On a toujours :

$$0 < P_{i_s, j_s}, P_{i_{s+1}, j_s} < 1.$$

Les $(i_s)_s$ et $(j_s)_s$ sont distincts puisque le cycle est le plus court. On pose :

$$\varepsilon = \min_{0 \leq s \leq p} \left\{ P_{i_s, j_s}, P_{j_s, i_{s+1}}, 1 - P_{i_s, j_s}, 1 - P_{j_s, i_{s+1}} \right\}$$

c'est-à-dire $0 < \varepsilon < 1$. En séparant le graphe en deux ensembles d'arêtes :

$$\mathcal{A} = \{(i_s, j_s)\}_{s=1}^{p} \quad \text{et} \quad \mathcal{B} = \{(j_s, i_{s+1})\}_{s=1}^{p}.$$

On pose Q et R telle que :

$$Q_{i,j} = \begin{cases} P_{i,j} & \text{si } (i,j) \notin \mathcal{A} \cup \mathcal{B}, \\ P_{i,j} + \varepsilon/2 & \text{si } (i,j) \in \mathcal{A}, \\ P_{i,j} - \varepsilon/2 & \text{si } (i,j) \in \mathcal{B}, \end{cases} \quad \text{et} \quad R_{i,j} = \begin{cases} P_{i,j} & \text{si } (i,j) \notin \mathcal{A} \cup \mathcal{B}, \\ P_{i,j} - \varepsilon/2 & \text{si } (i,j) \in \mathcal{A}, \\ P_{i,j} + \varepsilon/2 & \text{si } (i,j) \in \mathcal{B}, \end{cases}.$$

Par définition d'ε, on a $0 \leq Q_{i,j}, R_{i,j} \leq 1$. Puisque chaque arête gauche de $\mathcal{A}$ a une arête droite dans $\mathcal{B}$, (et réciproquement) la contrainte de somme sur les lignes (et sur les colonnes) est maintenue, donc $Q, R \in \mathcal{B}_n$. Finalement, on trouve : $P = (Q+R)/2$. ■

Corollaire 2.1 Pour $m = n$ et $a = b = \mathbb{1}_n$, il existe une solution optimale pour le problème Kantorovitch, qui est une matrice de permutation associée à une permutation optimale pour le problème Monge.

3 Métrique de Wasserstein

3.1 Formulation continue de Kantorovitch

On se place ici dans le cadre continu, où α et β sont des mesures de probabilité arbitraires sur $\mathcal{X}$, $\mathcal{Y}$. On notera P_1 et P_2 les projections

Définition 3.1 Un *couplage* entre α et β est une mesure de probabilité $\pi \in \mathcal{M}_+^1(\mathcal{X} \times \mathcal{Y})$ telle que $P_1 \sharp \pi = \alpha$ et $P_2 \sharp \pi = \beta$. On note $\mathcal{U}(\alpha, \beta)$ l'ensemble des couplages entre α et β.

En prenant α et β discrètes, on vérifie bien qu'on retrouve la forumlation de l'Equation Kantorovitch.

R Dans le cas où $\mathcal{U}(\alpha, \beta)$ est non vide, le produit tensoriel $\alpha \otimes \beta$ est un couplage, dit *indépendant*.

Définition 3.2 La formulation de Kantorovitch est le problème d'optimisation suivant :

$$\mathcal{L}_c(\alpha, \beta) = \inf_{\pi \in \mathcal{U}(\alpha, \beta)} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) \, \mathrm{d}\pi(x, y) = \inf_{(X,y) \sim (\alpha, \beta)} \mathbb{E}[c(X, Y)] \tag{Kantorovitch}$$

où $c : (\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}^+$ est la fonction de coût.

Proposition 3.1 On se place dans le cas $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$, $c(x, y) = \|x - y\|^2$. Si α a densité par rapport à la mesure de Lebesgue, avec $T = \nabla\varphi$ l'application optimale pour Monge, $\pi = (\mathrm{Id}, T)\sharp\alpha$ est l'application optimale pour le problème de Kantorovitch.

3.2 Propriété Métriques du Transport

Ici, on se place dans le cas $\mathcal{X} = \mathcal{Y}$ et on se donne une métrique d sur $\mathcal{X}$. On supposera $c = d^p$ pour un certain $p \geq 1$.

Définition 3.3 Pour $\alpha, \beta \in \mathcal{P}_p(\mathcal{X})$, on définit la distance p-Wasserstein comme :

$$\mathrm{W}_p(\alpha, \beta) = (\mathcal{L}_{d^p}(\alpha, \beta))^{1/p} = \left(\inf_{\pi \in \mathcal{U}(\alpha, \beta)} \int d^p(x, y) \, \mathrm{d}\pi(x, y) \right)^{\frac{1}{p}}$$

R On a une version définie à partir du problème de Monge :

$$\inf_{T \sharp \alpha = \beta} \int d^2(x, T(x)) \, \mathrm{d}\alpha(x)$$

mais c'est seulement une pseudo-distance.

Démonstration. Il reste à prouver que W_p définit bien une distance. Pour ça, on a simplement besoin de la séparation et de l'inégalité triangulaire.

Séparation On a : $\mathrm{W}_p(\alpha, \beta) = 0 \Rightarrow \int d(x, y) \, \mathrm{d}\pi^*(x, y) = 0 \Rightarrow d(x, y) = 0, \pi^*$ presque surement. Ceci implique que π^* est supporté sur la diagonale de $\mathcal{X}^2$, c'est-à-dire $P_1 \sharp \pi^* = \lambda = P_2 \sharp \pi^*$ et donc $\alpha = \beta$.

Inégalité Triangulaire On ne donne la preuve que dans le cas discret, $\alpha = \sum_i a_i \delta_{x_i}$, $\beta = \sum_j b_j \delta_{y_j}$, $\gamma = \sum_k c_k \delta_{z_k}$.

En prenant $\pi_{\alpha\beta}$ et $\pi_{\beta\gamma}$ des plans optimaux, on pose :

$$S_{ijk} = \frac{\pi_{\alpha\beta}(i,j)\pi_{\beta\gamma}(j,k)}{b_j}$$

si $b_j \neq 0$, 0 sinon. On a alors $\sum_i S_{ijk} = \pi_{\beta\gamma}(j,k)$ et $\sum_k S_{ijk} = \pi_{\alpha\beta}(i,j)$. Les trois marginales de S sont (α, β, γ). On définit alors :

$$\pi_{\alpha\gamma} = \sum_{i,k} \underbrace{(\sum_j S_{ijk})}_{\pi_{\alpha\gamma}(i,k)} \delta_{x_i, z_k}$$

On vérifie aisément que c'est bien un couplage entre α et γ. On a alors :

$$\begin{aligned}
W_p(\alpha,\gamma) &\leq \left(\sum_{i,k} \pi_{\alpha\gamma}(i,k) d\left(x_i, z_k\right)^p\right)^{1/p} \\
&= \left(\sum_{i,j,k} S_{ijk} d\left(x_i, z_k\right)^p\right)^{1/p} \\
&\leq \left(\sum_{i,j,k} S_{ijk}\left(d\left(x_i, y_j\right) + d\left(y_j, z_k\right)\right)^p\right)^{1/p} \\
&\leq \left(\sum_{i,j,k} S_{ijk} d\left(x_i, y_j\right)^p\right)^{1/p} + \left(\sum_{i,j,k} S_{ijk} d\left(y_j, z_k\right)^p\right)^{1/p} \\
&= \left(\sum_{ij} \pi_{\alpha\beta}(i,j) d\left(x_i, y_j\right)^p\right)^{1/p} + \left(\sum_{jk} \pi_{\beta\gamma}(j,k) d\left(y_j, z_k\right)^p\right)^{1/p} \\
&= W_p(\alpha,\beta) + W_p(\beta,\gamma)
\end{aligned}$$

Pour étendre la preuve à des mesures générales, on utilise le lemme de recollage.

■

Cette structure permet de munir l'espace $\mathcal{P}_p(\mathcal{X})$ d'une structure d'espace métrique. Dans la suite, on va supposer que $\mathcal{X}$ est compact, mais les définitions peuvent être étendues à $\mathbb{R}^d$ par exemple.

Proposition 3.2 Si $\mathcal{X}$ est borné, pour $1 \leq p \leq q$, on a :

$$\mathrm{W}_p(\alpha,\beta) \leq \mathrm{W}_q(\alpha,\beta) \leq (\mathrm{diam}\mathcal{X})^{\frac{q-p}{q}}\, \mathrm{W}_p(\alpha,\beta)^{p/q}$$

Démonstration. En posant $\varphi(s) = s^{q/p}$ (qui est convexe), par l'inégalité de Jensen :

$$\mathrm{W}_p(\alpha,\beta)^q \leq (\int d(x,y)^p \,\mathrm{d}\pi(x,y))^{q/p} \leq \int d(x,y)^q \,\mathrm{d}\pi(x,y)$$

Donc :

$$\mathrm{W}_p(\alpha,\beta) \leq \inf(\int d(x,y)^q \,\mathrm{d}\pi(x,y))^{1/q} = \mathrm{W}_q$$

La preuve est la même pour l'autre inégalité, en utilisant $d^q \leq (\mathrm{diam}\mathcal{X})^{q-p} d^p$. ■

R La propriété précédente montre que sur un compact, toutes les distances p-Wasserstein définissent la même topologie.

Définition 3.4 Une suite (α_n) converge $\star$-faiblement vers α, noté $\alpha_n \xrightarrow{\star} \alpha$ dans $\mathcal{M}_1^+(\mathcal{X})$ si pour tout f :

$$\int f \, \mathrm{d}\alpha_n \to \int f \, \mathrm{d}\alpha$$

R

- Dans le cas où $\alpha_n = \delta_{x_n}$, on a convergence vers $\alpha = \delta_x$ si et seuleement $x_n \to x$.
- La convergence $\star$-faible est la convergence en loi pour les variables aléatoires.

Définition 3.5 La topologie forte sur $\mathcal{M}_+^1(\mathcal{X})$ est celle définie par la distance de variation totale.

Proposition 3.3 Si on prend d la distance 0-1, alors $\mathrm{W}_p^p(\alpha, \beta) = \frac{1}{2} \|\alpha - \beta\|_{TV}$.

Proposition 3.4 Si $\mathcal{X}$ est compact, alors α_n converge $\star$-faiblement si et seulement si $W_p(\alpha_n, \alpha) \xrightarrow[n \to +\infty]{} 0$. W_p métrise donc la convergence $\star$-faible. Si $\mathcal{X}$ n'est pas compact, la convergence pour W_p équivaut à la convergence $\star$-faible et à la convergence des p-ème moments.

La convergence des sommes de Riemann est équivalente à la convergence $\star$-faible de $\frac{1}{n} \sum_k \delta k/n \to \mathcal{U}_{[0,1]}$.

3.3 Interpolation de McCann

On considère deux mesures de probabilité α, β sur $\mathcal{X} = \mathbb{R}^d$.

Définition 3.6 Pour π un plan optimal entre α et β, et pour $t \in [0,1]$, on définit :

$$\pi_t = P_t \sharp \pi \text{ où } P_t(x, y) = (1-t)x + ty$$

C'est une homotopie, qu'on appelle *Interpolation de McKann* ou *Interpolation de Déplacement*.

Proposition 3.5 L'interpolation de McCann vérifie :

$$\alpha_t \in \mathrm{argmin}_\rho (1-t) \, \mathrm{W}_2^2(\alpha, \rho) + t \, \mathrm{W}_2^2(\beta, \rho)$$

Proposition 3.6 L'espace $(\mathcal{P}_2(\mathbb{R}^d), \mathrm{W}_2)$ est un espace géodésique.

Démonstration. Si T est l'application optimale de Monge pour α, β :

$$\begin{aligned}
\mathrm{W}_2^2(\alpha_t, \alpha_s) &= \int \|(1-t)x + tT(x) - (1-s)x - sT(x)\|^2 \, \mathrm{d}\alpha(x) \\
&= \int \|(t-s)(T(x) - x)\|^2 \, \mathrm{d}\alpha(x) \\
&= |t-s|^2 \, W_2(\alpha, \beta)
\end{aligned}$$

■

4 Gaussiennes et Transport Optimal, Dualité

4.1 Dualité pour les mesures discrètes

Proposition 4.1 Si $\alpha = \sum a_i \delta_{x_i}$ et $\beta = \sum b_j \delta_{y_j}$, et C est une matrice $n \times m$ de coût, en notant $U(\alpha, \beta)$ l'ensemble des couplages entre α et β :

$$L_c(\alpha, \beta) = \max_{(u,v) \in R(\alpha,\beta)} \langle u, \alpha \rangle + \langle v, \beta \rangle$$

où $R(\alpha, \beta) = \{(u, v) \in \mathbb{R}^n \times \mathbb{R}^m \,|\, u \oplus v \leq C\}$.

Démonstration. On a $L_c(\alpha, \beta) = \min_{\pi \in U(\alpha,\beta)} \langle c, \pi \rangle$. On introduit donc :

$$\max_{(u,v) \in \mathbb{R}^n \times \mathbb{R}^m} \langle u, a - \pi \mathbb{1}_m \rangle + \left\langle v + b - {}^{\mathsf{t}}\pi \mathbb{1}_n \right\rangle = \begin{cases} 0 & \text{si } \pi \in U(\alpha, \beta) \\ +\infty & \text{sinon} \end{cases}$$

On a donc :

$$L_c(\alpha, \beta) = \min_{\pi \geq 0} \max_{(u,v) \in \mathbb{R}^n \times \mathbb{R}^m} L(\pi, u, v)$$

avec :

$$\underbrace{L(\pi, u, v))}_{\text{Lagrangien}} \langle c, \pi \rangle + \langle u, \alpha - \pi \mathbb{1}_m \rangle + \left\langle v + \beta - {}^{\mathsf{t}}\pi \mathbb{1}_n \right\rangle$$

Puisque dans le cadre de la programmation linéaire, avoir des solutions suffit pour avoir la dualité forte :

$$\begin{aligned} L_c(\alpha, \beta) &= \max_{u,v} \min_{\pi \geq 0} L(\pi, u, v) \\ &= \max_{u,v} \left(\langle u, \alpha \rangle + \langle v, \beta \rangle + \underbrace{\min_{\pi \geq 0} \langle c, \pi \rangle - \langle u, \pi \mathbb{1}_m \rangle - \left\langle v, {}^{\mathsf{t}}\pi \mathbb{1}_n \right\rangle}_{= \begin{cases} 0 & \text{si } u \oplus v \leq C \\ -\infty & \text{sinon} \end{cases}} \right) \end{aligned}$$

Finalement :

$$\boxed{L_c(\alpha, \beta) = \max_{(u,v) \in R(\alpha,\beta)} \langle u, \alpha \rangle + \langle v, \beta \rangle}$$

ce qui conclut la preuve. ■

R

- Si π est optimal, son support est inclus dans $\{(i, j) \,|\, u_i + v_j = C_{i,j}\}$
- $L_c(\alpha, \beta)$ est convexe en α, β (dual) mais est concave en C.

4.2 Dualité pour des mesures quelconques

Proposition 4.2 Si $\mathcal{X}, \mathcal{Y}$ sont compacts, alors :

$$L_C(\alpha, \beta) = \sup_{(f,g) \in R(C)} \int_{\mathcal{X}} f \,\mathrm{d}\alpha + \int_{\mathcal{Y}} g \,\mathrm{d}\beta$$

où :

$$R(C) = \{(f, g) \in \mathcal{C}(\mathcal{X}) \times \mathcal{C}(\mathcal{Y}) \,|\, \forall (x, y), f(x) + g(y) \leq C(x, y)\}$$

On dit que f, g sont des *potentiels de Kantorovitch*.

Dans le cas où $\mathcal{X}$ n'est pas compact, on remplace $\mathcal{C}$ par $\mathcal{C}_b$ et le résultat tient.

R

- Si π est optimal, de même, son support est tel que les potentiels sont égaux au coût.
- Dans le cas où α et β on retrouve le résultat discret.

4.3 c-transformations

On cherche à résoudre :

$$\max_{f\oplus g\leq c} \int g \,\mathrm{d}\beta, \quad f \text{ fixée}$$

On veut donc prendre g aussi grande que possible en ayant :

$$\begin{aligned} g(y) \leq & c(x,y) - f(x) \\ & \leq \inf_{x\in\mathcal{X}} c(x,y) - f(x) \end{aligned}$$

Définition 4.1 Étant donnée $f : \mathcal{X} \to \bar{\mathbb{R}}$, sa *c-transformée* est :

$$\begin{array}{rccl} f^c : & \mathcal{Y} & \to & \bar{\mathbb{R}} \\ & f^c(y) & \mapsto & \inf_{x\in\mathcal{X}} c(x,y) - f(x) \end{array}$$

La *$\bar{c}$-transformée* de $g : \mathcal{Y} \to \bar{R}$ est :

$$g^{\bar{c}}(x) = \inf_{y\in\mathcal{Y}} c(x,y) - g(y)$$

Si c est symmétrique, alors $f^c = f^{\bar{c}}$.

R La transformation $(f,g) \to (f,f^c)$ remplace les potentiels duaux par de meilleurs. De même pour $(f,g) \to (g^{\bar{c}}, g)$ et $(f,g) \to (g^{\bar{c}}, f^c)$.

