Transport Optimal et Architecture

Matthieu Boyer

Cours TalENS 2 2025-2026

Table des matières

Ce polycopié ne doit pas être vu comme un remplacement au cours, mais simplement comme un résumé du contenu qui permet d'être sûr de ne rien manquer.

1 Problème de Monge

Gaspard Monge, ingénieur militaire de Napoléon, est chargé par ce dernier du problème suivant :

Étant donné des grognards portant des sacs de sable depuis n camps, quelle est la manière optimale de construire n murs à des endroits différents, sachant que les endroits sont à des distances plus ou moins grandes de chaque camp ?

Ce problème fait partie de la grande classe des problèmes d'optimisation, et est fondamental en théorie des probabilités, puisqu'il est à la base de la théorie du transport optimal qui donne des manières de trouver des distributions de probabilités "moyennes".

1.1 Bases de Probabilités

On ne rentrera pas ici dans les "vraies" bases de la théorie de la mesure, et notamment sur la notion de tribu.

Définition 1.1 Une mesure sur un espace $\mathcal{X}$ muni d'une tribu ensemble Σ de parties de X est une application $\mu : \Sigma \to \mathbb{R}$ telle que :

- $\mu(\varnothing) = 0$
- $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$
- $\mu\left(\bigcup_{i=0}^{\infty} A_i\right) = \sum_{i=0}^{\infty} \mu(A_i)$ si les A_i sont deux à deux disjoints

- $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$.

On dit que $\mu(\mathcal{X})$ est la masse totale de μ.

Intuitivement, une mesure est une application qui donne la quantité d'éléments dans une partie d'un espace.

Dans notre cas, on s'intéressera principalement aux cas $|\mathcal{X}| \hookrightarrow \mathbb{N}$ et $X = \mathbb{R}^d$.

Définition 1.2 Sur $\mathcal{X}$ au plus dénombrable, on a une mesure dite de comptage qui a chaque partie de X associe son nombre d'éléments.

Sur $\mathcal{X} = \mathbb{R}^d$ on a une mesure λ dite mesure de Lebesgue qui à chaque partie de X associe son volume. En particulier, $\lambda\left(\prod_i [a_i, b_i]\right) = \prod_i |b_i - a_i|$.

Définition 1.3 Une mesure de probabilité est une mesure positive de masse 1. Une mesure μ a densité par rapport à la mesure de Lebesgue s'il existe une fonction p telle que

$$\mu(A) = \int_A p(x)\, \mathrm{d}\lambda(x)$$

Le symbole intégrale $\int$ signifie ici que pour tout point x dans A, on construit un petit pavé autour de A, qu'on calcule son volume $\mathrm{d}\lambda(x)$ et qu'on le multiplie par la densité de μ en x $p(x)$.

Autrement dit, dans le volume $\mathrm{d}\lambda(x)$ autour de x, il y a $p(x)$ particules. On trouve donc le nombre total de particules dans A en sommant les nombres de particules autour de tout point x de A.

Définition 1.4 Pour $x \in \mathbb{R}^d$, on définit le dirac en x comme la mesure de probabilité δ_x qui vaut 1 sur $\{x\}$ et 0 ailleurs.

Pour $\mu \in \mathbb{R}^d, \Sigma \in \mathcal{M}_{d,d}(\mathbb{R})$, on définit la gaussienne centrée en μ de covariance Σ par sa densité $p(x) = \exp\left(-\frac{1}{2}\Sigma^{-1}(x-\mu)^2\right)$.

Définition 1.5 Une variable aléatoire X à valeurs dans E est une fonction dite mesurable de $\mathcal{X}$ dans E. La probabilité que X soit à valeurs dans $A \subseteq E$ est la mesure $\mu(X^{-1}(A))$.

Autrement dit, c'est la quantité (au sens de μ) d'antécédents des éléments de A dans par X.

1.2 Formulation de Monge

On considère dans la suite deux mesures de probabilité $\alpha \in \mathcal{P}(\mathcal{X}), \beta \in \mathcal{P}(\mathcal{Y})$.

Définition 1.6 Pour $T : \mathcal{X} \to \mathcal{Y}$, on définit la mesure $T_\sharp\alpha$ par $T_\sharp\alpha(B) = \alpha(T^{-1}(B))$.

Autrement dit, $T_\sharp\alpha(B)$ est la quantité (selon α) d'antécédents des éléments de B par T.

La formulation de Monge du problème de transport optimal de α à β est la suivante :

Définition 1.7 Soit c une fonction de coût de $\mathcal{X}$ à $\mathcal{Y}$, c'est-à-dire une fonction positive de $\mathcal{X}$ dans $\mathcal{Y}$. Le problème de Monge associé à α, β, c est de calculer :

$$M = \inf_{T:\mathcal{X}\to\mathcal{Y}} \left\{ \int c(x, T(x))\, \mathrm{d}\alpha(x) \,\middle|\, T_\sharp\alpha = \beta \right\} \tag{1}$$

M représente la manière optimale de déplacer tout le poids de α vers β, sachant le coût du déplacement d'un point de $\mathcal{X}$ vers un point de $\mathcal{Y}$.

