Proposition de Projet - **M1 Mathématiques et Applications**

Matching de nuages de points aléatoires: étude combinatoire, probabiliste et informationnelle

Matteo D'ACHILLE & Luca GANASSALI
matteo.dachille@universite-paris-saclay.fr luca.ganassali@universite-paris-saclay.fr

3 janvier 2025

Introduction. Comment apparier deux nuages indépendants de points aléatoires de façon optimale ? Dans un cadre géométrique, on peut formaliser cette question comme suit. Soit (Ω, D) un espace métrique polonais, et soient $\mathcal{B} = (B_i)_{i=1}^n$ et $\mathcal{R} = (R_i)_{i=1}^n$ deux processus ponctuels binomiaux indépendants sur Ω, désignés respectivement par "bleu" et "rouge", et ayant respectivement pour mesures d'intensité $\nu_\mathcal{B}$ et $\nu_\mathcal{R}$. Pour $p \in \mathbb{R}$, considérons la variable aléatoire suivante :

$$\mathcal{H}_{\text{opt}} = \min_{\pi \in S_n} \sum_{i=1}^{n} D^p(B_i, R_{\pi(i)}) \ , \tag{1}$$

où S_n est le groupe des permutations de n éléments. Le triplet de l'espace métrique (Ω, D), de l'aléa des points $\nu_\mathcal{B}, \nu_\mathcal{R}$ et de l'exposant distance-énergie $p \in \mathbb{R}$ définit un **Euclidean Random Assignment Problem** (ERAP en abrégé).

Un autre point de vue sur le problème est celui du modèle dit de **matching planté**. Nous considérons un modèle P_1 dans lequel les nuages de points $(\mathcal{B}, \mathcal{R})$ ne sont plus tirés selon deux processus de points indépendants comme précédemment, mais où le nuage $\mathcal{R}$ est obtenu à partir du nuage $\mathcal{B}$ selon le modèle suivant :

$$R_{\pi^\star(i)} = B_i + \sigma \xi_i \ , \tag{2}$$

où $\sigma > 0$ et ξ_i sont des bruits i.i.d. centrés, et où $\pi^\star$ est une permutation uniforme dans S_n tirée indépendamment de tout le reste. Dans ce nouveau problème, nous nous intéresserons à la question suivante : dans le modèle P_1, est-il possible, en observant $(\mathcal{B}, \mathcal{R})$, de retrouver le matching planté $\pi^\star$?

Objectifs du projet. Après s'être familiarisés avec les outils de base du projet, tels que le théorème de Birkhoff–von Neumann, l'algorithme hongrois, les processus ponctuels de Poisson et binomiaux, et le lien entre la v.a. dans l'eq. (1) et les distances de Wasserstein ou Zolotaraev en transport optimal, nous nous concentrerons d'abord sur le cas où (Ω, D) est l'intervalle $[0, 1] \subset \mathbb{R}$ équipé de la distance euclidienne usuelle, avec $p \in [0, 1]$. Dans ce cadre :

— nous étudierons quelques propriétés combinatoires d'une permutation $\pi \in S_n$ réalisant le minimum dans l'eq. (1) ;
— nous comprendrons comment une telle structure combinatoire encode un arbre, et comment ce codage permet d'estimer, en fonction de l'exposant p, l'espérance de la v.a. dans l'eq. (1).

Nous nous intéresserons ensuite au modèle de matching planté de l'eq. (2), et :
— sous P_1, nous établirons le lien entre la loi a posteriori de $\pi^\star$ et le problème d'optimisation de l'eq. (1) lorsque les bruits sont à queues lourdes ;
— nous chercherons ensuite à établir un résultat de type transition de phase pour le problème de reconstruction exacte de $\pi^\star$;
— nous pourrons aussi étudier les aspects numériques pour la résolution ce problème, notamment le sampling selon la loi a posteriori, et illustrer la transition de phase via des simulations.

Durant toute cette étude, toute initiative visant à modifier les questions de recherche, les modèles, les directions, pourra être discutée.

Références

[1] Matteo D'Achille, *Statistical Properties of the Euclidean Random Assignment Problem*, PhD thesis, Paris-Saclay University, 2020. https://theses.hal.science/tel-03098672

[2] Luca Ganassali, *The graph alignment problem : fundamental limits and efficient algorithms*, PhD thesis, PSL Research University, Ecole normale supérieure, 2022. https://theses.hal.science/tel-03921009

[3] Yihong Wu and Jiaming Xu, *Statistical Problems with Planted Structures : Information-Theoretical and Computational Limits*, Cambridge University Press, 2021. https://arxiv.org/abs/1806.00118

[4] Robert J. McCann, *Exact Solutions to the Transportation Problem on the Line*, Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci., vol. 455, no. 1984, pp. 1341–1380, Apr. 1999. https://www.jstor.org/stable/53375

Dernière mise à jour : 3 janvier 2025.