Proposition 4.3 Si c est L-lipschitzienne en y, alors f^c est lipschitzienne.

Démonstration. Exercice. ■

À ce stade, on aurait envie d'itérer à partir de potentiels de base puis d'utiliser les c-transformées pour obtenir un meilleur résultat. La proposition ci-dessous montre que malheureusement ceci ne fonctionnera pas.

Proposition 4.4 Si on note $f^{c\bar{c}} = (f^c)^{\bar{c}}$ alors :

1. $f \leq \varphi \Rightarrow f^c \geq \varphi^c$;
2. $f^{c\bar{c}} \geq f$;
3. $g^{\bar{c}c} \geq g$;
4. $f^{c\bar{c}c} = f^c$.

Démonstration. 1. Par définition.

2. On a :

$$f^{c\bar{c}} = \inf_{y\in\mathcal{Y}} \left(c(x,y) - \underbrace{\underbrace{\inf_{x'\in\mathcal{X}} \left(c(x',y) - f\left(x'\right)\right)}_{\leq c(x,y)-f(x)}}_{\geq -(c(x,y)-f(x))} \right)$$
$$\geq \inf_{y\in\mathcal{Y}} \left(c(x,y) - c(x,y) + f(x)\right)$$

3. De même.
4. On a $f^{c\bar{c}} \geq f \Rightarrow f^{c\bar{c}c} \leq f^c$. Avec $g = f^c$, on a $f^{c\bar{c}c} \geq f^c$.

■

4.4 Quelques Cas particuliers

4.4.1 Cas Euclidien Quadratique

On veut ici calculer :

$$\min_{X\sim\alpha, Y\sim\beta} \mathbb{E}\left(\|X-Y\|^2\right) = K - 2 \max_{X\sim\alpha, Y\sim\beta} \mathbb{E}\left(\langle X, Y\rangle\right)$$

où K est une constante ne dépendant que de α et β. On va donc vérifier qu'écrire φ sous la forme $(\alpha, \nabla\varphi\sharp\alpha)$ fonctionne, c'est-à-dire redémontrer le théorème de Brenier :

Preuve du théorème de Brenier. On prend ici $c(x,y) = -\langle x, y\rangle$, qui est symmétrique, on obtient que :

$$f^c(y) = -\sup_x \langle x, y\rangle + f(x) = -(-f)^*(y)$$

pour $\cdot^*$ la transformation de Legendre-Fenchel. Puisqu'on sait que $(-f)^*$ est convexe, f^c est concave.

Si π est optimal pour L_c et f est optimal pour le dual, alors :

$$\operatorname{supp}(\pi) \subseteq \{(x,y) \mid f^{cc}(x) + f^c(y) = -\langle x, y\rangle\}$$

En notant $\varphi = -f^{cc}$, φ est convexe et donc :

$$\varphi^*(y) = \sup_x \langle x, y\rangle - \varphi(x) = -f^c(y)$$

On a donc :

$$\operatorname{supp}(\pi) \subseteq \underbrace{\{(x,y) \mid \varphi(x) + \varphi^*(y) = \langle x, y\rangle\}}_{\text{Sous-différentielle } \partial\varphi(x)}$$

Si φ est différentiable :

$$\operatorname{supp}(\pi) \subseteq \{(x, \nabla\varphi(x))\}$$

Puisque φ est convexe, φ est différentiable presque partout pour la mesure de Lebesgue, donc si α est absolument continue par rapport à la mesure de Lebesgue, φ est aussi différentiable presque partout pour α. ■

4.4.2 Cas Semi-discret

On suppose ici α absolument continue et $\beta = \sum_{j=1}^m b_j \delta_{y_j}$. On a :

$$\begin{aligned} L_c(\alpha,\beta) &= \max_{f,g\in\mathcal{C}(\mathcal{X})\times\mathcal{C}(\mathcal{Y}), f\oplus g\leq c} \int f\,\mathrm{d}\alpha + \int g\,\mathrm{d}\beta \\ &= \max_{f\oplus g\leq c} \int f\,\mathrm{d}\alpha + \sum_{j=1}^m b_j g(y_j) \end{aligned}$$

Avec $\varphi_v(x) = \min_j c(x,y_j) - v_j$, on peut montrer que :

$$\begin{aligned} L_c(\alpha,\beta) &= \max_{v\in(\mathbb{R}^d)^m} \int \varphi_v\,\mathrm{d}\alpha + \sum_{j=1}^m b_j v_j \\ &= \max_v \int \min_j \left(c\left(x,y_j\right) - v_j\right)\,\mathrm{d}\alpha + \sum_{j=1}^m b_j v_j \end{aligned}$$

Si on considère les cellules de Laguerre :

$$L_j(v) = \{x\in\mathcal{X}\,|\,\forall j'\neq j, c(x,y_j) - v_j \leq c(x,y_{j'}) - v_{j'}\}$$

On obtient :

$$L_c(\alpha,\beta) = \max_v \sum_j \int_{L_j(v)} \left(c(x,y_j) - v_j\right)\,\mathrm{d}\alpha + \langle b, v\rangle$$

4.4.3 Distance 1-Wasserstein

On se place dans le cas $c(x,y) = d(x,y)$ sur $\mathcal{X} = \mathcal{Y}$.

Proposition 4.5
1. f est c- concave si et seulement si f est δ-lipschitzienne pour $\delta \leq 1$.
2. Is $\mathrm{Lip}(f) \leq 1$, alors $f^c = -f0$

Proposition 4.6 Sous les hypothèses ci-dessus :

$$\mathrm{W}_1(\alpha,\beta) = \max_{\varphi,\mathrm{Lip}(\varphi)\leq 1} \int \varphi\,\mathrm{d}(\alpha-\beta)$$

R Dans le cas discret, $\alpha - \beta = \sum m_k \delta_{z_k}$ avec $\sum m_k = 0$. On a donc :

$$\mathrm{W}_1(\alpha,\beta) = \max_{u_k} \left\{ \sum_k u_k m_k \,\middle|\, \forall k,l, |u_k - u_l| \leq d(z_k,z_l) \right\}$$

R Si on est aussi dans le cas euclidien, la condition lipschitzienne globale peut se remplacer par :

$$\|\nabla\varphi\|_\infty \leq 1$$

5 Transport Optimal Tranché

5.1 Définition et propriétés

Définition 5.1 Given $u \in \mathbb{R}^d$, l'opérateur de tranchage (ou *slicing*) basé sur u est défini par :

$$p_u : x \in \mathbb{R}^d = \langle x, u \rangle$$

On va généraliser cette opération sur les mesures grâce à l'opérateur de poussage $\sharp$: Si $X \sim \mu$, alors $\langle u, X \rangle \sim p_u \sharp \mu$. Dans la suite on note P_u l'application de projection orthogonale sur u.

Définition 5.2 La distance de Wasserstein tranchée par c W_c entre μ et ν est la distance 1-dimensionnelle de Wasserstein basée sur $c : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ continue est :

$$S\,\mathrm{W}_p(\mu, \nu) = \mathbb{E}_{\theta \sim \mathcal{U}_{\mathbb{S}^{d-1}}} [\mathrm{W}(p_\theta \sharp \mu, p_\theta \sharp \nu)]$$

Démonstration. C'est bien une distance, tkt. ■

Proposition 5.1 Pour toutes deux distributions μ, ν sur $\mathbb{R}^d$:

$$S\,\mathrm{W}_p^p(\mu, \nu) \leq C_{d,p}\,\mathrm{W}_p^p(\mu, \nu)$$

où :

$$C_{d,p} = \frac{1}{d} \int_{\mathbb{S}^{d-1}} \|\theta\|_p^p \, \mathrm{d}\mathcal{U}_{\mathbb{S}^{d-1}}(\theta)$$

5.2 En pratique

On a l'approximation de Monte-Carlo suivante (puisque l'espérance sur la sphère n'est pas calculable en général) :

$$S\,\mathrm{W}_p^p(\mu, \nu) \simeq \frac{1}{K} \sum_{j=1}^{K} W_p(p_{\theta_j} \sharp \mu, p_{\theta_j} \sharp \nu)^p$$

6 Modèles de Flot et Diffusion

Dans cette section, on s'intéresse à l'espace $\mathbb{W}_p$ des mesures de probabilités sur $\mathbb{R}^d$ muni de la distance de p-Wasserstein.

6.1 Transport optimal dynamique

Dans la suite, $c(x, y) = \|x - y\|^2$. On s'intéresse d'abord aux géodésiques sur cet espaces, et à ce qu'elles nous apprennent en apprentissage.

Si on se donne α, β deux mesures de probabilités avec α à densité, et T une application optimale entre α et β, on définit μ_t la distribution de $X_t = (1 - t)X_0 + tT(X_0)$ avec $X_0 \sim \alpha$. L'application $t \mapsto \mu_t$ est donc une géodésique dans $\mathbb{W}_2$.

Définition 6.1 On définit $v_t(x) = (T - \mathrm{Id}) \circ T_t^{-1}(x)$, de sorte que $v_t \circ T_t(x) = T(x) - x$ est une constante. C'est le champ de vitesse constant sur les chemins de x à $T(x)$.

On a alors :

$$\int_0^1 \int \|v_t(x)\|^2 \, \mathrm{d}\mu_t \, \mathrm{d}t = \int_0^1 \int \|v_t \circ T_t(x)\|^2 \, \mathrm{d}\alpha \, \mathrm{d}t = \int_0^1 \int \|T(x) - x\|^2 \, \mathrm{d}\alpha \, \mathrm{d}t = \int \|T(x) - x\|^2 \, \mathrm{d}\alpha = \mathrm{W}_2(\alpha, \beta)$$

Définition 6.2 L'ensemble des chemins avec vitesses entre α et β deux mesures de probabilités est :

$$V(\alpha,\beta) = \{(\mu_t, v_t) \,|\, \mu_0 = \alpha, \mu_1 = \beta\}$$

où, pour tout t, $\mu_t \in \mathcal{P}(\mathbb{R}^d)$ et où $v : [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ est un champ de vitesses et où la paire (μ_t, v_t) vérifie l'équation de continuité (CE).

Proposition 6.1 Si $T = \nabla\varphi$ avec φ convexe, alors T est monotone et pour tout t, T_t est injective.

Définition 6.3 La paire (μ_t, v_t) satisfie l'équation de continuité si :

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}t} + \mathrm{div}(\mu_t v_t) = 0 \tag{CE}$$

où l'égalité est entendue au sens des distributions.

Il reste donc à vérifier que $\mu_t = T_t \sharp \alpha$ et $v_t = (T - \mathrm{Id}) \circ T_t^{-1}(x)$ vérifient l'équation de continuité. On se donne ψ a support compact sur $]0,1[\times\mathbb{R}^d$ et on a :

$$\int_0^1 \int \frac{\mathrm{d}\psi}{\mathrm{d}t}\,\mathrm{d}\mu_t\,\mathrm{d}t = \int_0^1 \int \frac{\mathrm{d}\psi_t}{\mathrm{d}t}(T_t(x))\,\mathrm{d}\alpha\,\mathrm{d}t$$
$$\int_0^1 \int_{\mathbb{R}^d} \langle \nabla_x \psi, v_t(x)\rangle\,\mathrm{d}\mu_t\,\mathrm{d}t = \int_0^1 \int_{\mathbb{R}^d} \langle \nabla_x \psi \circ T_t(x), T(x) - x\rangle\,\mathrm{d}\alpha\,\mathrm{d}t$$

Ainsi :

$$\int_0^1 \int_{\mathbb{R}^d} \left(\frac{\mathrm{d}\psi_t}{\mathrm{d}t} \circ T_t(x) + \langle \nabla_x \psi, v_t(x)\rangle \right) \mathrm{d}\mu_t\,\mathrm{d}t = \int_0^1 \int_{\mathbb{R}^d} \underbrace{\left(\frac{\mathrm{d}\psi_t}{\mathrm{d}t} \circ T_t(x) + \langle \nabla_x \psi \circ T_t(x), T(x) - x\rangle \right)}_{=\frac{\mathrm{d}}{\mathrm{d}t}(\psi(t, T(t,x)))} \mathrm{d}\alpha\,\mathrm{d}t$$

Finalement :

$$\int_0^1 \int_{\mathbb{R}^d} \left(\frac{\mathrm{d}\psi_t}{\mathrm{d}t} \circ T_t(x) + \langle \nabla_x \psi, v_t(x)\rangle \right) \mathrm{d}\mu_t\,\mathrm{d}t = \int_{\mathbb{R}^d} \underbrace{(\psi(1, T(1,x)) - \psi(0, T(0,x)))}_{=0}\,\mathrm{d}\alpha$$

On vérifie donc bien que :

Proposition 6.2

$$W_2^2(\alpha,\beta) \geq \inf_{V(\alpha,\beta)} \int_0^1 \int \|v_t(x)\|^2\,\mathrm{d}\mu_t\,\mathrm{d}t$$

On rappelle que le flot φ de v est défini par $\frac{\mathrm{d}\varphi_t(x)}{\mathrm{d}t} = v(t, \varphi_t(x))$ et $\varphi_0(x) = x$.

Avec la définition de v_t ci-dessus, on voit que T_t est le flot de v_t.

Théorème 6.1 Si v_t est uniformément borné et Lipschitz en x (uniformément en t), si φ_t est son flot et si de plus α est absoluement continue par rapport à la mesure de Lebesgue, alors, $\rho_t = \varphi_t \sharp \alpha$ est l'unique solution de (CE) avec la confition initiale $\rho_0 = \alpha$.

On obtient donc le Théorème suivant :

Théorème 6.2 — Bénamou-Brenier On a :

$$W_2^2(\alpha,\beta) = \min_{V(\alpha,\beta)} \int_0^1 \int \|v_t\|^2 \, \mathrm{d}\mu_t \, \mathrm{d}t$$

la solution étant unique et donnée par l'interpolation de McCann et sa vitesse correspondante v_t.

Démonstration. Utilisant l'équation de la Proposition 6.2 on a déjà une inégalité. En utilisant le Théorème 6.1 ci-dessus, on obtient :

$$\begin{aligned}
\int_0^1 \int_{\mathbb{R}^d} \|v_t(x)\|^2 \, \mathrm{d}\mu_t \, \mathrm{d}t &= \int_0^1 \int_{\mathbb{R}^d} \|v_t \circ \varphi_t(x)\|^2 \, \mathrm{d}\alpha \, \mathrm{d}t \\
&\overset{Fubini}{=} \int_{\mathbb{R}^d} \int_0^1 \|v_t \circ \varphi_t(x)\|^2 \, \mathrm{d}t \, \mathrm{d}\alpha \\
&\overset{Jensen}{\geq} \int_{\mathbb{R}^d} \left\| \int_0^1 \underbrace{v_t \circ \varphi_t(x)}_{\frac{\mathrm{d}\varphi_t}{\mathrm{d}t}(x)} \, \mathrm{d}t \right\|^2 \mathrm{d}\alpha \\
&= \int_{\mathbb{R}^d} \|\varphi(1,x) - x\|^2 \, \mathrm{d}\alpha \\
&\overset{\varphi_1 \sharp \alpha = \beta}{\geq} W_2^2(\alpha,\beta)
\end{aligned}$$

Pour l'unicité de la solution, on se donne deux solutions et on introduit leur moyenne, l'inégalité de Cauchy-Schwarz (et son cas d'égalité) nous permet de conclure la preuve. ■

6.2 Couplage par flot

On va appliquer le Théorème 6.1 et construire une paire $(\mu_t, v_t) \sim$ (CE) telle que $\mu_0 = \alpha$, $\mu_1 = \beta$. Le couplage par flot est une manière de construire un poussé en avant entre α et β, mais celui ci ne sera pas nécessairement optimal.

C'est une méthode applicable à la modélisation générative :

1. On échantillonne $\alpha = \mathcal{N}(0, \mathrm{Id})$.
2. On calcule un champ v_θ tel que $\mu_t, v_\theta \sim$ (CE) où μ_t est un chemin de α à β.
3. On intègre v_θ pour optenir φ^θ.
4. On échantillone β en échantillonnant α puis en considérant l'image par φ_1^θ.

Pour définir μ_t, v_t de manière efficace on a l'algorithme suivant :

1. On part d'un couplage $X_0, X_1 \sim \Pi \in \mathcal{U}(\alpha, \beta)$.
2. On définit $X_t = tX_1 + (1-t)X_0$ et μ_t la loi de X_t.
3. On calcule ensuite v_t comme l'espérance de $X_1 - X_0$ sachant $X_t = x$.

Proposition 6.3 La paire (μ_t, v_t) définie ci-dessus vérifie (CE), et donc $(\mu_t, v_t) \in V(\alpha, \beta)$.

7 Barycentre de Sinkhorn

7.1 Régularisation entropique

Considérons deux mesures discrètes $\alpha = \sum_{i=1}^n a_i \delta_{x_i}$ et $\beta = \sum_{j=1}^m b_j \delta_{y_j}$.

Définition 7.1 Le problème de Schrödinger statique s'écrit :

$$\min_{P\in\mathbb{R}_+^{n\times m}, P\mathbf{1}_m=a, P^\top\mathbf{1}_n=b} \langle C,P\rangle + \varepsilon H(P) \tag{Schr}$$

où $H(P)=\sum_{i,j} P_{i,j}\log P_{i,j}$ est l'entropie de Shannon négative de la matrice de transport P, avec la convention $0\log 0=0$. Le paramètre ε est appelé *paramètre de régularisation entropique*, ou *température* en physique statistique.