Définition 1.8 Pour $c = \|\cdot\|$ une norme, on définit la p-distance de Wasserstein associée à c comme

$$W_p(\alpha, \beta) = \inf_{T_\sharp \alpha = \beta} \sqrt[p]{\int \|x - T(x)\|^p \, \mathrm{d}\alpha(x)}.$$

2 Calcul et Applications

2.1 Calcul Discret

Dans le cas, plus simple, où $\alpha = \sum_{i=1}^N a_i \delta_{x_i}$ et $\beta = \sum_{j=1}^N b_j \delta_{y_j}$, les applications T telles que $T_\sharp \alpha = \beta$ correspondent à des assignations des x_i aux y_j, c'est à dire une permutation σ de l'ensemble $[\![1, N]\!]$ telle que $T(x_i) = y_{\sigma(i)}$.

Le problème d'optimisation qui correspond au transport optimal dans ce cas est le problème de couplage sur un graphe biparti complet :

Définition 2.1 Un graphe biparti complet est constitué d'un ensemble $V = A \sqcup B$ de *sommets* composé de deux parties A et B de tailles égales n. Pour chaque sommet de A on ajoute une *arête* à tout sommet de B avec des poids associés au coût entre les deux sommets. Un couplage sur un graphe biparti est le choix de n arêtes qui touchent tout sommet de A et tout sommet de B.

Dans l'Algorithme 1 on présente un pseudo-code pour l'algorithme hongrois qui permet de résoudre ce problème. L'idée de base étant de se ramener à chaque itération à un problème plus simple : une fois qu'un sommet est assigné, on le "supprime" virtuellement de la matrice de coût puis on recommence jusqu'à avoir terminé. Cet algorithme est dit *glouton* parce qu'il effectue à tout instant l'assignation la plus efficace.

Cet algorithme est assez efficace, pusiqu'il effectue un nombre d'opérations dit $\mathcal{O}(n^3)$ (lire « grand O de n^3 »), qui est de l'ordre de n^3 et plus mathématiquement, il existe une constante positive C telle que le nombre d'opérations $f(n)$ en fonction de n vérifie $f(n) \leq n^3$.

2.2 Applications du Transport Optimal

Algorithme 1 Algorithme Hongrois pour le problème d'assignation bipartite

procedure MAXZ(C) ▷ Input : Cost matrix C of size $t \times t$
  $Z \leftarrow \{\}$
  while $0 \in C_{i,j}$, $\forall (i,j) \notin Z$ **do**
    $x \leftarrow$ row such that $C_{x,y}$ has the fewest marked 0 elements
    $y \leftarrow$ column such that $C_{x,y} = 0$ and C_y has the fewest marked 0 elements
    Add (x, y) to set Z
    Mark $C_{x,y}$ as an independent zero
return Set Z containing independent zeros
end procedure

procedure MINCOVER(C, Z) ▷ Input : Cost matrix C and set Z
  $coveredRows \leftarrow \{\}$ ▷ Rows covered with horizontal lines
  $coveredCols \leftarrow \{\}$ ▷ Columns covered with vertical lines
  while $C_{i,j} = 0$ and $i \notin coveredRows$ **do**
    if $0 \in C_i$ and $i \in Z$ **then**
      $coveredRows \leftarrow i$ ▷ Cover row i
    if $C_{i,j} = 0$ and $i \notin coveredRows$ **then**
      $coveredCols \leftarrow j$ ▷ Cover column j
return $coveredRows, coveredCols$
end procedure

procedure HUNGARIAN(C) ▷ Input : Cost matrix C of size $t \times t$
  for every row i and column j in C **do**
    $C_i \leftarrow C_i - \min(C_i)$ ▷ Subtract row minimum
    $C_j \leftarrow C_j - \min(C_j)$ ▷ Subtract column minimum
  $Z \leftarrow$ MAXZ(C)
  $cx, cy \leftarrow$ MINCOVER(C, Z)
  while $\text{len}(cx) + \text{len}(cy) < n$ **do**
    $minVal \leftarrow \min(C_{ij})$ such that $i \notin cx$ and $j \notin cy$
    for every row $i \notin cx$ and column $j \notin cy$ **do**
      $C_{ij} \leftarrow C_{ij} - minVal$
    $Z \leftarrow$ MAXZ(C)
    $cx, cy \leftarrow$ MINCOVER(C, Z)
return Set Z containing index (i, j) of assignments
end procedure