



Paris, le 30/11/21

Offre de stage

Sujet : Optimisation non-convexe de problèmes parcimonieux régularisés avec le transport optimal

Possibilité de poursuivre sur une thèse

La Chaire Data Science and Artificial Intelligence for Digitized Industry and Services

Portée par Florence d'Alché-Buc, enseignante-chercheur dans le département Image, Données, Signal de Télécom ParisTech, la chaire DSAI réunit cinq partenaires industriels : Airbus Defence & Space, Engie, Idemia, Safran et Valeo Finance. Son objectif général est de développer, en liaison étroite entre les Parties, une formation et une recherche de niveau international.

Ses quatre principaux axes de recherche sont :

1. Analyse et prévision de séries temporelles (Predictive Analytics on Time Series) ;
2. Exploitation de données hétérogènes, massives et partiellement étiquetées (Exploiting Large Scale and Heterogeneous, Partially Labelled Data) ;
3. Apprentissage pour une prise de décision robuste et fiable (Learning for Trusted and Robust Decision) ;
4. Apprentissage dans un environnement dynamique (Learning through Interactions with a Changing Environment).

Description du stage

Encadrement

Hicham Janati

Lieu et dates du stage

Telecom Paris, 19 place marguerite perey, 91120 Palaiseau

Date de début du stage : Avril 2022

Équipe(s) d'accueil de la thèse

département IDS, équipe Signal Statistique et Apprentissage

Mots clés

Machine learning, transport optimal, sparsité, optimisation

Sujet détaillé

La théorie du transport optimal introduite par Gaspard Monge (1781) [1] puis revisitée par Leonid Kantorovich (1942) [2] permet de modéliser et calculer le coût minimal du transport d'une distribution source (de masse par ex.) à une distribution cible en tenant compte de la métrique de l'espace ambiant. Les récentes avancées numériques et computationnelles du transport régularisé (Cuturi, 2013) [3] ont permis de repositionner le transport optimal au centre de plusieurs applications en traitement de signal (computer vision, traitement de langage, imagerie médicale). En imagerie cérébrale par exemple, une carte d'activation fonctionnelle est caractérisée par son amplitude ainsi que son support sur le graphe modélisant le cerveau. L'utilisation du transport optimal dans ce contexte [4, 5, 6] est naturellement intéressante pour tenir compte de la géométrie du cerveau et des différences spatiales entre des cartes différentes. Néanmoins, ces comparaisons nécessitent des extensions à des mesures signées plus ou moins artificielles.

Le but de ce stage est d'éviter le recours à ces extensions en étudiant des modèles non-convexes où l'optimisation est alternée entre le support des mesures et les poids (signés ou non). Ceci permet de découpler la notion de support de celles des amplitudes (et du signe) en particulier quand il s'agit de régulariser le modèle.

Ce type de modèle a été étudié dans le contexte des modèles parcimonieux multi-tâches (Multi-level Lasso [7]). Le but de ce stage est d'étendre cette décomposition "multiplicative" dans le cadre d'un modèle régularisé par le transport optimal. Une première étape de ce travail sera d'étudier le modèle

$$\alpha = \sum_{i=1}^N a_i \delta_{x_i}$$

en supposant la taille du support

une formulation simple de type:

$$\min_{a,x} F(a\delta_x) + G_1(\delta_x) + G_2(a)$$

connue. Ceci permet

Où F est un terme d'attache aux données et G_1 et G_2 sont des régularisations éventuellement définies par le transport optimal où des méthodes de dualité [8] et de descente par coordonnées peuvent être déployées. La deuxième partie de ce stage sera consacrée à la relaxation de l'hypothèse "N connu" où les méthodes d'optimisation sur l'espace des mesures [9] seront explorées.

Après la conception théorique de la stratégie d'optimisation, l'implémentation de ces modèles se fera sur Python pour une évaluation sur des données publiques de Magnétoencephalographie (MEG).

Profil du candidat

Etudiant titulaire d'un master 2 recherche

- Apprentissage statistique / reconnaissance des formes
- Bon niveau en programmation (Python)
- Bon niveau d'anglais
- Intéressé par des problèmes théoriques et leurs applications
- Idéalement, des connaissances en transport optimal

Candidatures

A envoyer à hicham.janati100@gmail.com:

- Curriculum Vitae
- Lettre de motivation personnalisée expliquant l'intérêt du candidat sur le sujet (directement dans le corps du mail)
- Relevés de notes des années précédentes
- Contact d'une personne de référence

Les candidatures incomplètes ne seront pas examinées.

Références

- [1] Monge G. (1781). "Mémoire sur la théorie des déblais et des remblais". In: Histoire de l'Académie Royale des Sciences, pp. 666–704.
- [2] Kantorovich L. (1942). "On the translocation of masses". In: C.R. Acad. Sci. URSS.
- [3] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In Advances in Neural Information Processing Systems 26, pages 2292–2300.
- [4] Gramfort A., Peyré G., and Cuturi M. (2015). "Fast Optimal Transport Averaging of Neuroimaging Data". In: Proceedings of the Information Processing in Medical Imaging conference.
- [5] Janati H., Cuturi M., and Gramfort A. (2019). "Wasserstein regularization for sparse multi-task regression". In: Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics. Vol. 89. Proceedings of Machine Learning Research. PMLR.
- [6] Janati H., Bazeille T., Thirion B., Cuturi M., and Gramfort A. (2020b). "Multi-subject MEG/EEG source imaging with sparse multi-task regression". In: NeuroImage, p. 116847.
- [7] Lozano A. and Swirszcz G. (2012). "Multi-level Lasso for Sparse Multi-task Regression". Proceedings of the 29th International Conference on Machine Learning 2012.
- [8] Cuturi M. and Peyré G. (2016). "A smoothed dual approach for variational Wasserstein problems". In: SIAM Journal on Imaging Sciences 9.1, pp. 320–343.
- [9] Chizat L (2020). "Sparse optimization on measures with over-parameterized gradient descent"

Mathematical Programming (2021).