R H est strictement convexe, puisqu'on se trouve dans le simplexe des matrices à coefficients compris entre 0 et 1.

Proposition 7.1 Si $\varepsilon>0$, le problème de Schrödinger admet une unique solution P_ε. Si $\varepsilon=0$, on retrouve la formulation de Kantorovich, qui peut admettre plusieurs solutions.

Proposition 7.2 Lorsque $\varepsilon\to 0$, la solution du problème de Schrödinger converge vers la solution de Kantorovich qui maximise l'entropie :

$$P_\varepsilon \xrightarrow[\varepsilon\to 0]{} \operatorname{argmin}_P \{H(P) : P \text{ est une solution de Kantorovich}\}$$

À l'inverse, lorsque $\varepsilon\to+\infty$, on retrouve le couplage trivial :

$$P_\varepsilon \xrightarrow[\varepsilon\to+\infty]{} a\otimes b$$

L'intuition derrière cette régularisation entropique est l'ajout d'une fonction de barrière, de manière similaire aux méthodes de points intérieurs en optimisation convexe. Une différence importante étant que l'on utilise ici l'entropie de Shannon négative, et non pas une fonction logarithmique classique. Cela permet de garantir la positivité des coefficients de la matrice de transport P.

Théorème 7.1 Supposons sans perte de généralité que $a_i>0$ pour tout i et $b_j>0$ pour tout j. On a alors :

$$P \text{ est solution du problème de Schrödinger} \iff \begin{cases} P\mathbf{1}_m=a, P^\top\mathbf{1}_n=b \\ \exists u\in\mathbb{R}^n, v\in\mathbb{R}^m, P_{i,j}=u_iK_{i,j}v_j \text{ avec } K_{i,j}=e^{-C_{i,j}/\varepsilon} \end{cases}$$

De manière équivalente, un couplage P est solution s'il existe des vecteurs $u\in\mathbb{R}^n$ et $v\in\mathbb{R}^m$ tels que :

$$P=\operatorname{diag}(u)K\operatorname{diag}(v)$$

où K est la matrice de noyau exponentiel défini par $K_{i,j}=e^{-C_{i,j}/\varepsilon}$.

Démonstration. On veut résoudre le problème de Schrödinger :

$$\min_{P\mathbf{1}_m=a, P^\top\mathbf{1}_n=b} f(P):=\langle C,P\rangle+\varepsilon H(P)$$

Commençons par montrer par l'absurde que si $P^\star$ est solution, alors $P^\star_{i,j}>0$ pour tout (i,j). Supposons qu'il existe (i_0,j_0) tel que $P^\star_{i_0,j_0}=0$. Posons $P^t(1-t)P^\star+t(a\otimes b)$ pour $t\in[0,1]$. Alors P^t est admissible pour tout t. De plus, si l'on pose $g(t)=f(P^t)$, alors $g'(0)=-\infty$ car la dérivée de l'entropie en 0 est infinie. Donc pour t suffisamment petit, $f(P^t)<f(P^\star)$, ce qui contredit le fait que $P^\star$ est solution. Ainsi, $P^\star_{i,j}>0$ pour tout (i,j). Ceci justifie (peu rigoureusement) le fait que l'on peut omettre la contrainte de positivité dans la suite de la preuve.

Dérivons le problème dual de Lagrange. On introduit les multiplicateurs de Lagrange $\{f_i\}_{i=1}^n$ et $\{g_j\}_{j=1}^m$ associés aux contraintes de marges. Le lagrangien s'écrit :

$$L(P,f,g)=\langle C,P\rangle+\varepsilon H(P)+\langle a-P\mathbf{1}_m,f\rangle+\left\langle b-P^\top\mathbf{1}_n,g\right\rangle$$

(Ici, les conditions de Slater sont vérifiées car le problème est convexe et admet une solution intérieure, on peut donc échanger le min et le max.)

Remarquons que $\langle a - P\mathbf{1}_m, f\rangle = \langle a, f\rangle - \langle P\mathbf{1}_m, f\rangle = \langle a, f\rangle - \left\langle P, f\mathbf{1}_m^\top\right\rangle$. On a alors :

$$\nabla_P L = C + \varepsilon(\log P + 1) - f\mathbf{1}_m^\top - \mathbf{1}_n g^\top.$$

En annulant ce gradient, on obtient :

$$P_{i,j} = e^{(f_i + g_j - C_{i,j})/\varepsilon - 1} = e^{-1} e^{f_i/\varepsilon} e^{-C_{i,j}/\varepsilon} e^{g_j/\varepsilon}.$$

En posant $u_i = e^{f_i/\varepsilon - 1/2}$ et $v_j = e^{g_j/\varepsilon - 1/2}$, on retrouve bien la forme annoncée. ∎

7.2 Algorithme de Sinkhorn

On note le produit de Kronecker $\operatorname{diag}(u)z = u \odot z = (u_i z_i)_i$ le produit terme à terme entre les vecteurs u et z. On a donc $P\mathbf{1}_m = u \odot (Kv)$ qui doit valoir a, et de même $P^\top \mathbf{1}_n = v \odot (K^\top u) = b$. On a ainsi le système suivant :

$$\begin{cases} u \odot (Kv) = a \\ v \odot (K^\top u) = b \end{cases} \tag{SchrSys}$$

Théorème 7.2 Le problème de Schrödinger (Schr) est équivalent à la résolution du système (SchrSys).

Proposition 7.3 L'algorithme de Sinkhorn pour la résolution du problème de Schrödinger est le suivant :

- $v \leftarrow \mathbf{1}_m$
- Répéter jusqu'à convergence :
  - $u \leftarrow a/(Kv)$
  - $v \leftarrow b/(K^\top u)$

où l'on note z/w le quotient terme à terme entre les vecteurs z et w.

Cet algorithme est très simple et facilement parallélisable. Chaque itération coûte $\mathcal{O}(n^2)$ opérations, donnant une complexité totale de $\mathcal{O}(Tn^2)$ pour atteindre une précision ε en T étapes ; ceci est à comparer à des algorithmes comme celui du simplexe, qui a une complexité cubique.

Théorème 7.3 Il suffit de $T = \frac{1}{\varepsilon^2}$ itérations pour atteindre une précision ε.

R On a donc une complexité totale de $\mathcal{O}\left(\frac{n^2}{\varepsilon^2}\right)$ pour atteindre une précision ε. Par rapport aux méthodes à points intérieurs de complexité $\mathcal{O}(n^3 \log(\varepsilon))$, la méthode de Sinkhorn est donc plus efficace en termes d'échantillons n, mais moins efficace en termes de précision ε.

Démonstration. S'intéresse à la preuve par contraction qui donne une convergence linéaire, contrairement à une preuve par projection itérative qui donnerait une convergence sous-linéaire, la constante étant meilleure dans le cas linéaire. *Voir les notes de cours pour la preuve.* ∎

7.3 Reformulation en divergence de Kullback-Leibler

Définition 7.2 La divergence de Kullback-Leibler entre deux matrices $P, Q \in \mathbb{R}_+^{n\times m}$ est définie par :

$$\mathrm{KL}(P|Q) = \sum_{i,j} P_{i,j} \log\left(\frac{P_{i,j}}{Q_{i,j}}\right) - P_{i,j} + Q_{i,j}$$

Proposition 7.4 $\mathrm{KL}(P|Q) \geq 0$ avec égalité si et seulement si $P = Q$.

L'idée va être d'utiliser $Q = a \otimes b$ comme mesure de référence.

Proposition 7.5 Le problème de Schrödinger (Schr) s'écrit de manière équivalente :

$$\min_{P\in\mathbb{R}_+^{n\times m}, P\mathbf{1}_m = a, P^\top \mathbf{1}_n = b} \langle C, P\rangle + \varepsilon\,\mathrm{KL}(P|a\otimes b).$$

Démonstration. $\mathrm{KL}(P|a\otimes b)$ et $\mathrm{KL}(P|a'\otimes b')$ diffèrent d'une constante additive indépendante de P, donc le minimum est le même. ■

Définition 7.3 Soit $\frac{\pi}{\xi} \in \mathcal{P}(X\times Y)$. Si $\frac{\mathrm{d}\pi}{\mathrm{d}\xi}$ existe, on définit la divergence de Kullback-Leibler entre π et ξ par :

$$\mathrm{KL}(\pi|\xi) = \int_{X\times Y} \log\left(\frac{\mathrm{d}\pi}{\mathrm{d}\xi}\right) \mathrm{d}\pi - \pi(X\times Y) + \xi(X\times Y)$$

Si $\frac{\mathrm{d}\pi}{\mathrm{d}\xi}$ n'existe pas, on pose $\mathrm{KL}(\pi|\xi) = +\infty$.

Définition 7.4 Le **problème de Schrödinger général** s'écrit :

$$\inf_{\pi\in\mathcal{P}(X\times Y)} \left\{\int c\ \mathrm{d}\pi + \varepsilon\ \mathrm{KL}(\pi|a\otimes b) : \pi_1 = \alpha,\ \pi_2 = \beta\right\} \tag{SchrGen}$$

R L'information mutuelle entre deux variables aléatoires X et Y de loi jointe π est donnée par $I(X,Y) = \mathrm{KL}(\pi|\pi_1\otimes\pi_2)$. Ainsi, le problème de Schrödinger cherche un couplage π qui minimise le coût total plus une pénalisation de l'information mutuelle entre les deux variables aléatoires. Lorsque $I(X,Y) = 0$, les variables sont indépendantes, et le couplage est donc le produit tensoriel des marges.

Calculons le problème dual de Schrödinger dans ce cadre général. En réutilisant les notations des multiplicateurs de Lagrange, on a :

$$\min_P \max_{f,g} \langle C, P\rangle + \varepsilon\,\mathrm{KL}(P|a\otimes b) + \langle \alpha - P\mathbf{1}, f\rangle + \left\langle \beta - P^\top\mathbf{1}, g\right\rangle$$

Puisque les conditions de Slater sont vérifiées, on peut échanger le min et le max :

$$\max_{f,g} \langle a, f\rangle + \langle b, g\rangle + \min_P \left\langle C - f\mathbf{1}^\top - \mathbf{1}g^\top, P\right\rangle + \varepsilon\,\mathrm{KL}(P|a\otimes b),$$

ce qui est équivalent à :

$$\max_{f,g} \langle a, f\rangle + \langle b, g\rangle - \max_P \left\langle C - f\mathbf{1}^\top - \mathbf{1}g^\top, P\right\rangle - \varepsilon\,\mathrm{KL}(P|a\otimes b).$$

On remarque alors que ceci correspond à la transformation de Legendre-Fenchel de $\mathrm{KL}(\cdot|a\otimes b)$ évaluée en $(f\mathbf{1}^\top + \mathbf{1}g^\top - C)/\varepsilon$. On a donc :

$$\max_{f,g} \langle a, f\rangle + \langle b, g\rangle - \varepsilon\,\mathrm{KL}^\star(z|ab^\top)$$

où $z = (f\mathbf{1}^\top + \mathbf{1}g^\top - C)/\varepsilon$.

Lemme 7.1 La transformation de Legendre-Fenchel de $\mathrm{KL}(\cdot|a\otimes b)$ est donnée par :

$$\mathrm{KL}^\star(Z|Q)=\sum_{i,j}Q_{i,j}\exp(Z_{i,j})-1$$

On en déduit le dual de Schrödinger :

$$\max_{f,g}\langle a,f\rangle+\langle b,g\rangle-\varepsilon\sum_{i,j}a_ib_j\exp\left(\frac{f_i+g_j-C_{i,j}}{\varepsilon}\right)+\varepsilon$$

Dans le cas général, pour des mesures non-discrètes, on retrouve une formulation similaire :

$$\inf_{f\in\mathcal{C}(X),g\in\mathcal{C}(Y)}\int f(x)\,\mathrm{d}\alpha(x)+\int g(y)\,\mathrm{d}\beta(y)-\varepsilon\int e^{(f(x)+g(y)-c(x,y))/\varepsilon}\,\mathrm{d}\alpha(x)\,\mathrm{d}\beta(y)+\varepsilon$$

Proposition 7.6 L'**algorithme de Sinkhorn** dans ce cadre s'écrit alors :

- Initialiser g
- Répéter jusqu'à convergence :
  — $f\leftarrow g^{c,\varepsilon}:=\operatorname{argmin}_f D(f,g)$
  — $g\leftarrow f^{c,\varepsilon}:=\operatorname{argmin}_g D(f,g)$

où l'on dénote $(\cdot)^{c,\varepsilon}$ l'opérateur de c-transformée réguliarisée par l'entropie, et D la fonctionnelle duale de Schrödinger.

Proposition 7.7 L'**opérateur de c-transformée réguliarisée par l'entropie** s'écrit :

$$f^{c,\varepsilon}(y)=-\varepsilon\log\left(\int e^{(f(x)-c(x,y))/\varepsilon}\,\mathrm{d}\alpha(x)\right)$$

et de manière similaire :

$$g^{c,\varepsilon}(x)=-\varepsilon\log\left(\int e^{(g(y)-c(x,y))/\varepsilon}\,\mathrm{d}\beta(y)\right)$$

8 Flots de Gradients de Wasserstein

On va ici étudier les dynamiques de $\alpha_t\in\mathcal{P}(\mathbb{R}^d)$ pour t une variable de temps dans $\mathbb{R}^+$.

8.1 Formulation en vitesse

Dans notre cas, on va utiliser une approche *à la Lagrange*, on va prendre v_t tel que :

$$\partial_t\mu_t+\operatorname{div}(\mu_tv_t)=0$$

Si on suppose que v_t est Lipschitz, il existe une unique solution en μ_t. Toutefois, il en existe beaucoup en v_t, puisqu'il existe beaucoup de champs vectoriels de divergence nulle. Dans le cas où μ_t a densité ρ_t, on écrira plutôt

$$\partial_t\rho_t+\operatorname{div}(\rho_tv_t)=0 \tag{Speed}$$

Pour se simplifier, on utilisera plutôt, dans la vraie vie, la formulation faible de l'équation différentielle ci-dessus. On intègre par rapport à $\varphi_t(x)$ à support compact :

$$\int\partial_t\rho_t\varphi_t(x)\,\mathrm{d}x+\int\operatorname{div}(\rho_tv_t)\varphi_t(x)\,\mathrm{d}x\,\mathrm{d}t$$

et on vérifiera à chaque fois qu'on peut réécrire nos propriétés via

$$\forall \varphi_t \quad -\int \partial_t \varphi_t \,\mathrm{d}\alpha_t \,\mathrm{d}t - \int \langle \nabla \varphi_t(x), v_t \rangle \,\mathrm{d}\alpha \,\mathrm{d}t = 0 \tag{WeakSpeed}$$

On retrouve le transport optimal dynamique du théorème 6.1. On remarque notamment que la vitesse v_t optimale est un gradient.

8.2 Optimisation sur particules

On cherche à optimiser $f(\alpha)$ sur $\alpha \in \mathcal{P}(\mathbb{R}^d)$. Ici α représente une densité de particule, par exemple. Dans le cas où $\alpha_t = \delta_{x_t}$ est une seule particule, on retrouve la descente de gradient, si $g(x) = f(\delta_x)$ on résout en prenant $\dot{x} = -\nabla g(x)$.

Dans le cas de n particules, i.e. $\alpha_t = \frac{1}{n}\sum_{i=1}^n \delta_{x_i(t)}$,

$$X(t) = (x_1(t), \dots, x_n(t)) \in \mathbb{R}^{d\times n},$$

donc en écrivant $g(X) = f\left(\frac{1}{n}\sum \delta_{x_i}\right)$ on obtient la descente de particules $\dot{X} = -\nabla g(X)$, avec

$$g(X) = \frac{1}{n^2}\sum_{i,j=1}^n k(x_i, x_j) \quad \text{avec} \quad k(x, x') = \frac{1}{\|x - x'\| + \varepsilon}$$

On retrouve ainsi $f(\alpha) = \iint k(x, x')\,\mathrm{d}\alpha(x)\,\mathrm{d}\alpha(x')$.

On va donc s'intéresser à ce qui se passe quand $n \to \infty$ et essayer de faire sens de l'entropie dans ce cadre.

Pour cela on utilise la méthode JKO (Jordan-Kinderlheren-Otto), une généralisation de la méthode implicite d'Euler pour Wasserstein. On se donne $\tau > 0$ un pas. On prend $\alpha_{t+\tau} = \operatorname{argmin} \frac{1}{2\tau}\,\mathrm{W}_2^2(\alpha_t, \alpha) + f(\alpha)$.

Quand $\tau \to 0$, si on suppose que $(\alpha_t)_t$ converge vers une courbe continue $t \in \mathbb{R}^+ \mapsto \alpha_t$, alors elle converge vers une équation aux dérivées partielles (Speed), avec v_t à définir. On va avoir besoin de calculer deux dérivées différentes : la dérivée classique (aussi appelée verticale, première variation ou Fréchet) et la dérivée de Wasserstein (aussi appelée dérivée d'Otto ou dérivée horizontale). La dérivée verticale est une fonction réelle $\delta f(\alpha) \in \mathcal{C}(\mathbb{R}^d, \mathbb{R})$ tandis que la dérivée horizontale $\nabla_{\mathrm{W}} f(\alpha) \in \mathcal{C}(\mathbb{R}^d, \mathbb{R}^d)$ est un champ de vecteur.

