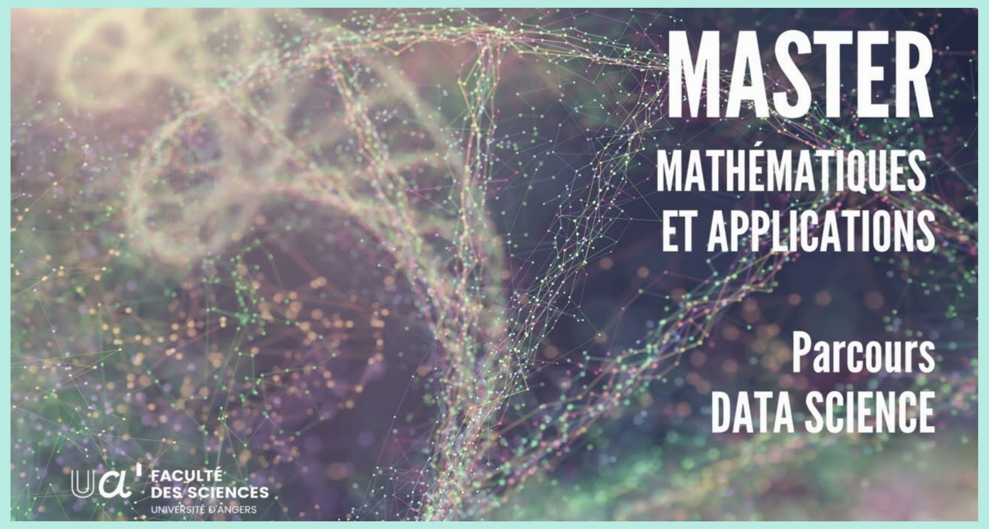


Transport optimal et applications

université angers

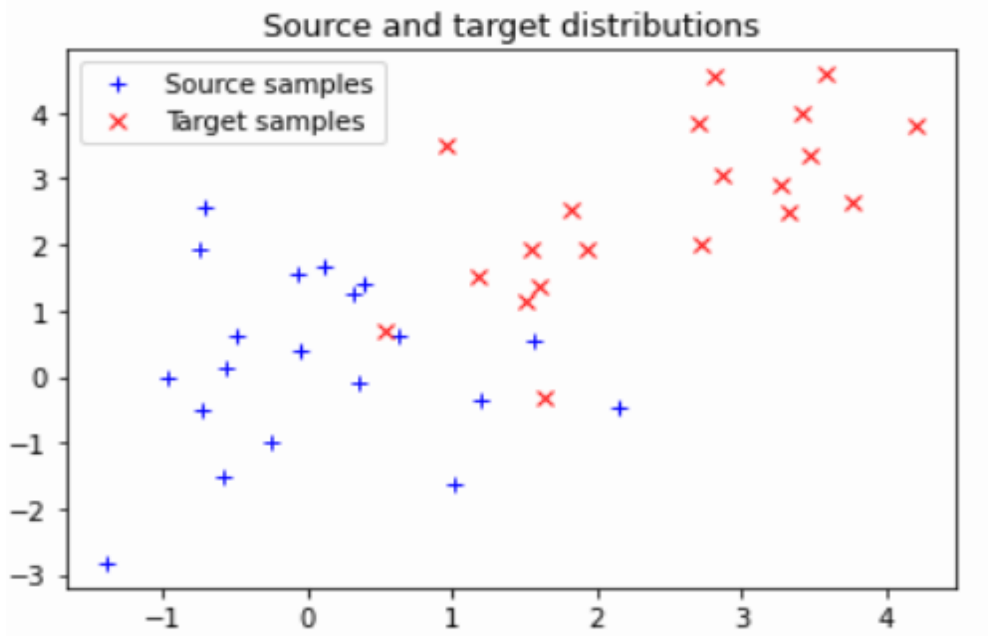
Introduction :

Le transport optimal est un problème mathématique qui consiste à trouver le coût minimal pour transporter une mesure de probabilité source vers une autre mesure de probabilité cible. Cette notion fut introduite pour la première fois par le mathématicien français Gaspard Monge en 1781 dans son ouvrage "Mémoire sur la théorie des déblais et des remblais". Celui-ci cherchait à déplacer de manière optimale des tas de sable vers des fossés pour des fins militaires.

