

Recherche proposée - Reconstruction de surfaces en présence de bruit: garanties topologiques pour des approches bayésiennes nonparamétriques

Problématique: Les algorithmes de reconstruction de surfaces sont très largement utilisés en statistique appliquée pour l'élimination du bruit, la réduction de la dimensionnalité et l'estimation du support d'une variable aléatoire. Dans le domaine de l'intelligence artificielle, ils servent à identifier le contour d'un objet et à en modéliser la forme globale. Une application classique vient de l'imagerie médicale par résonance magnétique: à partir de numérisations du cerveau, on cherchera à identifier de potentielles tumeurs ainsi que leur nombre, taille et forme. Ces informations peuvent ensuite aider le personnel médical au cours d'une opération. Il est très important, dans ce cas, de pouvoir assurer la précision de l'identification des tumeurs. Les garanties recherchées à cette fin sont dites *métrique* et *topologique*. La première assure une bonne estimation des tailles et positions, tandis que la seconde permet l'identification du bon nombre et des bonnes formes. Bien que le domaine de la géométrie computationnelle ait contribué des algorithmes et résultats fonctionnant en l'absence de bruit, la considération additionnelle de possibles erreurs dans les données, comme il y en a toujours dans la réalité, échappe largement à l'analyse dans plusieurs contextes. La quantification de l'incertitude des algorithmes bayésiens, sur laquelle s'appuient des décisions souvent critiques, doit aussi être validée.

Objectifs visés: Développer un cadre théorique permettant de vérifier la précision métrique et topologique d'algorithmes bayésiens nonparamétriques de reconstruction de surfaces courbes lorsque les données observées sont perturbées par du bruit. Déterminer les propriétés fréquentistes de la quantification de l'incertitude et proposer des critères de validation.

Méthodologie et procédure proposées: La recherche proposée se décompose en trois volets. Le premier relève des propriétés déterministes de l'approximation de surfaces et vise à mesurer la complexité intrinsèque du problème. Une littérature assez vaste donne déjà des indications à cet égard, mais nous aurons à raffiner certains résultats. Les contraintes ainsi soulevées permettront, en un deuxième temps, d'obtenir des résultats généraux de convergence statistique par l'application de résultats récents sur l'asymptotique des procédures bayésiennes. La principale difficulté réside ici dans le type de convergence que nous devons obtenir: nos résultats préliminaires indiquent la nécessité d'une convergence dans un espace de Sobolev, où la théorie actuelle est peu développée. Finalement, nous considérerons l'usage pratique des algorithmes vérifiant nos conditions et développerons pour ceux-ci la quantification de l'incertitude et les critères de validation.

Connaissances actuelles: Ce projet s'appuie sur des résultats préliminaires [1], ayant fait l'objet d'une affiche et d'un article dans le *Notes from the Margin*, indiquant le potentiel de recherche et l'accessibilité de certains théorèmes. Bien que la considération du bruit pose toujours problème dans plusieurs pans de la littérature sur le sujet de la reconstruction de surfaces [2], nous estimons que l'application d'idées statistiques de la théorie bayésienne puisse éclairer la situation. Je serai en visite en novembre à l'université Texas A&M, sur l'invitation de Prof. Debdeep Pati, pour discuter, entre autres, d'une collaboration liée à ce projet de recherche. J'ai aussi été invité à Duke University cet automne par Prof. Rebecca Steorts et Prof. David Dunson pour présenter au *MLBytes Speaker Series* et discuter de la possibilité d'un doctorat sur ce projet de recherche.

[1] **Binette, O.** (2018) Topologie et apprentissage machine. *Notes from the Margin*. XIII: 5-6.

[2] Pati, D. & Dunson, D. (2014) *Bayesian Closed Surface Fitting through Tensor Products*. Accepted pending minor revisions at the *Journal of Machine Learning Research*.