Définition 8.1 La dérivée verticale de f en α est, étant donné $\gamma = \beta - \alpha$ une mesure de masse nulle définie par l'expansion de Taylor

$$f(\alpha + t\gamma) = f(\alpha)t\,\langle \gamma, \delta f(\alpha)\rangle + o(t),$$

le produit scalaire étant entendu au sens du produit fonction mesure.

Dans le cas où $f(\alpha) = \int g(x)\,\mathrm{d}\alpha(x)$, on a $\nabla_{\mathrm{W}} f(\alpha) = t\int g\,\mathrm{d}\gamma = g$. Dans le cas où $f(\alpha) = \int U(p(x))\,\mathrm{d}x$ avec $p(x) = \frac{\mathrm{d}\alpha}{\mathrm{d}x}$, $f(\alpha) = +\infty$ et U une fonction d'énergie interne, si α n'a pas de densité, on calcule (avec $\eta = \frac{\mathrm{d}\gamma}{\mathrm{d}x}$)

$$\begin{aligned}
f(\alpha + t\gamma) &= \int U(p(x) + t\eta(x))\,\mathrm{d}x \\
&= \int \left[U(p(x)) + tU(p(x))\eta(x) + o(t)\right]\,\mathrm{d}x \\
&= \int f(\alpha) + t\int U'(p(x))\,\mathrm{d}\gamma(x) + o(t)
\end{aligned}$$

Donc $\delta f(\alpha)(x) = U'(p(x))$.

Définition 8.2 Le gradient de Wasserstein s'écrit donc $\nabla_{\mathrm{W}} f(\alpha) = \nabla_{\mathbb{R}^d}[\delta f(\alpha)]$.

Théorème 8.1 Si α_t existe, les équations aux dérivées partielles

$$\partial_t \alpha + \operatorname{div}(\alpha v) = 0 \quad \text{et} \quad v = -\nabla_{\mathrm{W}} f(\alpha)$$

Si on reprend l'exemple avec l'énergie interne, on obtient :

$$\partial_t \rho + \operatorname{div}(\rho v) + 0$$

avec $v = \nabla_{\mathrm{W}} f(\alpha) = \nabla \delta f(\alpha) = \nabla(\log \rho + 1) = \frac{\nabla \rho}{\rho}$, c'est à dire donc $\partial_t \rho = \Delta \rho$. On retrouve l'équation de la chaleur.

Optimal Transport : Exercices

1 Monge's Formulation

1.1 Warm-up (discrete assignment)

On énumère les $3! = 6$ permutations dans S_3, et l'on calcule pour chaque permutation σ le coût associé :

$$\frac{1}{3}\sum_{i=1}^{3} C_{i,\sigma(i)}$$

On trouve alors deux permutations optimales : $\sigma_1 = (1,2,3)$ et $\sigma_2 = (2,1,3)$, de coût 5.

1.2 Two points to two points

On écrit explicitement le coût associé avec Id et celui associé à Swap :

$$c_{\mathrm{Id}} = \|x_1 - y_1\|^2 + \|z - y_2\|^2, \qquad c_{\mathrm{Swap}} = \|x_1 - y_2\|^2 + \|z - y_1\|^2.$$

L'ensemble des z tels que Id est préférable à Swap est alors donné par :

$$\left\{ z \in \mathbb{R}^d \;\middle|\; c_{\mathrm{Id}}(z) \leqslant c_{\mathrm{Swap}}(z) \right\}.$$

En développant les normes au carré, on trouve :

$$\begin{aligned} & c_{\mathrm{Id}}(z) \leqslant c_{\mathrm{Swap}}(z) \\ \iff \quad & \|x_1\|^2 - 2\langle x_1 \,|\, y_1\rangle + \|y_1\|^2 + \|z\|^2 - 2\langle z \,|\, y_2\rangle + \|y_2\|^2 \\ & \leqslant \|x_1\|^2 - 2\langle x_1 \,|\, y_2\rangle + \|y_2\|^2 + \|z\|^2 - 2\langle z \,|\, y_1\rangle + \|y_1\|^2 \end{aligned}$$

En simplifiant, on trouve que l'ensemble cherché est :

$$\left\{ z \in \mathbb{R}^d \;\middle|\; \langle z \,|\, y_1 - y_2\rangle \leqslant \langle x_1 \,|\, y_1 - y_2\rangle \right\},$$

ce qui est un demi-espace d'équation $\langle z \,|\, y_1 - y_2\rangle = \langle x_1 \,|\, y_1 - y_2\rangle$.

1.3 Discrete to continuous and continuous to discrete

1.3.1

Soit $\alpha, \beta \in \mathcal{M}_1^+(\mathbb{R})$, avec $\alpha = \sum_{i=1}^N a_i \delta_{x_i}$ et β absolument continue par rapport à la mesure de Lebesgue. Supposons que nous ayons T telle que $T_\sharp \alpha = \beta$. Alors par définition du *pushforward* :

$$\sum_{i=1}^{N} a_i \delta_{T(x_i)} = \beta,$$

et puisque pour tout i, $\lambda\{T(x_i)\} = 0$, on a $\beta(\{T(x_i)\}) = 0$. Dès lors, $\sum_i a_i \delta_{T(x_i)}(\{T(x_i)\}) = a_i = 0$, ce qui est absurde.

1.3.2

On cherche une partition $(\Omega_j)_j$ de $\mathcal{X} = [0,1]$ telle que :

$$\mathcal{X} = [0,1] = \bigcup_j \Omega_j, \quad \text{avec} \quad \forall j, \alpha(\Omega_j) = b_j.$$

On peut alors définir T telle que $T(x) = y_j$ pour $x \in \Omega_j$. Par construction, on a bien $T_\sharp \alpha = \beta$.

On peut prendre n'importe quelle partition, par exemple pour une permutation σ, $\Omega_j = \Big[f^{-1}(\sum_{k=1}^{j-1} b_{\sigma(k)}), f^{-1}(\sum_{k=1}^{j} b_{\sigma(k)})\Big)$, la dernière partie étant fermée, où $f(x) = \int_0^x \rho(t)dt$ est la fonction de répartition de α. Cette construction est plus simple dans le cas α uniforme (question suivante), puisque la fonction de répartition est l'identité.

1.3.3

On cherche T qui minimise :

$$\inf \left\{ \int_{\mathbb{R}} \|x - T(x)\|^p \, \mathrm{d}\alpha(x) \,\middle|\, T_\sharp \alpha = \beta \right\}.$$

D'après le théorème de Brenier, on sait que l'optimum est atteint par la fonction de répartition inverse de β, c'est-à-dire $T(x) = F_\beta^{-1}(x)$, où $F_\beta(x) = \int_{-\infty}^x \rho(t)dt$. Dans le cas où β est discrète, cela revient à prendre $\Omega_j = \Big[\sum_{k=1}^{j-1} b_k, \sum_{k=1}^{j} b_k\Big)$ et $T(x) = y_j$ pour $x \in \Omega_j$.

1.3.4

Dans ce cas, le transport optimal assigne chaque segment à son centre. Puisque l'on a $\sum_{j=1}^n b_j = j/n$, on trouve :

$$W_2^2(\alpha,\beta) = \int_0^1 \min_j |x - T(x)|^2 \, \mathrm{d}x = \sum_{j=1}^n \int_{\frac{j-1}{n}}^{\frac{j}{n}} |x - y_j|^2 \, \mathrm{d}x = \sum_{j=1}^n \int_{\frac{j-1}{n}}^{\frac{j}{n}} \left|x - \frac{j}{n}\right|^2 \, \mathrm{d}x = \frac{1}{12n^2}.$$

1.4 Uniform laws on intervals with quadratic cost

1.4.1

On cherche T de la forme $T(x) = ux + v$. Puisque $T_\sharp \mathcal{U}_s = \mathcal{U}_{T(s)}$, on envoie a sur c et b sur d. On a donc :

$$T(x) = \frac{c-d}{a-b} x + \frac{c(d-b)}{a-b}.$$

1.4.2

On calcule, et on doit obtenir :

$$W_2^2(\alpha,\beta) = \int_a^b |x - T(x)|^2 \, \mathrm{d}x = (m_\alpha - m_\beta)^2 + \frac{1}{12}(\ell_\alpha - \ell_\beta)^2,$$

où $m_\alpha = (a+b)/2$ et $\ell_\alpha = b - a$ (et de même pour β).

1.5 Many Monge solutions for $|x-y|$

1.5.1

On commence par montrer que $W_1(\alpha,\beta) \geqslant 1$:

$$W_1(\alpha,\beta) = \inf_{T_\sharp\alpha=\beta} \int |x-T(x)| d\alpha(x) \geqslant \left| \int x \, \mathrm{d}\alpha(x) - \int y \, \mathrm{d}\beta(y) \right| = \frac{1}{2} \left| \frac{1}{2} 2^2 - \frac{1}{2}(3^2-1) \right| = 1.$$

On peut vérifier $T_\sharp \alpha = \beta$. De plus, un calcul direct montre que T_1 a un coût de 1, donc que $W_1(\alpha,\beta) \leqslant 1$. Le coût optimal est donc 1, et T_1 est une application de Monge optimale.

1.5.2

Un calcul direct montre que T_2 a un coût de 1, donc que T_2 est aussi optimale.

1.5.3

On pose pour $\theta \in [1,2]$:

$$T_\theta(x) = \begin{cases} x+\theta & x \in [0,\theta] \\ x+2-\theta & x \in (\theta, 2] \end{cases}$$

On a alors pour tout $\theta \in [1,2]$ que le coût de T_θ est 1, et que $T_{\theta\sharp}\alpha = \beta$. On a donc une infinité d'applications de Monge optimales.

1.6 Affine images

1.6.1 1-D case

D'après le théorème de Brenier, puisque la fonction de coût est $\|x-y\|^2$, le transport optimal existe, est unique, et est le gradient d'une fonction convexe. En posant $T(x) = ax+b$, on remarque que $T_\sharp\mu = \nu$ par définition, et que T est bien le gradient d'une fonction convexe, car T est croissante ($a>0$). T est donc le transport optimal. On a alors :

$$\begin{aligned} W_2^2(\mu,\nu) &= \int_{\mathbb{R}} (x-(ax+b))^2 d\mu(x) = \mathbb{E}[((1-a)X-b)^2] \\ &= \mathbb{E}[((1-a)(X-m_\mu) - (1-a)m_\mu - b)^2] = (1-a)^2\mathbb{V}(X) + ((1-a)m_\mu - b)^2. \end{aligned}$$

1.6.2 Higher Dimension

En appliquant à nouveau Brenier, on obtient que $T = \nabla\varphi$ pour φ convexe. Puisque $\nabla^2\varphi = A$, on en déduit par caractérisation des fonctions convexes que A est semi-définie positive ($A \succeq 0$). On a alors :

$$\begin{aligned} W_2^2(\mu,\nu) &= \int_{\mathbb{R}} \|x-(Ax+b)\|^2 \, \mathrm{d}\mu(x) = \mathbb{E}\left[\|(I_d-A)X-b\|^2\right] \\ &= \mathbb{E}\left[\|(I_d-A)(X-m_\mu) - (I_d-A)m_\mu - b\|^2\right] \\ &= \mathbb{E}\left[\mathrm{Tr}\left((x-m_\mu)^\top (I_d-A)^\top (I_d-A)(x-m_\mu)\right)\right] + \|b+(I_d-A)m_\mu\|^2 \\ &= \mathbb{E}\left[\mathrm{Tr}\left((I_d-A)^\top (x-m_\mu)^\top (x-m_\mu)(I_d-A)\right)\right] + \|b+(I_d-A)m_\mu\|^2 \\ &= \mathrm{Tr}\left((I_d-A)^\top \Sigma_\mu (I_d-A)\right) + \|b+(I_d-A)m_\mu\|^2. \end{aligned}$$

1.7 Piecewise constant densities on $[0, 1)$

Le coût est quadratique, donc T est croissante par Brenier. On trouve alors :

$$T(x) = \begin{cases} \frac{x}{2} & \text{si } x \in [0, \frac{1}{2}] \\ x - \frac{1}{4} & \text{si } x \in [\frac{1}{2}, \frac{3}{4}] \\ 2x - \frac{1}{4} & \text{si } x \in [\frac{3}{4}, 1] \end{cases}$$

On peut alors calculer le coût :

$$W_2^2(\alpha, \beta) = \int_{\mathbb{R}} |x - T(x)|^2 \, \mathrm{d}\alpha(x) = \int_0^1 |x - T(x)|^2 \frac{1}{3/2} \, \mathrm{d}x$$

1.8 Radial symmetry

1.8.1

On commence par montrer que $QTQ^{-1} = T$. On a $QTQ_\sharp^{-1}\alpha = QT_\sharp\alpha = Q_\sharp\beta = \beta$. De plus, le coût associé est :

$$\int \left\|x - QTQ^{-1}(x)\right\|^2 \, \mathrm{d}\alpha(x) = \int \|Q(x) - QT(x)\| \, |\det(Q^{-1})| \, \mathrm{d}Q_\sharp^{-1}\alpha(x) = \int \|x - T(x)\|^2 \, \mathrm{d}\alpha(x)$$

où l'on a utilisé que $Q_\sharp^{-1}\alpha = \alpha$, que $\|Q(x) - Q(y)\|^2 = \|x - y\|^2$, puisque Q est une isométrie donc préserve le produit scalaire. Ainsi, QTQ^{-1} est une application de transport optimale entre α et β. Par unicité du transport, on a $QTQ^{-1} = T$.

Montrons désormais que T se met sous la forme $T(x) = \frac{x}{\|x\|} t\left(\|x\|\right)$. Puisque T commute avec les rotations, on a toujours $T(\Lambda) \subseteq \Lambda$ pour une droite Λ. En effet, toute rotation fixant x va fixer $T(x)$. Ainsi, on sait que toujours, $T(x) = t(x)x$ où $t(x)$ est un scalaire. Mais on sait de plus que t ne peut pas dépendre de la direction de x, car sinon, T ne commuterait pas avec les rotations (on le voit géométriquement en considérant le facteur d'agrandissement selon certaines droites). Ainsi :

$$T(x) = \hat{t}(\|x\|)x = t(\|x\|)\frac{x}{\|x\|}$$

1.8.2

$$F_\beta(\rho) = F_{T_\sharp\alpha}(\rho) = \mathbb{P}(\|T(X)\| \leqslant \rho) = \mathbb{P}(t(\|X\|) \leqslant \rho) = \mathbb{P}(\|X\| \leqslant t^{-1}(\rho)) = F_\alpha(t^{-1}(\rho))$$

1.8.3

Dans ce cas, on a $F_\alpha(r) = (r/R_\alpha)^2$, et $F_\beta^{-1}(x) = R_\beta\sqrt{x}$. Ainsi, $t(r) = F_\beta^{-1}(F_\alpha(r)) = rR_\beta/R_\alpha$, et :

$$T(x) = \frac{R_\beta}{R_\alpha}x$$

1.9 Translation decomposition for W_2

1.9.1

Puisque les distributions α et β ont des moments d'ordre 2 finis, en particulier, elles ont des moments d'ordre 1 finis μ_α et μ_β dans $\mathbb{R}^d$. En considérant a et b l'opposé de ces moments, on vérifie :

$$\int x \mathrm{d}\alpha_0(x) = \int x \mathrm{d}\alpha(x + \mu) \overset{x:=x-\mu}{=} \int (x) \mathrm{d}\alpha(x) = \int x \mathrm{d}\alpha(x) - \mu = \mu_\alpha - \mu$$

On procède de la même manière pour β. L'équation ci dessus montre qu'on a 0 si et seulement si on prend $\mu = \mu_\alpha = a$.

1.9.2

On se donne un couplage π de α et β et on considère le couplage poussé $\pi_0 = (T_{-a} \times T_{-b})_\sharp \pi$. C'est un couplage de α_0 et β_0. Pour x, y dans le support de π on écrit $x = (x-a) + a$ et $y = (y-b) + b$. On a :

$$\|x-y\|^2 = \|(x-a)-(y-b)+(a-b)\|^2 = \|(x-a)-(y-b)\|^2 + 2(a-b)\left((x-a)-(y-b)\right) + \|a-b\|^2$$

On intègre ceci sur π, le terme produit étant nul car :

$$2(a-b)\int \left((x-a)-(y-b)\right) \mathrm{d}\pi(x,y) = 2(a-b) \times \left(\int (x-a)\mathrm{d}\alpha - \int (y-b)\mathrm{d}\beta\right) = 2(a-b)(0-0) = 0$$

On obtient alors :

$$\int \|x-y\|^2 \,\mathrm{d}\pi \overset{x=x-a}{\underset{y=y-b}{=}} \int \|x-y\|^2 \,\mathrm{d}\pi_0 + \|a-b\|^2$$

En passant à la borne inférieure, on a bien la décomposition :

$$W_2^2(\alpha,\beta) = \|a-b\|^2 + W_2^2(\alpha_0,\beta_0)$$

Par ailleurs, on voit aisément qu'un couplage optimal pour α_0, β_0 passe à un couplage optimal pour α, β en le poussant par $(T_a \times T_b)$.