La modélisation du problème de transport optimal par Monge est déterministe et particulièrement contraignante, elle ne trouve pas toujours de réponses et l'unicité de celle-ci n'est pas garantie. L'approche du problème fût assouplie par les travaux de Leonid Kantorovich et Tjalling Koopmans en 1975. Ils ont tous deux reçu le prix Nobel d'économie pour leur travail sur cette question.

Le transport optimal est un concept mathématique source de recherche actuellement. Cette théorie est utilisée dans des domaines comme l'économie, l'informatique, la physique, la biologie et bien d'autres. Il peut servir à résoudre des problèmes de planification de production dans l'industrie. Par exemple dans la grande distribution, pour faciliter les transports des marchandises en connaissant les stocks dans les entrepôts et les demandes des clients, il est possible d'optimiser les envois de marchandises en utilisant le transport optimal.

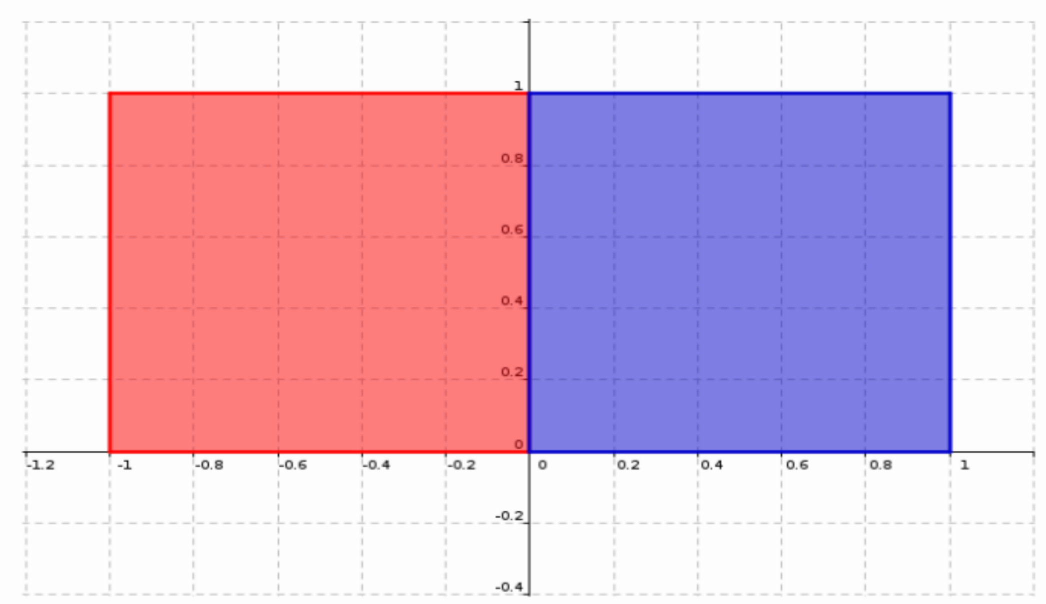
Modélisation du problème par Monge :



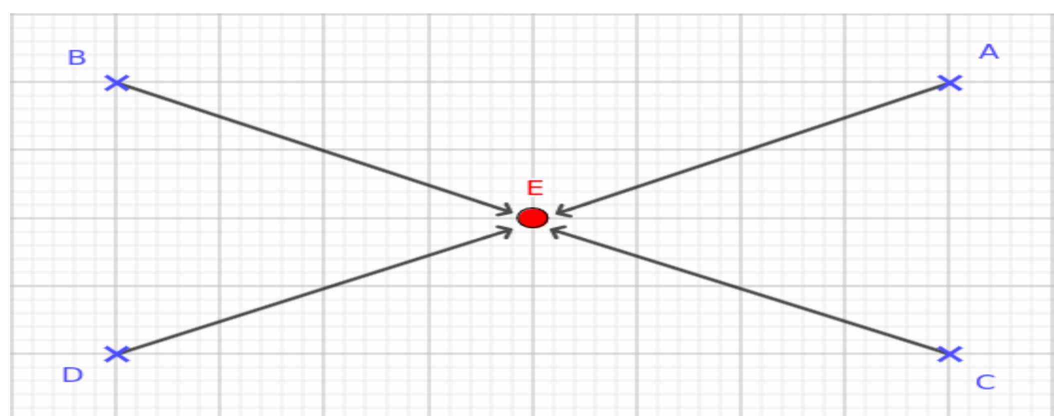
- Soit $\mu = \sum_{i=1}^{n} \alpha_i \delta_{\{x_i\}}$ une mesure de probabilité source.
- Soit $\nu = \sum_{i=1}^{m} \beta_i \delta_{\{y_i\}}$ une mesure de probabilité cible.

