Inférence d'objets différentiels par interpolation polynômiale : vitesses minimax pour l'estimation de support, plans tangents et courbure.

Clément LEVRARD, LPMA - Université Paris Diderot

Eddie AAMARI, INRIA & LMO, Paris Saclay

Mots-clés : Inférence géométrique, plans tangents, polynômes locaux, vitesse de convergence, minimax

Certains types de données, comme la répartition des galaxies dans l'univers, des points sur une surface ou encore des paramètres physiques soumis à des contraintes, peuvent être modélisés comme s'organisant autour d'une structure de dimension réduite, une sous-variété M de dimension d de l'espace ambiant $\mathbb{R}^D$. Ces structures, pour peu qu'elles soient suffisamment régulières, peuvent être approchées localement par des sous-espaces vectoriels de dimension réduite. Sur cette hypothèse d'approximation linéaire locale sont basés plusieurs outils d'inférence géométrique des caractéristiques de cette variété support M, que l'objet d'intérêt soit la variété elle-même, via les algorithmes de reconstruction tels le Tangential Delaunay Complex [1], ou d'autres quantités structurelles telles le reach.
Pour beaucoup de ces algorithmes, la qualité d'estimation des espaces tangents joue un rôle déterminant. Dès lors, plusieurs études ont cherché à déterminer des estimateurs $\hat{T}_x M$ du plan tangent $T_x M$, avec des garanties théoriques sur la déviation en angle

$$\angle(\hat{T}_x M, T_x M) := \|\pi_{\hat{T}_x M} - \pi_{T_x M}\|_{op} = \max_{j=1,\ldots,d} |\sin(\theta_j)|,$$

où π_T est la projection orthogonale de $\mathbb{R}^D$ sur T, et les θ_j sont les angles principaux.
L'estimateur le plus fréquemment proposé dans un cadre probabiliste pour $T_x M$ ([3] par exemple) est basé sur une ACP locale du nuage de points au voisinage du point x. De ce point de vue les directions tangentes sont assimilées aux directions capturant au mieux l'étalement du nuage de point constitué par les voisins de x. Les différents résultats obtenus pour cet estimateur $\hat{T}_{PCA,x}$ mettent en relief l'influence de la régularité de la variété M. Informellement, pour une variété de classe $\mathcal{C}^2$, [3] montre que $\angle(\hat{T}_{PCA,x} M, T_x M) \leq n^{-1/d}$, tandis que pour une variété de classe $\mathcal{C}^3$, [4] obtient $\angle(\hat{T}_{PCA,x} M, T_x M) \leq n^{-3/(d+2)}$.
Cependant, des estimateurs prenant en compte, pour des régularités plus élevées, le caractére localement polynômial des paramétrisations de ces variétés semblent mener vers des vitesses de convergence plus élevées. Par exemple, [2] prouve qu' un estimateur $\hat{T}_x M$ basé sur l'interpolation locale des points y autour de x par une formule de type $y - x = \pi(y-x) + (y-x)^t T_2 (y-x) + O(\|y-x\|^3)$ mène à une vitesse de convergence de type $\angle(\hat{T}_x M, T_x M) \leq (1/n)^{-2/d}$, où l'estimateur $\hat{T}$ est le sous-espace vectoriel associé au projecteur interpolant π dans la formule ci-dessus. En nous basant sur cette idée, nous nous sommes interrogés sur l'optimalité de ce type de procédure lorsqu'un ordre de régularité k est fixé. Par ailleurs, nous avons cherché à déterminer le rôle précis de la dimension ambiante dans les vitesses d'estimation.

Références

[1] Boissonnat, Jean-Daniel and Ghosh, Arijit, *Manifold reconstruction using tangential Delaunay complexes*, Discrete & Computational Geometry, 2014.

[2] Siu-Wing Cheng and Man-Kwun Chiu, *Tangent Estimation from Point Samples*, Discrete & Computational Geometry, 2016.

[3] Eddie Aamari and Clément Levrard, *Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction*, ArXiv e-prints, 2015.

[4] Singer, A. and Wu, H.-T., *Vector diffusion maps and the connection Laplacian*, Communications on Pure and Applied Mathematics, 2012.

[5] Frédéric Cazals and Marc Pouget, *Estimating differential quantities using polynomial fitting of osculating jets*, Computer Aided Geometric Design, 2005.

Clément LEVRARD, LPMA, Bâtiment Sophie Germain, Université Paris Diderot, 75013 Paris
levrard@math.univ-paris-diderot.fr