1.9.3

En prenant T_0 le transport de Brenier de α_0 à β_0 et en définissant :

$$T(x) = T_0(x-a) + b$$

on voit que $x \sim \alpha$ implique $T(x) \sim \beta$ et donc que $T_\sharp \alpha = \beta$.

De plus, si $T_0 = \nabla \varphi_0$ où φ_0 est convexe, $T(x) = \nabla\left(\varphi_0(x-a) + bx\right)$ est le gradient d'une fonction convexe. Par conséquent, T est le transport de Brenier de α à β.

1.10 Adding separable terms ; quadratic cost vs. $-\langle x \,|\, y \rangle$

1.10.1

On calcule le coût associé à $\tilde{c}$:

$$\begin{aligned}\int \tilde{c}(x, T(x))\,\mathrm{d}\alpha(x) &= \int c(x, T(x))\,\mathrm{d}\alpha(x) + \int u(x)\,\mathrm{d}\alpha(x) + \int v(T(x))\,\mathrm{d}\alpha(x) \\ &= \int c(x, T(x))\,\mathrm{d}\alpha(x) + \int u(x)\,\mathrm{d}\alpha(x) + \int v(y)\,\mathrm{d}\beta(y),\end{aligned}$$

où la dernière égalité vient du fait que $T_\# \alpha = \beta$. Ainsi, le coût associé à $\tilde{c}$ diffère de celui associé à c d'une constante indépendante du transport T, donc les T optimaux sont les mêmes.

1.10.2

Le coût associé diffère dans ce cas d'une constante multiplicatrice positive indépendante du transport T, donc les T optimaux sont les mêmes.

1.10.3

Puisque $c_{\text{quad}}(x,y) = \|x\|^2 + \|y\|^2 - 2\langle x \,|\, y\rangle$, on peut écrire $c_{\text{quad}}(x,y) = \lambda c_{\text{in}}(x,y) + u(x) + v(y)$ avec $u(x) = \|x\|^2$, $v(y) = \|y\|^2$, et $\lambda = 2 > 0$. Par les questions précédentes, les transports optimaux pour c_{quad} et c_{in} sont les mêmes.

1.11 Quantile coupling : explicit computation

$F_\mu(x) = \int_0^x \mathrm{d}x = x$ et $F_\nu(y) = \int_0^y 2t \,\mathrm{d}t = y^2$ pour $x, y \in [0,1]$. On en déduit que $F_\mu^{-1}(x) = x$ et $F_\nu^{-1}(y) = \sqrt{y}$. Par monotonicité du transport en 1D, on a donc $T = F_\nu^{-1} \circ F_\mu$, soit $T(x) = \sqrt{x}$. En calculant $\int_0^1 |x - T(x)|^2 \,\mathrm{d}x$, on retrouve bien $\int_0^1 |F_\mu^{-1}(x) - F_\nu^{-1}(x)|^2 \,\mathrm{d}x = 1/30$.

1.12 Cumulative functions

1.12.1

Let $X \sim \alpha$ be a random variable with cumulative distribution function C_α. We want to show that $C_\alpha(X)$ is uniformly distributed on $[0,1]$. To prove this, we compute the cumulative distribution function of $C_\alpha(X)$.

For any $t \in [0,1]$,

$$P(C_\alpha(X) \leqslant t) = P(\alpha((-\infty, C_\alpha^{-1}(t)])) = P(X \leqslant C_\alpha^{-1}(t)).$$

Since $C_\alpha(x) = \alpha((-\infty, x])$, we have

$$C_\alpha^{-1}(t) = \inf\{x \in \mathbb{R} : C_\alpha(x) \geqslant t\}.$$

Now, since α has a density with respect to the Lebesgue measure, we know that C_α is continuous. Therefore, the random variable $C_\alpha(X)$ has the distribution of the uniform random variable on $[0,1]$, meaning

$$P(C_\alpha(X) \leqslant t) = t, \quad \text{for all } t \in [0,1].$$

Thus, $(C_\alpha)\sharp\alpha = U[0,1]$, and the density assumption (continuity of α) is essential for the uniform distribution.

1.12.2

Now let $\alpha = \sum_i p_i \delta_{x_i}$. The cumulative distribution function $C_\alpha(x)$ for such a discrete measure is given by

$$C_\alpha(x) = \sum_{i : x_i \leqslant x} p_i.$$

The pushforward of $X \sim \alpha$ under C_α is given by

$$(C_\alpha)\sharp\alpha = \sum_i p_i \delta_{C_\alpha(x_i)}.$$

For $X \sim \alpha$, we have $P(C_\alpha(X) \leqslant t) = P(\alpha((-\infty, C_\alpha^{-1}(t)]))$. Since C_α is a step function with jumps at the points x_i, the probability distribution of $C_\alpha(X)$ will have jumps at the points $C_\alpha(x_i)$, and the probability mass will be concentrated at these jump points. More precisely, for any $t \in [0,1]$,

$$P(C_\alpha(X) \leqslant t) = \sum_{i : C_\alpha(x_i) \leqslant t} p_i.$$

The probability mass is distributed at the jump points of C_α, and for any t corresponding to a jump point, the probability will include equality at that jump.

Thus, for a discrete measure α, $C_\alpha(X)$ is a discrete random variable taking values in the set of jump points of C_α, with probabilities corresponding to the mass of α at those points.

1.12.3

1.13 Pushforward of a uniform law by a nonlinear map

1.13.1

Let $X \sim \alpha = U[-1, 1]$ and $Y = T(X) = X^2$.

Since X is uniformly distributed on $[-1, 1]$, its probability density function is given by :

$$f_X(x) = \frac{1}{2}, \quad x \in [-1, 1].$$

We define $Y = X^2$. The values of Y range from 0 to 1, since $Y = X^2$ is always non-negative. Now, to find the distribution of Y, we use the change of variable formula. For $T(x) = x^2$, we find the density of Y by noting that the probability density function $f_Y(y)$ satisfies

$$f_Y(y) = \sum_{x:T(x)=y} \frac{f_X(x)}{|T'(x)|},$$

where $T'(x) = 2x$. Since $T(x) = x^2$ is a two-to-one function on the interval $[-1, 1]$, we need to account for both $x = \sqrt{y}$ and $x = -\sqrt{y}$, both of which map to the same y. Thus, we have :

$$f_Y(y) = \frac{f_X(\sqrt{y})}{|2\sqrt{y}|} + \frac{f_X(-\sqrt{y})}{|2(-\sqrt{y})|} = 2 \cdot \frac{\frac{1}{2}}{2\sqrt{y}} = \frac{1}{2\sqrt{y}}, \quad y \in [0, 1].$$

Therefore, the density of $\beta = T\#\alpha$ is :

$$f_Y(y) = \frac{1}{2\sqrt{y}}, \quad y \in (0, 1].$$

1.13.2

We have

$$\int_0^1 f_Y(y)\, dy = \int_0^1 \frac{1}{2\sqrt{y}}\, dy.$$

We make the substitution $u = \sqrt{y}$, so $du = \frac{1}{2\sqrt{y}}\, dy$, and the limits of integration change accordingly : when $y = 0$, $u = 0$, and when $y = 1$, $u = 1$. Thus, the integral becomes :

$$\int_0^1 \frac{1}{2\sqrt{y}}\, dy = \int_0^1 du = 1.$$

2 Discrete Kantorovich Formulation

2.1 1D discrete $\implies$ optimal Kantorovich plan

La matrice de coût C ayant toutes ses lignes identiques (égales à $\begin{bmatrix}(0.1)^2 & (0.2)^2 & (0.1)^2\end{bmatrix}$), tous les couplages ont le même coût. La matrice P est donc :

$$P = \begin{bmatrix} 1 & 0.1 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0.1 & 1 \end{bmatrix}$$

2.2 Parametric 1D cases

2.2.1

We have $a = (1+\varepsilon, 1-\varepsilon)$ and $b = (1,1)$ with $|\varepsilon| \leqslant 1$.

The optimal transport plan $P \in \mathbb{R}_+^{2\times 2}$ must satisfy :

— Row constraints : $P_{11} + P_{12} = 1 + \varepsilon$, $P_{21} + P_{22} = 1 - \varepsilon$
— Column constraints : $P_{11} + P_{21} = 1$, $P_{12} + P_{22} = 1$

Case 1 : $\varepsilon \geqslant 0$ The optimal plan is :

$$P = \begin{pmatrix} 1 & \varepsilon \\ 0 & 1-\varepsilon \end{pmatrix}$$

Case 2 : $\varepsilon < 0$ The optimal plan is :

$$P = \begin{pmatrix} 1+\varepsilon & 0 \\ -\varepsilon & 1 \end{pmatrix}$$

2.2.2

We have $a = (1+\varepsilon, 1+\eta, 1-\varepsilon-\eta)$ and $b = (1,1,1)$ with $1+\varepsilon \geqslant 0$, $1+\eta \geqslant 0$, $1-\varepsilon-\eta \geqslant 0$.

Case : $\varepsilon > 0$**,** $\eta > 0$ The optimal monotone plan is :

$$P = \begin{pmatrix} 1 & \varepsilon & 0 \\ 0 & 1-\varepsilon & \eta \\ 0 & 0 & 1-\varepsilon-\eta \end{pmatrix}$$

Case (ii) : $\varepsilon < 0$**,** $\eta > 0$ The optimal monotone plan is :

$$P = \begin{pmatrix} 1+\varepsilon & 0 & 0 \\ -\varepsilon & 1 & \eta \\ 0 & 0 & 1-\varepsilon-\eta \end{pmatrix}$$

2.3 Couplings in basic cases

2.3.1

Le seul couplage entre δ_x et δ_y est la matrice I_1.

2.3.2

Le seul couplage entre δ_x et $\frac{1}{2}(\delta_{y_1} + \delta_{y_2})$ est la matrice $P = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$.

2.3.3

L'ensemble des couplages entre α et β de masses distinctes est vide. En effet, si P est un tel couplage, on doit avoir $\sum_j P_{ij} = \alpha_i$ et $\sum_i P_{ij} = \beta_j$. En sommant sur i puis sur j, on obtient $\sum_i \alpha_i = \sum_j \beta_j$, ce qui contredit l'hypothèse.

2.4 Discrete α_t from a coupling

2.4.1

Montrons que $\alpha_0 = (P_0)_\sharp \pi$. Soit h une fonction continue sur $\mathbb{R}^d$. On a :

$$\begin{aligned}
\int_{\mathbb{R}^d\times\mathbb{R}^d} h(P_0(x,y))\,\mathrm{d}\pi(x,y) &= \int_{\mathbb{R}^d\times\mathbb{R}^d} h(x)\,\mathrm{d}\pi(x,y) \\
&= \sum_i \sum_j h(x_i) P_{i,j} \\
&= \sum_i h(x_i) \left(\sum_j P_{i,j} \right) \\
&= \sum_i h(x_i) a_i \\
&= \int_{\mathbb{R}^d} h(x)\,\mathrm{d}\alpha_0(x).
\end{aligned}$$

On montre de la même façon que $\alpha_1 = (P_1)_\sharp \pi$.

2.4.2

On a :

$$\alpha_t = \sum_{i,j} P_{i,j} \delta_{(1-t)x_i + t y_j},$$

donc α_t est discrète pour tout $t \in [0,1]$, et contient au plus nm masses de Dirac.

2.4.3

Voir ci-dessus.

2.4.4

Supposons $n = m$. Pour que α_t ait exactement n masses de Dirac pour tout $t \in [0,1]$, il faut que le couplage optimal P soit tel que tous les points $(1-t)x_i + t y_j$ avec $P_{i,j} > 0$ soient distincts pour tout $t \in [0,1]$.

Une condition suffisante est que P soit une matrice de permutation, c'est-à-dire qu'il existe une permutation $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ telle que :

$$P_{i,j} = \begin{cases} a_i & \text{si } j = \sigma(i) \\ 0 & \text{sinon} \end{cases}$$

Dans ce cas, pour que cela soit un couplage optimal, il faut que la permutation σ minimise le coût total :

$$\sum_{i=1}^n a_i c(x_i, y_{\sigma(i)}).$$

Exemple concret : Prenons $n = m = 2$, avec :

$$\begin{aligned}
\alpha_0 &= \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1, \\
\alpha_1 &= \frac{1}{2}\delta_0 + \frac{1}{2}\delta_2.
\end{aligned}$$

Donc $x_1 = 0$, $x_2 = 1$, $y_1 = 0$, $y_2 = 2$, et $a_1 = a_2 = b_1 = b_2 = \frac{1}{2}$.

Pour le coût euclidien $c(x, y) = |x - y|^2$, la matrice de coût est :

$$C = \begin{pmatrix} 0 & 4 \\ 1 & 1 \end{pmatrix}.$$

Le couplage optimal est :

$$P^* = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

qui est une matrice diagonale (un type de matrice de permutation).

L'interpolation correspondante est :

$$\alpha_t = \frac{1}{2}\delta_{(1-t)\cdot 0 + t\cdot 0} + \frac{1}{2}\delta_{(1-t)\cdot 1 + t\cdot 2} = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_{1+t}.$$

Pour tout $t \in [0, 1]$, les points 0 et $1 + t$ sont distincts, donc α_t a exactement $n = 2$ masses de Dirac pour tout $t \in [0, 1]$.

2.5 When does a finite-cost coupling exist ?

2.5.1 Dimension 2 by hand

C'est possible exclusivement dans 6 cas sur 16, à savoir lorsque le support de π couvre au moins une fois chaque ligne et chaque colonne de la matrice C.

2.5.2 General uniform case

Soit P un couplage de coût fini. On définit $Q := nP$ qui est doublement stochastique. Par théorème de Birkhoff-von Neumann, Q s'écrit comme une combinaison convexe de matrices de permutation, soit :

$$nP = \sum_k \lambda_k P_k$$

avec $\lambda_k > 0$, et P_k des matrices de permutation. Soit σ la permutation associée à P_1. Alors $[nP]_{j,\sigma(j)} = \sum_i \lambda_i [P_i]_{j,\sigma(j)} \geqslant \lambda_1 > 0$, donc $[P]_{j,\sigma(j)} > 0$. Puisque le coût de P est fini, on a donc $C_{j,\sigma(j)} < +\infty$. σ convient donc.

Réciproquement, si σ est une permutation telle que $C_{j,\sigma(j)} < +\infty$ pour tout j, alors la matrice de permutation P associée à σ est un couplage de coût fini.

3 Continuous Kantorovich Formulation

3.1 Couplings in base cases

3.1.1

Let α, β be finite positive measures with *different* total masses. We want to determine $\Pi(\alpha, \beta)$, the set of couplings.

Recall that a coupling between α and β is a measure π on the product space such that the marginals satisfy :

$$\begin{aligned} \pi(A \times Y) &= \alpha(A) \quad \text{for all measurable } A, \\ \pi(X \times B) &= \beta(B) \quad \text{for all measurable } B. \end{aligned}$$

For such a coupling to exist, we need :

$$\pi(X \times Y) = \alpha(X) = \beta(Y).$$

However, since α and β have different total masses, we have $\alpha(X) \neq \beta(Y)$, which means :

$$\Pi(\alpha, \beta) = \varnothing.$$

When α and β have different total masses, there are no couplings between them.

3.1.2

(i) Let δ_x be the Dirac measure at point x and $\mathcal{U}[0,1]$ be the uniform measure on $[0,1]$ (Lebesgue measure restricted to $[0,1]$).

We have :

$$\begin{aligned} \delta_x(X) &= 1, \\ \mathcal{U}[0,1]([0,1]) &= 1. \end{aligned}$$

So the total masses are equal, and couplings exist. A coupling $\pi \in \Pi(\delta_x, \mathcal{U}[0,1])$ must satisfy :

1. First marginal : $\pi(A \times [0,1]) = \delta_x(A)$ for all measurable A.
2. Second marginal : $\pi(X \times B) = \mathcal{U}[0,1](B) = |B|$ (Lebesgue measure of B) for all measurable $B \subseteq [0,1]$.

From the first marginal condition, all the mass of π must be concentrated on the set $\{x\} \times [0,1]$.

Therefore, π must have the form :

$$\pi = \delta_x \otimes \nu,$$

where ν is a probability measure on $[0,1]$.

From the second marginal condition :

$$\pi(X \times B) = (\delta_x \otimes \nu)(X \times B) = \delta_x(X) \cdot \nu(B) = \nu(B).$$

For this to equal $|B|$, we need $\nu = \mathcal{U}[0,1]$. There is thus a unique coupling :

$$\Pi(\delta_x, \mathcal{U}[0,1]) = \{\delta_x \otimes \mathcal{U}[0,1]\}.$$

Explicitly, for any measurable set $A \times B$:

$$\pi(A \times B) = \begin{cases} |B| & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

(ii) Let $\alpha = \frac{1}{2}(\delta_x + \delta_y)$ and $\beta = \mathcal{U}[0,1]$.

We have :

$$\begin{aligned} \alpha(X) &= 1, \\ \beta([0,1]) &= 1. \end{aligned}$$

So couplings exist. A coupling $\pi \in \Pi(\alpha, \beta)$ must satisfy :

1. First marginal : $\pi(A \times [0,1]) = \alpha(A) = \frac{1}{2}(\delta_x(A) + \delta_y(A))$ for all measurable A.
2. Second marginal : $\pi(X \times B) = |B|$ for all measurable $B \subseteq [0,1]$.