Peut-on construire une application $T : \mathbb{R}^2 \to \mathbb{R}^2$ transportant μ sur ν sous une certaine contrainte et avec un coût minimal ? Autrement dit : comment transporter les points bleus sur les points rouges avec un coût de transport minimal ?

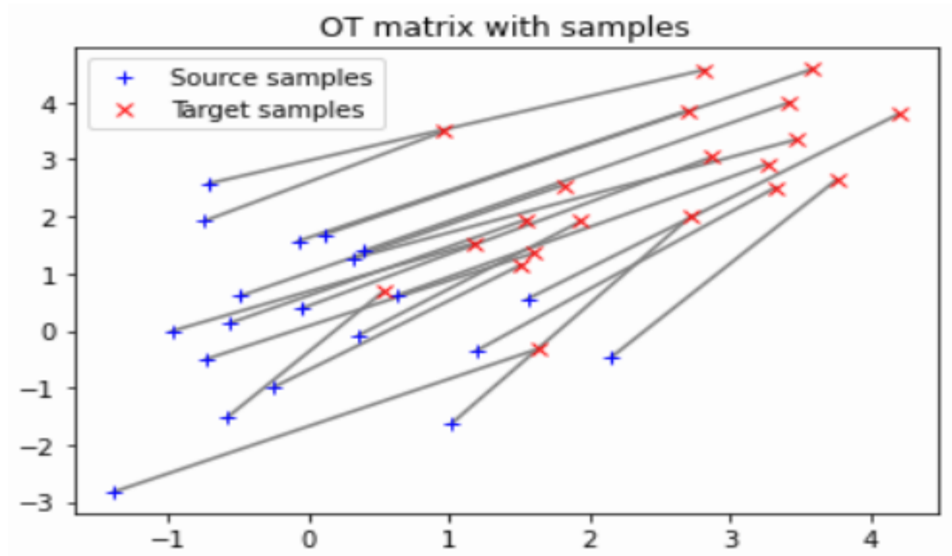
- La contrainte sur T :
  la *mesure image* de μ par T est égale à ν.
  $$T_{\#\mu}(\cdot) = \mu(T^{-1}(\cdot)) = \nu(\cdot)$$
- Le coût d'un transporteur :
- Une fonction de coût peut être la distance induite par la norme euclidienne $\|\cdot\|$ sur $\mathbb{R}^2$ par exemple.
- Le coût d'un transporteur du plan T_1 est donné par :
  $$C(T_1) = \int_{\mathbb{R}^2} \|x - T_1(x)\| \, \mu(dx).$$
  Dans le cas discret :
  $$C(T_1) = \sum_{i=1}^{n} \|x_i - T_1(x_i)\| \times \alpha_i.$$
- Transporteur optimal :
  T^* une application réalisant l'infimum :
  $$T^* = \text{arginf}\{C(T);\ T : \mathbb{R}^2 \to \mathbb{R}^2,\ T_{\#\mu} = \nu\}$$



Il y a plusieurs manières de transporter la "masse uniforme" bleue sur la "masse uniforme" rouge. Vous pouvez utiliser une translation, une symétrie ou encore une rotation. Les coûts associés à ces transporteurs pour la distance euclidienne varient. Comment choisir le meilleur transporteur ? Ici la translation et la symétrie sont les transporteurs optimaux, il n'y a pas toujours unicité de la solution pour le problème de transport optimal.



