Olivier Binette

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

a. 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.

b. 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.

c. 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. En particulier, il nous faut des critères exacts permettant d'assurer qu'une surface approximante est bel et bien isotopique à la surface cible. Des approches actuelles sur le sujet s'appuient sur la notion de *compatibilité normale* [5], mais nous pouvons renforcer certains résultats à l'aide de techniques de déformation plus flexibles. Cela permettra une meilleure pondération de la complexité des modèles par rapport à leurs propriétés d'approximation.

Les contraintes soulevées précédemment 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 devrons obtenir: nos résultats préliminaires indiquent la nécessité d'une convergence dans un espace de Sobolev, où la théorie actuelle est peu développée. Nous devrons donc investiguer l'efficacité des procédures bayésiennes nonparamétriques pour l'estimation sous les fonctions de pertes correspondantes.

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. Ces critères seront basés sur la stabilité topologique locale des surfaces estimées: les zones où des variations probables peuvent causer des bris dans la structure topologique (comme l'introduction d'auto-intersections, de trous et des déchirures) seront identifiées comme problématiques et incertaines. L'implémentation de cette idée pourra s'appuyer sur des outils de persistance topologique, bien que l'approche soit ici entièrement probabiliste.

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.

d. Contribution à l'avancement des connaissances

Notre projet comprend l'approfondissement de plusieurs champs de connaissances. D'une part, la théorie s'appuie sur le développement de méthodes d'approximation constructives de surfaces et le développement de la théorie asymptotique bayésienne pour les convergences dans les espaces de Sobolev. De plus, nos considérations pratiques requièrent des outils computationnels et graphiques pour la quantification de l'incertitude sur les surfaces, ainsi que de nouvelles techniques de validation basées sur des aspects probabilistes de la stabilité topologique. Nous croyons que ce projet mènera à des contributions significatives dans plus d'un domaine et qu'il aidera à populariser l'usage des algorithmes de reconstruction de surface dans différents domaines des sciences et génies par le biais de la proposition d'outils plus complets.

Bibliographie

[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*. Journal of Machine Learning Research; accepted pending minor revisions.

[3] Bobrowski, O., Mukherjee, S., & Taylor, J. E. (2017). Topological consistency via kernel estimation. *Bernoulli*, *23*(1), 288-328.

[4] Yang, Y., & Dunson, D. B. (2016). Bayesian manifold regression. *The Annals of Statistics*, *44*(2), 876-905.

[5] Chazal, F., Lieutier, A., & Rossignac, J. (2007). Normal-map between normal-compatible manifolds. *International Journal of Computational Geometry & Applications*, *17*(05), 403-421.