We can decompose π as :

$$\pi = \pi_x + \pi_y,$$

where π_x is concentrated on $\{x\} \times [0,1]$ and π_y is concentrated on $\{y\} \times [0,1]$.

From the first marginal condition :

$$\pi_x(\{x\} \times [0,1]) = \frac{1}{2},$$
$$\pi_y(\{y\} \times [0,1]) = \frac{1}{2}.$$

We can write :

$$\pi_x = \delta_x \otimes \nu_x,$$
$$\pi_y = \delta_y \otimes \nu_y,$$

where ν_x and ν_y are measures on $[0,1]$ with total mass $\frac{1}{2}$ each.

From the second marginal condition :

$$\pi(X \times B) = \nu_x(B) + \nu_y(B) = |B|.$$

This means $\nu_x + \nu_y = \mathcal{U}[0,1]$.

The set of couplings is thus :

$$\Pi\left(\frac{1}{2}(\delta_x + \delta_y), \mathcal{U}[0,1]\right) = \left\{\delta_x \otimes \nu_x + \delta_y \otimes \nu_y : \nu_x, \nu_y \geqslant 0,\ \nu_x + \nu_y = \mathcal{U}[0,1],\ \nu_x([0,1]) = \nu_y([0,1]) = \frac{1}{2}\right\}.$$

Intuitively, for any measurable partition A, B of $[0,1]$ with $|A| = |B| = \frac{1}{2}$:

$$\nu_x = \mathcal{U}|_A,$$
$$\nu_y = \mathcal{U}|_B.$$

3.2 Gaussian couplings in $\mathbb{R}$ (zero means)

3.2.1

The quadratic optimal transport cost is :

$$\mathbb{E}[(X-Y)^2],$$

where $(X,Y) \sim \pi = \mathcal{N}(0, \Sigma)$.

We have $X \sim \mathcal{N}(0, \sigma_\alpha^2)$, $Y \sim \mathcal{N}(0, \sigma_\beta^2)$, and $\mathrm{Cov}(X,Y) = c$.

Using the expansion :

$$\begin{aligned}
\mathbb{E}[(X-Y)^2] &= \mathbb{E}[X^2] - 2\mathbb{E}[XY] + \mathbb{E}[Y^2] \\
&= \mathbb{E}[X^2] - 2\mathrm{Cov}(X,Y) - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2] \\
&= \mathbb{E}[X^2] - 2\mathrm{Cov}(X,Y) + \mathbb{E}[Y^2] \quad (\text{since } \mathbb{E}[X] = \mathbb{E}[Y] = 0) \\
&= \mathrm{Var}(X) - 2c + \mathrm{Var}(Y) \\
&= \sigma_\alpha^2 - 2c + \sigma_\beta^2.
\end{aligned}$$

Finally :

$$\mathbb{E}[(X-Y)^2] = \sigma_\alpha^2 + \sigma_\beta^2 - 2c$$

3.2.2

For Σ to be positive semi-definite, all principal minors must be non-negative :

1. First order : $\sigma_\alpha^2 \geqslant 0$ and $\sigma_\beta^2 \geqslant 0$ (already given).
2. Second order :
$$\det(\Sigma) = \sigma_\alpha^2\sigma_\beta^2 - c^2 \geqslant 0.$$

From the determinant condition :
$$c^2 \leqslant \sigma_\alpha^2\sigma_\beta^2.$$

3.2.3

We minimize $\mathbb{E}[(X-Y)^2] = \sigma_\alpha^2 + \sigma_\beta^2 - 2c$ subject to the constraint $|c| \leqslant \sigma_\alpha\sigma_\beta$.

Since the cost is a decreasing function of c, we want to maximize c subject to the constraint.

The maximum value of c is :
$$c^* = \sigma_\alpha\sigma_\beta.$$

The optimal covariance matrix is thus :
$$\Sigma^* = \begin{pmatrix} \sigma_\alpha^2 & \sigma_\alpha\sigma_\beta \\ \sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{pmatrix}.$$

and the optimal cost is :
$$\begin{aligned} \mathbb{E}[(X-Y)^2]^* &= \sigma_\alpha^2 + \sigma_\beta^2 - 2\sigma_\alpha\sigma_\beta \\ &= (\sigma_\alpha - \sigma_\beta)^2. \end{aligned}$$

3.2.4

The rank of Σ depends on the determinant and the values of $\sigma_\alpha, \sigma_\beta, c$:

Case 1 : $\sigma_\alpha = \sigma_\beta = 0$

Then $\Sigma = 0$ (the zero matrix), so $\text{rank}(\Sigma) = 0$.

In this case, both α and β are δ_0, and $\pi = \delta_{(0,0)}$.

Case 2 : Exactly one of $\sigma_\alpha, \sigma_\beta$ is zero

Without loss of generality, suppose $\sigma_\alpha = 0$ and $\sigma_\beta > 0$. Then from $|c| \leqslant \sigma_\alpha\sigma_\beta = 0$, we must have $c = 0$.

Thus :
$$\Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_\beta^2 \end{pmatrix}, \quad \text{rank}(\Sigma) = 1.$$

The coupling is $\pi = \delta_0 \otimes \mathcal{N}(0, \sigma_\beta^2)$.

Case 3 : $\sigma_\alpha, \sigma_\beta > 0$ and $|c| < \sigma_\alpha\sigma_\beta$

Then $\det(\Sigma) = \sigma_\alpha^2\sigma_\beta^2 - c^2 > 0$, so Σ is positive definite and $\text{rank}(\Sigma) = 2$.

The support of π is all of $\mathbb{R}^2$ (since the Gaussian has full support).

Case 4 : $\sigma_\alpha, \sigma_\beta > 0$ and $|c| = \sigma_\alpha \sigma_\beta$ (the optimal case)

Then $\det(\Sigma) = 0$, so $\mathrm{rank}(\Sigma) = 1$ (since $\Sigma \neq 0$).

For $c = \sigma_\alpha \sigma_\beta$, we can factor :

$$\Sigma = \begin{pmatrix} \sigma_\alpha^2 & \sigma_\alpha \sigma_\beta \\ \sigma_\alpha \sigma_\beta & \sigma_\beta^2 \end{pmatrix} = \begin{pmatrix} \sigma_\alpha \\ \sigma_\beta \end{pmatrix} \begin{pmatrix} \sigma_\alpha & \sigma_\beta \end{pmatrix}.$$

This means Σ has rank 1, and the Gaussian is degenerate, concentrated on the line :

$$\{(x, y) \in \mathbb{R}^2 : y = \frac{\sigma_\beta}{\sigma_\alpha} x\}.$$

More precisely, we can write $(X, Y) = Z(\sigma_\alpha, \sigma_\beta)$ where $Z \sim \mathcal{N}(0, 1)$.

For $c = -\sigma_\alpha \sigma_\beta$, the support is on the line :

$$\{(x, y) \in \mathbb{R}^2 : y = -\frac{\sigma_\beta}{\sigma_\alpha} x\}.$$

In summary :

$$\mathrm{rank}(\Sigma) = \begin{cases} 0 & \text{if } \sigma_\alpha = \sigma_\beta = 0, \\ 1 & \text{if exactly one of } \sigma_\alpha, \sigma_\beta \text{ is zero, or } |c| = \sigma_\alpha \sigma_\beta, \\ 2 & \text{if } \sigma_\alpha, \sigma_\beta > 0 \text{ and } |c| < \sigma_\alpha \sigma_\beta. \end{cases}$$

Brenier's theorem states that for absolutely continuous measures α, β on $\mathbb{R}^d$ (with finite second moments), there exists a unique optimal transport map $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $T_\sharp \alpha = \beta$ and T is the gradient of a convex function.

Moreover, the optimal coupling is concentrated on the graph of T :

$$\pi^* = (Id \times T)_\sharp \alpha.$$

For one-dimensional Gaussians $\alpha = \mathcal{N}(0, \sigma_\alpha^2)$ and $\beta = \mathcal{N}(0, \sigma_\beta^2)$, the optimal transport map is :

$$T(x) = \frac{\sigma_\beta}{\sigma_\alpha} x.$$

and the optimal coupling is :

$$\pi^* = (Id \times T)_\sharp \alpha = \{(X, \tfrac{\sigma_\beta}{\sigma_\alpha} X) : X \sim \mathcal{N}(0, \sigma_\alpha^2)\}.$$

This is exactly the Gaussian coupling with $c = \sigma_\alpha \sigma_\beta$, which has rank 1 and is supported on the line $y = \frac{\sigma_\beta}{\sigma_\alpha} x$.

The optimal Gaussian coupling (Case 4 with $c = \sigma_\alpha \sigma_\beta$) thus matches Brenier's theorem :

- It is a deterministic coupling (rank 1, supported on a graph).
- The support is $\{(x, T(x)) : x \in \mathbb{R}\}$ where $T(x) = \frac{\sigma_\beta}{\sigma_\alpha} x$ is the optimal transport map.
- The map T is monotone (increasing when $\sigma_\beta / \sigma_\alpha > 0$) and is the gradient of the convex function $\varphi(x) = \frac{\sigma_\beta}{2\sigma_\alpha} x^2$.

3.3 Uniform -> two Diracs ; two Diracs -> uniform

3.3.1

For $T_\sharp \alpha = \beta$, we need T to map sets of measure $\frac{1}{2}$ to each of x and y.

By Brenier's theorem, the optimal map for quadratic cost is monotone. Assuming $x < y$:

$$\boxed{T(u) = \begin{cases} x & \text{if } u \in [0, \frac{1}{2}), \\ y & \text{if } u \in [\frac{1}{2}, 1]. \end{cases}}$$

The optimal cost is :

$$\begin{aligned} W_2^2(\alpha, \beta) &= \int_0^{1/2} |u - x|^2 \, du + \int_{1/2}^1 |u - y|^2 \, du \\ &= \frac{1}{3} - \frac{x}{4} - \frac{3y}{4} + \frac{x^2 + y^2}{2}. \end{aligned}$$

3.3.2

Any map $S : \mathbb{R} \to [0, 1]$ gives :

$$S_\sharp \beta = \frac{1}{2} \delta_{S(x)} + \frac{1}{2} \delta_{S(y)},$$

which is a discrete measure supported on at most 2 points. This cannot equal $\alpha = \mathcal{U}[0, 1]$, which is absolutely continuous with support on the entire interval $[0, 1]$.

3.3.3

The optimal coupling is given by the Monge map from the first question :

$$\pi^* = (Id \times T)_\sharp \alpha.$$

Explicitly :

$$\boxed{\pi^* = \mathcal{U}[0, \tfrac{1}{2}] \otimes \delta_x + \mathcal{U}[\tfrac{1}{2}, 1] \otimes \delta_y}$$

This transports the uniform distribution on $[0, \frac{1}{2}]$ entirely to x, and the uniform distribution on $[\frac{1}{2}, 1]$ entirely to y.

The coupling is supported on two horizontal line segments : $[0, \frac{1}{2}] \times \{x\}$ and $[\frac{1}{2}, 1] \times \{y\}$.

3.4 Bottleneck cost : metric property and closed-form OT value

3.4.1

We verify the metric axioms :

Positivity $d(x, y) \geqslant 0$ for all $x, y \in X$ since $m > 0$ everywhere.

Separation
- If $x = y$, then $d(x, y) = 0$ by definition.
- If $d(x, y) = 0$, then either $x = y$ (in which case we're done), or $x \neq y$ and $m(x) + m(y) = 0$. Since $m > 0$ everywhere, the latter is impossible. Thus $x = y$.

Symmetry $d(x, y) = m(x) + m(y) = m(y) + m(x) = d(y, x)$ for $x \neq y$, and $d(x, x) = 0 = d(x, x)$.

Triangle inequality : Suppose $x \neq z \neq y$:

If x, y, z are all distinct, then :

$$\begin{aligned} d(x,y) + d(y,z) &= [m(x) + m(y)] + [m(y) + m(z)] \\ &= m(x) + 2m(y) + m(z) \\ &\geqslant m(x) + m(z) = d(x,z). \quad (\text{since } m(y) > 0) \end{aligned}$$

Therefore, $d(x,y)$ is a metric on X.

3.4.2

We have :

$$W_c(\alpha, \beta) := \inf_{\pi \in \Pi(\alpha,\beta)} \int c(x,y)\, d\pi(x,y).$$

The cost for a coupling π is :

$$\begin{aligned} \int c(x,y)\, d\pi(x,y) &= \int_{x=y} 0\, d\pi(x,y) + \int_{x \neq y} [m(x) + m(y)]\, d\pi(x,y) \\ &= \int_{x \neq y} m(x)\, d\pi(x,y) + \int_{x \neq y} m(y)\, d\pi(x,y). \end{aligned}$$

Since

$$\begin{aligned} \int_{x \neq y} m(x)\, d\pi(x,y) &= \int m(x)\, d\pi(x,y) - \int_{x=y} m(x)\, d\pi(x,y) \\ &= \int m(x)\, d\alpha(x) - \int m(x)\, d\mu(x), \end{aligned}$$

where $\mu(A) := \pi(A \times A \cap \{(x,x) : x \in X\})$ is the mass on the diagonal, which satisfies $\mu \leqslant \alpha$ and $\mu \leqslant \beta$.

Similarly :

$$\int_{x \neq y} m(y)\, d\pi(x,y) = \int m(y)\, d\beta(y) - \int m(y)\, d\mu(y).$$

Therefore :

$$\int c(x,y)\, d\pi(x,y) = \int m\, d\alpha + \int m\, d\beta - 2 \int m\, d\mu.$$

To minimize this, we want to maximize $\int m\, d\mu$ subject to $\mu \leqslant \alpha$ and $\mu \leqslant \beta$.

The maximum is achieved when $\mu = \alpha \wedge \beta$ (the maximal common submeasure) :

$$\mu^*(A) := \min\{\alpha(A), \beta(A)\} \quad \text{for all } A \in \mathcal{B}.$$

Therefore :

$$\boxed{W_c(\alpha, \beta) = \int m\, d\alpha + \int m\, d\beta - 2 \int m\, d(\alpha \wedge \beta)}$$

3.4.3

The optimal coupling is constructed as follows. Let $\mu := \alpha \wedge \beta$ be the maximal common submeasure. Define :

$$
\begin{aligned}
\alpha^+ &:= \alpha - \mu, \quad \text{(excess mass of } \alpha\text{)} \\
\beta^+ &:= \beta - \mu. \quad \text{(excess mass of } \beta\text{)}
\end{aligned}
$$

The optimal plan is :

$$\boxed{\pi^* = \mu \otimes \delta_{\text{diag}} + \alpha^+ \otimes \beta^+}$$

More explicitly :

- The measure $\mu \otimes \delta_{\text{diag}}$ is supported on the diagonal $\{(x,x) : x \in X\}$ and satisfies $(\mu \otimes \delta_{\text{diag}})(A \times B) = \mu(A \cap B)$.
- The measure $\alpha^+ \otimes \beta^+$ is the product measure on $X \times X$.

To verify it actually is optimal, we compute the marginals

$$
\begin{aligned}
\pi^*(A \times X) &= \mu(A) + \alpha^+(A) \cdot \beta^+(X) = \mu(A) + \alpha^+(A) = \alpha(A), \\
\pi^*(X \times B) &= \mu(B) + \alpha^+(X) \cdot \beta^+(B) = \mu(B) + \beta^+(B) = \beta(B).
\end{aligned}
$$

and the total cost :

$$
\begin{aligned}
\int c(x,y)\, d\pi^*(x,y) &= \int_{\text{diag}} 0\, d(\mu \otimes \delta_{\text{diag}}) + \int_{X \times X} [m(x) + m(y)]\, d(\alpha^+ \otimes \beta^+) \\
&= 0 + \int m\, d\alpha^+ + \int m\, d\beta^+ \\
&= \int m\, d(\alpha - \mu) + \int m\, d(\beta - \mu) \\
&= \int m\, d\alpha + \int m\, d\beta - 2 \int m\, d\mu,
\end{aligned}
$$

which matches the optimal value from part (b).

3.5 Gaussian couplings in $\mathbb{R}^d$ (zero means)

3.5.1

We have :

$$\mathbb{E}\|X - Y\|^2 = \mathbb{E}\|X\|^2 + \mathbb{E}\|Y\|^2 - 2\mathbb{E}\langle X, Y\rangle = \mathrm{Tr}(\Sigma_\alpha) + \mathrm{Tr}(\Sigma_\beta) - 2\mathrm{Tr}(K).$$

3.5.2

Since $\Sigma_\alpha \succ 0$, Schur complement gives

$$\Sigma \succeq 0 \quad \Longleftrightarrow \quad \Sigma_\beta - K^\top \Sigma_\alpha^{-1} K \succeq 0.$$

Equivalently, with $M := \Sigma_\alpha^{-1/2} K \Sigma_\beta^{-1/2}$, this is $I - M^\top M \succeq 0$, i.e. $\|M\|_{\mathrm{op}} \leqslant 1$. Thus the feasible K are exactly

$$\boxed{K = \Sigma_\alpha^{1/2} M \Sigma_\beta^{1/2} \quad \text{with} \quad \|M\|_{\mathrm{op}} \leqslant 1.}$$

Another equivalent form is $K \Sigma_\beta^{-1} K^\top \preceq \Sigma_\alpha$.