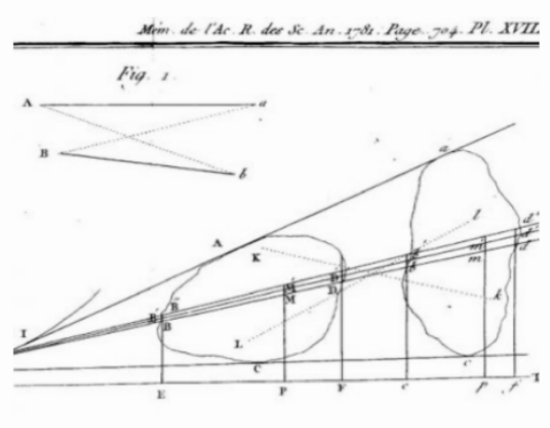
La modélisation du problème de transport optimal proposée par Gaspard Monge trouve ses limites dans la division de la masse. Une application transport est déterministe, tout antécédent à une unique image par un transporteur T. Nous pouvons transporter les points bleus sur l'unique point rouge cible. En revanche le problème inverse ne trouve pas de solution au sens de Monge. Il est impossible d'envoyer le point rouge à quatre endroits différents.



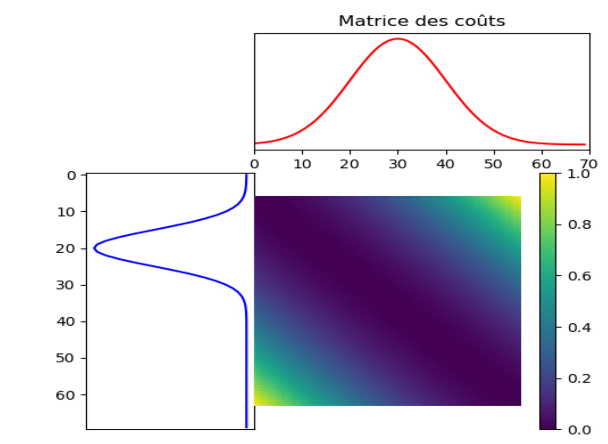
Voici le résultat obtenu pour le problème de transport optimal vu précédemment. Nous avons modélisé les points bleus et les points rouges à l'aide de deux familles de vecteurs Gaussiens avec des paramètres distincts. Le couplage entre les points s'effectue via un script Python qui utilise le module "POT: Python Optimal Transport".

Assouplissement des contraintes, problème de Monge-Kantorovich :

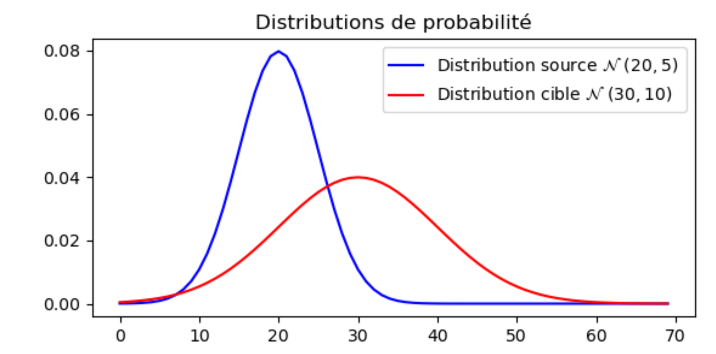

Gaspard Monge (1746-1818)