3.5.3

Using the parametrization in (b),

$$\mathrm{Tr}(K) = \mathrm{Tr}(M\,S), \qquad S := \Sigma_\beta^{1/2}\Sigma_\alpha^{1/2}.$$

By von Neumann's trace inequality,

$$\max_{\|M\|_{\mathrm{op}}\leqslant 1} \mathrm{Tr}(MS) = \mathrm{Tr}\left((S^\top S)^{1/2}\right) = \mathrm{Tr}\left(\left(\Sigma_\beta^{1/2}\Sigma_\alpha\Sigma_\beta^{1/2}\right)^{1/2}\right),$$

achieved for $M = UV^\top$ when $S = U\,\mathrm{diag}(s)V^\top$ is an SVD. Therefore an optimal cross-covariance is

$$\boxed{K^* = \Sigma_\alpha^{1/2}\left(\Sigma_\alpha^{1/2}\Sigma_\beta\Sigma_\alpha^{1/2}\right)^{1/2}\Sigma_\alpha^{1/2}.}$$

Plugging into (a) yields the Gaussian-restricted Kantorovich value

$$\boxed{W_2^2(\alpha,\beta) = \mathrm{Tr}(\Sigma_\alpha) + \mathrm{Tr}(\Sigma_\beta) - 2\mathrm{Tr}\left(\left(\Sigma_\beta^{1/2}\Sigma_\alpha\Sigma_\beta^{1/2}\right)^{1/2}\right).}$$

This is attained by the deterministic linear map

$$\boxed{T(x) = Ax, \qquad A = \Sigma_\alpha^{-1/2}\left(\Sigma_\alpha^{1/2}\Sigma_\beta\Sigma_\alpha^{1/2}\right)^{1/2}\Sigma_\alpha^{-1/2},}$$

for which $\mathrm{Cov}(X, T(X)) = \Sigma_\alpha A^\top = \Sigma_\alpha A = K^*$.

3.5.4

Write the joint covariance as

$$\Sigma = \begin{pmatrix} \Sigma_\alpha & K \\ K^\top & \Sigma_\beta \end{pmatrix}.$$

— *Strict feasibility :* if $\Sigma_\beta - K^\top\Sigma_\alpha^{-1}K \succ 0$ (equivalently $\|M\|_{\mathrm{op}} < 1$), then $\Sigma \succ 0$ and $\mathrm{rank}(\Sigma) = 2d$. The Gaussian π has a smooth density on $\mathbb{R}^{2d}$ (full support).

— *At the optimizer $K = K^*$:* the Schur complement vanishes,

$$\Sigma_\beta - (K^*)^\top\Sigma_\alpha^{-1}K^* = 0,$$

so $\mathrm{rank}(\Sigma) = d$. In fact (X, Y) satisfies $Y = AX$ a.s., hence

$$\mathrm{supp}(\pi) = \{(x, Ax) : x \in \mathbb{R}^d\},$$

a d-dimensional linear subspace (the graph of T).

— *Brenier comparison :* since α is absolutely continuous, Brenier's theorem asserts that the *globally* W_2-optimal coupling is induced by a map $T = \nabla\varphi$. For Gaussians, T is linear, symmetric positive semidefinite, and equals the A above. Thus the Gaussian-restricted optimizer coincides with the Brenier map, and the joint law is concentrated on its graph.

3.6 A non-Monge optimal plan in $\mathbb{R}^2$

3.6.1

Write $x = (0, s)$ and $y = (\xi, t)$ with $\xi \in \{-1, +1\}$ and $s, t \in [0, 1]$. Then,

$$\|x - y\|^2 = (\xi - 0)^2 + (t - s)^2 = 1 + (t - s)^2 \geqslant 1.$$

Integrating over any coupling π gives $\int \|x - y\|^2 \, d\pi \geqslant 1$.

3.6.2

By construction, for any Borel $J \subset [0, 1]$,

$$\pi^*(S_- \times (\{-1\} \times J)) = \int_{s \in J} \frac{1}{2} \, ds = \frac{1}{2}|J|,$$

and similarly for S_+. Hence $(\mathrm{proj}_y)_\sharp \pi^* = \beta$, and trivially $(\mathrm{proj}_X)_\sharp \pi^* = \alpha$. Moreover, under π^* we have $t = s$ a.s., so

$$\int \|x - y\|^2 \, d\pi^* = \int \Big(1 + (t - s)^2\Big) \, d\pi^* = \int 1 \, d\pi^* = 1.$$

Combined with (a), this proves that π^* is optimal and the Kantorovich value equals 1.

3.6.3

Suppose $T_\sharp \alpha = \beta$ with $T(0, s) = (\xi(s), u(s))$, where $\xi(s) \in \{-1, +1\}$ and $u(s) \in [0, 1]$. Because β assigns mass $\frac{1}{2}$ to each vertical segment with *uniform* t-coordinate, we must have

$$\lambda(\{s : \xi(s) = -1\}) = \lambda(\{s : \xi(s) = +1\}) = \frac{1}{2},$$

and the pushforward of Lebesgue on $A_- := \{s : \xi(s) = -1\}$ by u is uniform on $[0, 1]$ (same for A_+). Now compute the cost :

$$\int \|x - T(x)\|^2 \, d\alpha(x) = \int_0^1 \Big(1 + (u(s) - s)^2\Big) \, ds = 1 + \int_0^1 (u(s) - s)^2 \, ds.$$

If this equalled 1, then $(u(s) - s)^2 = 0$ a.e., hence $u(s) = s$ a.e. But then the t-marginal on the left segment is the image of Lebesgue by $s \mapsto s$ restricted to A_-, i.e. Lebesgue on A_-, which is *not* uniform on $[0, 1]$ unless $A_- = [0, 1]$ (impossible since $\lambda(A_-) = \frac{1}{2}$). Therefore $\int_0^1 (u(s) - s)^2 \, ds > 0$, and so every Monge map has cost > 1. Consequently, no Monge map is optimal.

3.6.4

The optimizer π^* is supported on two graphs :

$$\mathrm{supp}(\pi^*) = \{((0, s), (-1, s)) : s \in [0, 1]\} \cup \{((0, s), (+1, s)) : s \in [0, 1]\}.$$

It is not contained in the graph of any single-valued map T, since for every $x = (0, s)$ the conditional law assigns positive mass to two distinct points $(-1, s)$ and $(+1, s)$. This does not contradict Brenier's theorem : that theorem requires the source to be absolutely continuous with respect to *Lebesgue measure on* $\mathbb{R}^2$. Here α is supported on a 1D set (a segment), so the assumption fails and an optimal plan need not be induced by a map.

4 Wasserstein Distances and duality

4.1 Total variation and 0/1 cost

Let $d(x,y) = \mathbf{1}_{x\neq y}$ be the discrete 0/1 distance, and let $\alpha = \sum_i a_i\delta_{x_i}$ and $\beta = \sum_j b_j\delta_{x_j}$ be probability measures on the same finite set $\{x_i\}_i$.

For any coupling $\pi \in \Pi(\alpha,\beta)$:

$$\begin{aligned}\int d(x,y)^p \, d\pi(x,y) &= \int \mathbf{1}_{x\neq y} \, d\pi(x,y) \\ &= \pi(\{(x,y) : x \neq y\}) \\ &= 1 - \pi(\{(x,y) : x = y\}) \\ &= 1 - \sum_i \pi(\{(x_i,x_i)\}).\end{aligned}$$

Let $m_i = \pi(\{(x_i,x_i)\})$ be the mass on the diagonal at point x_i. Then :

$$\int d(x,y)^p \, d\pi(x,y) = 1 - \sum_i m_i.$$

From the marginal constraints :

$$\begin{aligned}\sum_j \pi(\{(x_i,x_j)\}) &= a_i, \\ \sum_i \pi(\{(x_i,x_j)\}) &= b_j.\end{aligned}$$

The mass on the diagonal satisfies $m_i \leqslant \min(a_i,b_i)$ for all i.

To minimize the cost, we want to maximize $\sum_i m_i$. The optimal choice is :

$$m_i^* = \min(a_i,b_i).$$

Therefore :

$$\begin{aligned}W_p^p(\alpha,\beta) &= 1 - \sum_i \min(a_i,b_i) \\ &= 1 - \sum_i \min(a_i,b_i).\end{aligned}$$

Now recall the total variation distance :

$$\begin{aligned}\|\alpha-\beta\|_{\mathrm{TV}} &= \sup_{A\subseteq X} |\alpha(A)-\beta(A)| \\ &= \frac{1}{2}\sum_i |a_i-b_i| \\ &= \frac{1}{2}\left(\sum_i |a_i-b_i|\right).\end{aligned}$$

Using the identity $|a-b| = a+b-2\min(a,b)$ for non-negative a, b :

$$\begin{aligned}
\sum_i |a_i - b_i| &= \sum_i (a_i + b_i - 2\min(a_i, b_i)) \\
&= \sum_i a_i + \sum_i b_i - 2\sum_i \min(a_i, b_i) \\
&= 1 + 1 - 2\sum_i \min(a_i, b_i) \\
&= 2\left(1 - \sum_i \min(a_i, b_i)\right).
\end{aligned}$$

Therefore :

$$\|\alpha - \beta\|_{\mathrm{TV}} = \frac{1}{2} \cdot 2\left(1 - \sum_i \min(a_i, b_i)\right) = 1 - \sum_i \min(a_i, b_i).$$

Hence :

$$\boxed{W_p^p(\alpha, \beta) = \frac{1}{2}\|\alpha - \beta\|_{\mathrm{TV}}}$$

for any $p \geqslant 1$.

4.2 Equivalence of topologies on a finite discrete metric space

Let (X, d) be a finite discrete metric space with

$$d_{\min} := \min_{x \neq y} d(x, y), \qquad d_{\max} := \max_{x,y} d(x, y).$$

For any coupling $\pi \in \Pi(\alpha, \beta)$:

$$\begin{aligned}
W_1(\alpha, \beta) &= \inf_\pi \int d(x, y)\, d\pi(x, y) \\
&\geqslant \inf_\pi \int d_{\min} \mathbf{1}_{x \neq y}\, d\pi(x, y) \\
&= d_{\min} \inf_\pi \pi(\{(x, y) : x \neq y\}) \\
&= d_{\min} \cdot W_1^1(\alpha, \beta) \quad \text{(for 0/1 distance)} \\
&= d_{\min} \cdot \frac{1}{2}\|\alpha - \beta\|_{\mathrm{TV}} \\
&= \frac{d_{\min}}{2}\|\alpha - \beta\|_{\mathrm{TV}}.
\end{aligned}$$

Similarly :

$$\begin{aligned}
W_1(\alpha, \beta) &= \inf_\pi \int d(x, y)\, d\pi(x, y) \\
&\leqslant \inf_\pi \int d_{\max} \mathbf{1}_{x \neq y}\, d\pi(x, y) \\
&= d_{\max} \cdot \frac{1}{2}\|\alpha - \beta\|_{\mathrm{TV}} \\
&= \frac{d_{\max}}{2}\|\alpha - \beta\|_{\mathrm{TV}}.
\end{aligned}$$

Therefore :

$$\boxed{\frac{d_{\min}}{2}\|\alpha-\beta\|_{\mathrm{TV}} \leqslant W_1(\alpha,\beta) \leqslant \frac{d_{\max}}{2}\|\alpha-\beta\|_{\mathrm{TV}}}$$

Since $d_{\min}, d_{\max} > 0$ are fixed constants for a finite discrete metric space, the inequalities show that :

- $W_1(\alpha,\beta) \to 0$ if and only if $\|\alpha-\beta\|_{\mathrm{TV}} \to 0$.
- A sequence (α_n) is Cauchy in the W_1 metric if and only if it is Cauchy in the TV metric.

Therefore, the W_1-topology and the total variation topology coincide on $\mathcal{P}(X)$.

4.3 c-transforms

Recall the c-transform and $\bar{c}$-transform definitions :

$$\begin{aligned} f^c(y) &:= \inf_{x\in X}[c(x,y)-f(x)], \\ g^{\bar{c}}(x) &:= \inf_{y\in Y}[c(x,y)-g(y)]. \end{aligned}$$

4.3.1

If $f(x) \leqslant f'(x)$ for all $x \in X$, then :

$$\begin{aligned} f^c(y) &= \inf_{x\in X}[c(x,y)-f(x)] \\ &\geqslant \inf_{x\in X}[c(x,y)-f'(x)] \\ &= f'^c(y). \end{aligned}$$

Therefore, $f^c \geqslant f'^c$.

4.3.2

By definition :

$$\begin{aligned} f^{c\bar{c}}(x) &= \inf_{y\in Y}[c(x,y)-f^c(y)] \\ &= \inf_{y\in Y}\left[c(x,y) - \inf_{x'\in X}[c(x',y)-f(x')]\right] \\ &= \inf_{y\in Y}\sup_{x'\in X}[c(x,y)-c(x',y)+f(x')]. \end{aligned}$$

Taking $x' = x$:

$$f^{c\bar{c}}(x) \geqslant c(x,y) - c(x,y) + f(x) = f(x)$$

for any y. Therefore, $f^{c\bar{c}} \geqslant f$.

4.3.3

This follows by symmetry from part (b), exchanging the roles of c and $\bar{c}$.

4.3.4

From part (b), we have $f^{c\bar{c}} \geqslant f$.

Applying the c-transform and using part (a) :

$$(f^{c\bar{c}})^c \leqslant f^c.$$

That is, $f^{c\bar{c}c} \leqslant f^c$.

For the reverse inequality, let $h = f^c$. Then from part (c) :

$$h^{\bar{c}c} \geqslant h,$$

which gives $(f^c)^{\bar{c}c} \geqslant f^c$, i.e., $f^{c\bar{c}c} \geqslant f^c$.

Therefore, $f^{c\bar{c}c} = f^c$.

5 Sliced Optimal Transport

5.1 Comparing isotropic centered Gaussians

5.1.1

For any $\theta \in \mathbb{S}^{d-1}$, the projection $(p_\theta)_\sharp\mu$ is the distribution of $\langle\theta, X\rangle$ where $X \sim \mathcal{N}(0, \sigma^2 I_d)$.

Since $\langle\theta, X\rangle = \theta^\top X$ and $X \sim \mathcal{N}(0, \sigma^2 I_d)$, we have :

$$\langle\theta, X\rangle \sim \mathcal{N}(0, \sigma^2\|\theta\|^2) = \mathcal{N}(0, \sigma^2).$$

Similarly, $(p_\theta)_\sharp\nu = \mathcal{N}(0, \tau^2)$.

For one-dimensional centered Gaussians $\mathcal{N}(0, \sigma^2)$ and $\mathcal{N}(0, \tau^2)$, the Wasserstein-2 distance is :

$$W_2^2(\mathcal{N}(0, \sigma^2), \mathcal{N}(0, \tau^2)) = (\sigma - \tau)^2.$$

Since this is independent of θ, the Sliced-Wasserstein distance is :

$$\begin{aligned}\mathbf{SW}_2^2(\mu, \nu) &= \int_{\mathbb{S}^{d-1}} W_2^2((p_\theta)_\sharp\mu, (p_\theta)_\sharp\nu)\, d\sigma(\theta) \\ &= \int_{\mathbb{S}^{d-1}} (\sigma - \tau)^2\, d\sigma(\theta) \\ &= (\sigma - \tau)^2.\end{aligned}$$

5.1.2

For isotropic Gaussians $\mu = \mathcal{N}(0, \sigma^2 I_d)$ and $\nu = \mathcal{N}(0, \tau^2 I_d)$, the Wasserstein-2 distance can be computed using the formula from Exercise 2 of the Gaussian couplings :

$$\begin{aligned}W_2^2(\mu, \nu) &= \mathrm{Tr}(\Sigma_\mu) + \mathrm{Tr}(\Sigma_\nu) - 2\mathrm{Tr}\left(\left(\Sigma_\nu^{1/2}\Sigma_\mu\Sigma_\nu^{1/2}\right)^{1/2}\right) \\ &= \mathrm{Tr}(\sigma^2 I_d) + \mathrm{Tr}(\tau^2 I_d) - 2\mathrm{Tr}\left(\left((\tau^2 I_d)^{1/2}(\sigma^2 I_d)(\tau^2 I_d)^{1/2}\right)^{1/2}\right) \\ &= d\sigma^2 + d\tau^2 - 2\mathrm{Tr}\left((\sigma^2\tau^2 I_d)^{1/2}\right) \\ &= d\sigma^2 + d\tau^2 - 2d\sigma\tau \\ &= d(\sigma - \tau)^2.\end{aligned}$$

Therefore :

$$W_2(\mu,\nu) = \sqrt{d}\,|\sigma - \tau|.$$

Comparison :

$$\boxed{\mathbf{SW}_2(\mu,\nu) = \frac{1}{\sqrt{d}} W_2(\mu,\nu)}$$

For isotropic centered Gaussians, the Sliced-Wasserstein distance is exactly $1/\sqrt{d}$ times the Wasserstein distance.