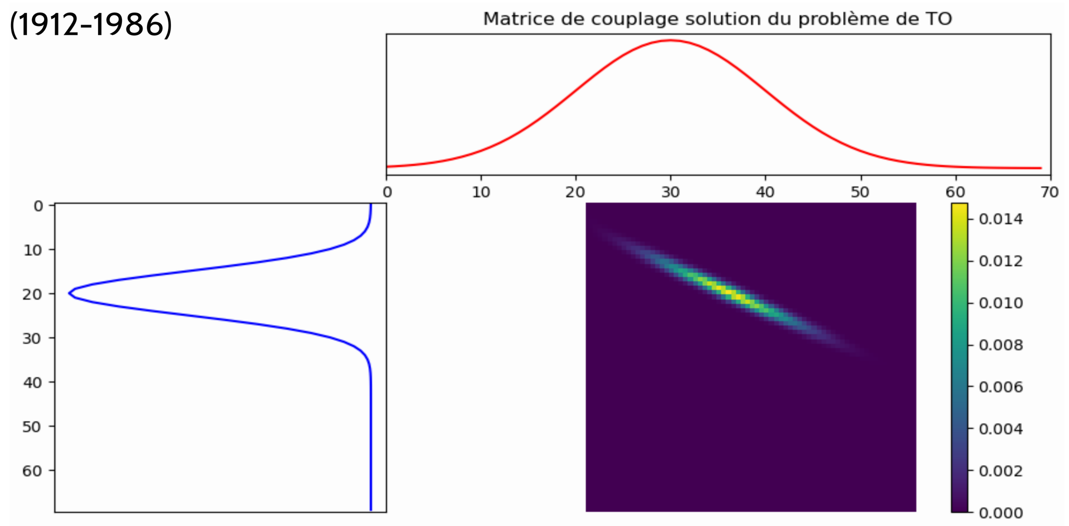

Extrait de la théorie des déblais et des remblais


Leonid Vitalievitch Kantorovitch (1912-1986)



- Une vision probabiliste et non déterministe du problème.
- Soit $\mu, \nu \in \mathcal{P}(\mathbb{R})$ et (X, Y) un couple de variables aléatoires de loi marginale respective μ, ν. On note $\pi \in \mathcal{P}(\mathbb{R} \times \mathbb{R})$ la loi du couple.
- $\Pi(\mu, \nu)$ l'ensemble des couplages possibles.

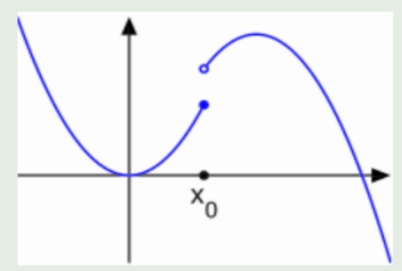




- Un transport est donné par un couplage π. Pour c une fonction de coût on obtient le coût du transport :
  $$I_c(\pi) = \iint_{\mathbb{R} \times \mathbb{R}} c(x, y)\pi(dxdy)$$
- Problème de Monge-Kantorovich $\Rightarrow$ optimisation sur l'espace des couplages possibles.
- Si le minimum est réalisé, $I_c(\pi^*) = T_c(\mu, \nu)$, on dit que π^* est un couplage optimal.
  $$\mathcal{T}_c(\mu, \nu) = \inf\{I_c(\pi)\ :\ \pi \in \Pi(\mu, \nu)\}$$
  $$\pi^* = \text{argmin}\{I_c(\pi)\ :\ \pi \in \Pi(\mu, \nu)\}$$

Définition

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

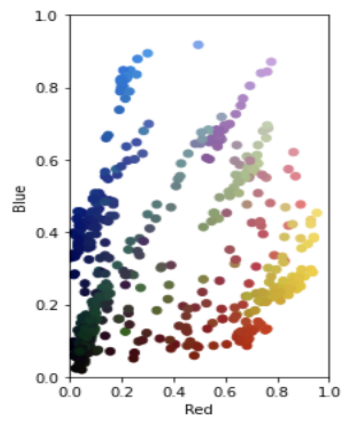


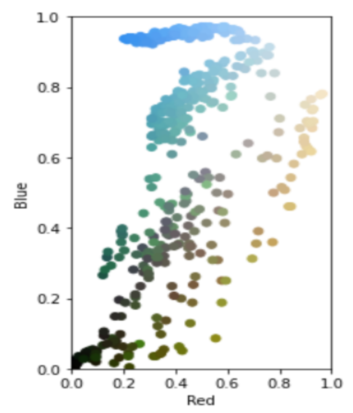
Théorème

Soit $c : E \times E \to [0, +\infty]$ une fonction de coût sci. Pour toutes mesures de probabilité $\mu, \nu \in \mathcal{P}(E)$, il existe $\pi^* \in \Pi(\mu, \nu)$ tel que
$$I_c(\pi^*) = \iint_{E\ E} c(x, y)\pi^*(dxdy) = T_c(\mu, \nu) = \inf\{I_c(\pi)\ :\ \pi \in \Pi(\mu, \nu)\}$$

Transfert de couleurs entre deux images via le Transport Optimal :

ÉCHANTILLONNAGE 200 PIXELS


Pixels Sources


Pixels Cibles

Couplage entre les pixels grâce à un algorithme de transport Optimal

Identification des couleurs et visualisation

Travail encadré par Fabien Panloup