5.2 Projected optimal transport plans

5.2.1

We just need to verify that γ_θ has the correct marginals.

$$\begin{aligned}
\text{First marginal of } \gamma_\theta &= (p_1)_\sharp \gamma_\theta \\
&= (p_1)_\sharp (p_\theta \otimes p_\theta)_\sharp \gamma \\
&= (p_\theta \circ p_1)_\sharp \gamma \\
&= (p_\theta)_\sharp (p_1)_\sharp \gamma \\
&= (p_\theta)_\sharp \mu,
\end{aligned}$$

where p_1 denotes projection onto the first coordinate.

Similarly, the second marginal of γ_θ is $(p_\theta)_\sharp \nu$. Therefore, $\gamma_\theta \in \Pi((p_\theta)_\sharp \mu, (p_\theta)_\sharp \nu)$.

5.2.2

We use the fact that projections are 1-Lipschitz, which implies they are cost-reducing, to prove the optimality. For any $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$:

$$\begin{aligned}
|p_\theta(x) - p_\theta(y)| &= |\langle \theta, x \rangle - \langle \theta, y \rangle| \\
&= |\langle \theta, x - y \rangle| \\
&\leqslant \|\theta\| \cdot \|x - y\| \\
&= \|x - y\|.
\end{aligned}$$

Therefore :

$$\begin{aligned}
W_2^2((p_\theta)_\sharp \mu, (p_\theta)_\sharp \nu) &\leqslant \int_{\mathbb{R}^d \times \mathbb{R}^d} |p_\theta(x) - p_\theta(y)|^2 \, d\gamma_\theta(s,t) \\
&= \int_{\mathbb{R}^d \times \mathbb{R}^d} |p_\theta(x) - p_\theta(y)|^2 \, d\gamma(x,y) \\
&\leqslant \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^2 \, d\gamma(x,y) \\
&= W_2^2(\mu,\nu).
\end{aligned}$$

But we also have :

$$\begin{aligned}
\int |s - t|^2 \, d\gamma_\theta(s,t) &= \int |p_\theta(x) - p_\theta(y)|^2 \, d\gamma(x,y) \\
&\leqslant \int \|x - y\|^2 \, d\gamma(x,y) \\
&= W_2^2(\mu,\nu).
\end{aligned}$$

Now, by the triangle inequality in reverse and optimality of projections for the quadratic cost on the line :

$$W_2^2((p_\theta)_\sharp\mu, (p_\theta)_\sharp\nu) = \int |s-t|^2 \, d\gamma_\theta(s,t),$$

which shows γ_θ is optimal.

5.3 Translation decomposition for $\mathbf{SW}_2$

5.3.1

For $\theta \sim \mathcal{U}_{\mathbb{S}^{d-1}}$ and $X \sim \mu$:

$$\begin{aligned}
\mathbb{E}_{s\sim(p_\theta)_\sharp\mu}[s] &= \mathbb{E}_{X\sim\mu}[p_\theta(X)] \\
&= \mathbb{E}_{X\sim\mu}[\langle\theta, X\rangle] \\
&= \langle\theta, \mathbb{E}_{X\sim\mu}[X]\rangle \\
&= \langle\theta, \mathbb{E}_{X\sim\mu}[X]\rangle.
\end{aligned}$$

Similarly, $\mathbb{E}_{t\sim(p_\theta)_\sharp\nu}[t] = \langle\theta, \mathbb{E}_{Y\sim\nu}[Y]\rangle$.

5.3.2

Let $u = \mathbb{E}_{x\sim\mu}[x]$ and $v = \mathbb{E}_{y\sim\nu}[y]$ be the means of μ and ν respectively.

Define $\mu_0 = (T_{-u})_\sharp\mu$ and $\nu_0 = (T_{-v})_\sharp\nu$, where $T_a(x) = x + a$ is translation by a.

Then μ_0 and ν_0 are centered : $\mathbb{E}_{x\sim\mu_0}[x] = 0$ and $\mathbb{E}_{y\sim\nu_0}[y] = 0$.

For $(p_\theta)_\sharp\mu$, the centered version is :

$$\begin{aligned}
((p_\theta)_\sharp\mu)_0 &= \text{centered version of } (p_\theta)_\sharp\mu \\
&= (T_{-m_\theta})_\sharp(p_\theta)_\sharp\mu,
\end{aligned}$$

where $m_\theta = \mathbb{E}_{s\sim(p_\theta)_\sharp\mu}[s] = \langle\theta, u\rangle$.

Now :

$$\begin{aligned}
(p_\theta)_\sharp\mu_0 &= (p_\theta)_\sharp(T_{-u})_\sharp\mu \\
&= (p_\theta \circ T_{-u})_\sharp\mu \\
&= (T_{-\langle\theta,u\rangle} \circ p_\theta)_\sharp\mu \\
&= (T_{-m_\theta})_\sharp(p_\theta)_\sharp\mu \\
&= ((p_\theta)_\sharp\mu)_0.
\end{aligned}$$

Similarly, $((p_\theta)_\sharp\nu)_0 = (p_\theta)_\sharp\nu_0$.

5.3.3

For any $\theta \in \mathbb{S}^{d-1}$:

$$\begin{aligned}
W_2^2((p_\theta)_\sharp\mu, (p_\theta)_\sharp\nu) &= W_2^2((p_\theta)_\sharp\mu_0 + \langle\theta, u\rangle, (p_\theta)_\sharp\nu_0 + \langle\theta, v\rangle) \\
&= |\langle\theta, u\rangle - \langle\theta, v\rangle|^2 + W_2^2((p_\theta)_\sharp\mu_0, (p_\theta)_\sharp\nu_0) \\
&= |\langle\theta, u - v\rangle|^2 + W_2^2((p_\theta)_\sharp\mu_0, (p_\theta)_\sharp\nu_0).
\end{aligned}$$

Integrating over $\theta \sim \mathcal{U}_{\mathbb{S}^{d-1}}$:

$$\mathbf{SW}_2^2(\mu,\nu) = \int_{\mathbb{S}^{d-1}} |\langle \theta, u-v\rangle|^2 \, d\sigma(\theta) + \mathbf{SW}_2^2(\mu_0,\nu_0).$$

For the first term, using the hint $\mathbb{E}_{\theta\sim\mathcal{U}_{\mathbb{S}^{d-1}}}[\theta\theta^\top] = \frac{1}{d} I_d$:

$$\begin{aligned}
\int_{\mathbb{S}^{d-1}} |\langle \theta, u-v\rangle|^2 \, d\sigma(\theta) &= \int_{\mathbb{S}^{d-1}} (u-v)^\top \theta\theta^\top (u-v)\, d\sigma(\theta) \\
&= (u-v)^\top \left(\int_{\mathbb{S}^{d-1}} \theta\theta^\top \, d\sigma(\theta)\right)(u-v) \\
&= (u-v)^\top \frac{1}{d} I_d (u-v) \\
&= \frac{1}{d}\|u-v\|^2.
\end{aligned}$$

Therefore :

$$\boxed{\mathbf{SW}_2^2(\mu,\nu) = \frac{1}{d}\|u-v\|^2 + \mathbf{SW}_2^2(\mu_0,\nu_0)}$$

6 Flow Matching and Dynamic Optimal Transport

6.1 Velocity of couplings

6.1.1

To prove it suffices to see that the diffeomorphism $(X,Y) \leftrightarrow (AX+b, AY+b)$ is an affine one.

As such, $\mathbb{E}[AX_1 + b - AX_0 - b \mid X_t' = x] = A\mathbb{E}[X_1 - X_0 \mid AX_t + b = x]$ where X_t' is the interpolation in this case. Since $AX_t + b = x \Leftrightarrow X_t = A^{-1}(x-b)$, we get the result.

6.1.2

Just computations to do with $(1-r)X_0 + rX_1 = \frac{x}{1-t+tc}$.

6.2 Transformation of transport plan

In Exercise 1, we considered three transformations of couplings :

(i) Affine transformation : $(AX_0 + b, AX_1 + b)$

(ii) Translation : $(X_0, X_1 + b)$

(iii) Scaling : (X_0, cX_1)

6.2.1

If the original coupling has optimal map $T^* : \alpha \to \beta$, consider the transformed measures :

$$\alpha' = (Ax+b)_\# \alpha, \quad \beta' = (Ay+b)_\# \beta$$

The optimal transport map between α' and β' is :

$$T'^*(x) = AT^*(A^{-1}(x-b)) + b$$

This can be verified by checking that $(T')_\# \alpha' = \beta'$:

$$(T')_\# \alpha' = (T')_\# [(Ax+b)_\# \alpha] = [A \circ T^* \circ A^{-1} \circ (\cdot - b) + b]_\# [(A \cdot + b)_\# \alpha]$$

By the push-forward property, this equals $(AT^* + b)_\# \alpha = \beta'$.

6.2.2

If the original coupling has optimal map $T^* : \alpha \to \beta$, consider :

$$\alpha' = \alpha, \quad \beta' = (y+b)_{\#}\beta$$

The optimal transport map becomes :

$$T'^*(x) = T^*(x) + b$$

This is simply a translation of the original map.

6.2.3

If the original coupling has optimal map $T^* : \alpha \to \beta$, consider :

$$\alpha' = \alpha, \quad \beta' = (cy)_{\#}\beta$$

The optimal transport map becomes :

$$T'^*(x) = cT^*(x)$$

This is a scaling of the original map by factor c.

The transformations of the optimal transport map follow naturally from the transformations of the measures :

- Affine : $T'^*(x) = AT^*(A^{-1}(x-b)) + b$
- Translation : $T'^*(x) = T^*(x) + b$
- Scaling : $T'^*(x) = cT^*(x)$

These results follow from the equivariance properties of optimal transport under measure transformations.

6.2.4 Flow is optimal

We need to show that if the velocity field $v_t(x) = \mathbb{E}[X_1 - X_0 | X_t = x]$ generates a unique flow ϕ_t, then ϕ_1 is an optimal transport map.

The flow ϕ_t satisfies the ODE :

$$\frac{d}{dt}\phi_t(x) = v_t(\phi_t(x)), \quad \phi_0(x) = x$$

We have supposed this flow is unique and well-defined on $[0,1]$.

For the coupling (X_0, X_1) with interpolation $X_t = (1-t)X_0 + tX_1$, we have :

$$\frac{d}{dt}X_t = X_1 - X_0$$

The velocity field at X_t is :

$$v_t(X_t) = \mathbb{E}[X_1 - X_0 | X_t] = \frac{d}{dt}X_t.$$

Let μ_t be the law of X_t. Then $\mu_0 = \alpha$ and $\mu_1 = \beta$.

The continuity equation states :

$$\frac{\partial \mu_t}{\partial t} + \nabla \cdot (v_t \mu_t) = 0$$

This means that $\mu_t = (\phi_t)_\# \mu_0 = (\phi_t)_\# \alpha$ for all $t \in [0, 1]$.

In particular :

$$\mu_1 = (\phi_1)_\# \alpha = \beta$$

So ϕ_1 is a transport map from α to β.

The optimal transport cost for the squared Euclidean distance can be written as, by Benamou-Brenier

$$W_2^2(\alpha, \beta) = \inf \int_0^1 \int_{\mathbb{R}^d} |v_t(x)|^2 d\mu_t(x)\, dt$$

where the infimum is over all velocity fields v_t and curves μ_t satisfying the continuity equation with $\mu_0 = \alpha$ and $\mu_1 = \beta$.

For our velocity field $v_t(x) = \mathbb{E}[X_1 - X_0 | X_t = x]$ and the coupling (X_0, X_1) :

$$\begin{aligned}\int_0^1 \int_{\mathbb{R}} |v_t(x)|^2 d\mu_t(x)\, dt &= \int_0^1 \mathbb{E}[|v_t(X_t)|^2]\, dt \\ &= \int_0^1 \mathbb{E}[|\mathbb{E}[X_1 - X_0 | X_t]|^2]\, dt\end{aligned}$$

By the tower property and Jensen's inequality :

$$\mathbb{E}[|\mathbb{E}[X_1 - X_0 | X_t]|^2] \leqslant \mathbb{E}[|X_1 - X_0|^2]$$

with equality when $X_1 - X_0$ is X_t-measurable, which happens when (X_0, X_1) is a deterministic coupling (i.e., $X_1 = T(X_0)$ for some map T).

If the velocity field generates a unique flow and the coupling satisfies the displacement interpolation property, then :

$$\int_0^1 \mathbb{E}[|v_t(X_t)|^2]\, dt = \mathbb{E}[|X_1 - X_0|^2]$$

This achieves the Benamou-Brenier formula, and therefore $\phi_1 = T^*$ is an optimal transport map.

6.3 Gaussian Flow

6.3.1

We have $v_t(x) = E[X_1 \mid X_t = x] - E[X_0 \mid X_t = x]$ which gives from the Gaussian vector formula for conditional expectancy, as $E[X_1] = E[X_0] = 0$:

$$\mathbb{E}[X_1 | X_t = x] = \text{Cov}(X_1, X_t)\text{Cov}(X_t, X_t)^{-1} x,$$

and similarly in X_0. We get :

$$\mathbb{E}[X_1 - X_0 | X_t = x] = \text{Cov}(X_1, X_t)\text{Cov}(X_t, X_t)^{-1} x - \text{Cov}(X_0, X_t)\text{Cov}(X_t, X_t)^{-1} x.$$

which is the wanted formula with the definition of Σ_t :

$$v_t(x) = \frac{1}{1-t}\Big(((1-t)\Sigma_0 + t\Sigma_1)\Sigma_t^{-1} - \text{Id}\Big)\, x.$$

6.3.2

If Σ_0 and Σ_1 can be jointly diagonalized, then there exists an orthogonal matrix Q such that :

$$\Sigma_0 = Q\Lambda_0 Q^T, \quad \Sigma_1 = Q\Lambda_1 Q^T,$$

where Λ_0 and Λ_1 are diagonal matrices. In this case, the velocity field $v_t(x)$ can be simplified, and the optimal transport map ϕ_t between α and β can be written as :

$$\phi_t(x) = \exp\left(\int_0^t v_s(x)ds\right),$$

which corresponds to the flow induced by the velocity field $v_t(x)$. Since the matrices Σ_0 and Σ_1 are diagonal, the optimal transport map between α and β can be directly derived using the flow ϕ_t defined above.

Thus, ϕ_t is the optimal transport map between α and β.

6.4 Velocity of gaussian mixtures

6.4.1

$$v_t(x) := \mathbb{E}[X_1 - X_0 | X_t = x].$$

From the mixture model assumption, we know that X_0 and X_1 are distributed as a mixture of Gaussian distributions. We need to compute the conditional expectation of $X_1 - X_0$ given $X_t = x$.

For each Gaussian component k, we can compute the conditional expectation of $X_1 - X_0$ given $X_t = x$. Since the joint distribution of X_0 and X_1 is Gaussian, we use the following conditional expectation formula :

$$\mathbb{E}[X_1 | X_t = x] = m_1^k + \Sigma_{1t}^k (\Sigma_t^k)^{-1}(x - m_t^k),$$
$$\mathbb{E}[X_0 | X_t = x] = m_0^k + \Sigma_{0t}^k (\Sigma_t^k)^{-1}(x - m_t^k),$$

where Σ_{0t}^k and Σ_{1t}^k are the covariances between X_0, X_t and X_1, X_t, and Σ_t^k is the covariance matrix of X_t.

The velocity field for each component k is :

$$v_t^k(x) = \mathbb{E}[X_1 - X_0 | X_t = x] = (m_1^k - m_0^k) + \Sigma_{1t}^k (\Sigma_t^k)^{-1}(x - m_t^k) - \Sigma_{0t}^k (\Sigma_t^k)^{-1}(x - m_t^k).$$

This simplifies to :

$$v_t^k(x) = (m_1^k - m_0^k) + (\Sigma_{1t}^k - \Sigma_{0t}^k)(\Sigma_t^k)^{-1}(x - m_t^k).$$

Now we need to compute the overall velocity field by considering the weighted sum of the contributions from each Gaussian component. The weight for each component is given by $\alpha^k(x)$, which is the posterior probability of the k-th Gaussian given $X_t = x$:

$$\alpha^k(x) = \frac{\pi_k p_t^k(x)}{\sum_{j=1}^K \pi_j p_t^j(x)}.$$

Thus, the overall velocity field is :

$$v_t(x) = \sum_{k=1}^K \alpha^k(x) v_t^k(x) = \sum_{k=1}^K \alpha^k(x) \left((m_1^k - m_0^k) + (\Sigma_{1t}^k - \Sigma_{0t}^k)(\Sigma_t^k)^{-1}(x - m_t^k)\right).$$

This gives the desired expression for the velocity field $v_t(x)$.

6.4.2

In this special case, if X_0 is a standard Gaussian, i.e., $X_0 \sim \mathcal{N}(0, I)$, and X_1 follows a discrete distribution with K Dirac masses, the distributions $\mathcal{N}(m^k, \Sigma^k)$ become degenerate (i.e., each mixture component is a Dirac delta function at m_1^k). In this case, the covariance matrices Σ_0^k and Σ_1^k reduce to simple matrices with Dirac mass characteristics, and the velocity field $v_t(x)$ takes a more simplified form :

$$v_t(x) = \sum_{k=1}^{K} \pi_k \left(m_1^k - m_0^k \right).$$

In this case, the velocity field is constant and does not depend on x, reflecting the fact that X_1 is a discrete distribution of K Dirac masses.