

Stage de Master 2 : Statistique nonparamétrique pour des données directionnelles

Date et lieu

Printemps 2023
Champs-sur-Marne (77) et/ou Villetaneuse (93)

Encadrement

Claire Lacour¹ claire.lacour@univ-eiffel.fr,
et Thanh Mai Pham Ngoc² phamngoc@math.univ-paris13.fr

1. Laboratoire d'Analyse et de Mathématiques Appliquées
Université Gustave Eiffel
Cité Descartes, Bâtiment Copernic
5 boulevard Descartes
77454 Marne-la-Vallée cedex 2

2. LAGA CNRS UMR 7539
Institut Galilée
Université Sorbonne Paris Nord
99 avenue Jean Baptiste Clément
93430 Villetaneuse

Sujet

Ce projet propose d'étudier des données directionnelles qui représentent des directions de \mathbb{R}^d ou plus généralement sont des observations sur des variétés riemanniennes compactes. Typiquement, sur le cercle unité \mathbb{S}^1 on parlera de données circulaires et sur l'hypersphère \mathbb{S}^{d-1} de données sphériques. Les données circulaires sont fréquemment utilisées en sciences de la terre (direction du vent par exemple) mais aussi en écologie (direction prise par des animaux) ou encore en médecine (modélisation d'un moment de la journée), quant aux données sphériques on les rencontre par exemple en astrophysique, en paléomagnétisme ou en géologie. Les données directionnelles nécessitent un traitement spécifique de par la topologie sous-jacente ; les méthodes statistiques classiques utilisées dans un cadre euclidien pour des variables dans \mathbb{R} ou \mathbb{R}^d ne peuvent plus être mises en oeuvre. De nombreux surveys pour l'étude des données circulaires existent : [Mardia and Jupp, 2000], [Jammalamadaka and SenGupta, 2001], [Pewsey, 2004] ou plus récemment [Ley and Verdebout, 2017].

Dans ce stage, à l'aide des ouvrages précédemment cités, on s'intéressera progressivement aux notions suivantes : représentation des données, fonction caractéristique, lois classiques et leurs récents développements, fonctions de profondeur, estimation de paramètres, test d'hypothèses.

Des simulations numériques en R ou Python pourront être réalisées. Des estimations ou tests sur données réelles pourront compléter la partie numérique.

Enfin, selon l'avancement du stage, on s'intéressera à des statistiques non-paramétriques, à travers l'étude de l'article de recherche récent [Lacour and Pham Ngoc, 2022] qui étudie le modèle de mélange semiparamétrique sur le cercle suivant : pour α et β deux angles inconnus, la densité des observations X_i est supposée être de la forme $g(x) = pf(x - \alpha) + (1 - p)f(x - \beta)$ où la proportion p du mélange ainsi que la densité f sont supposées inconnues et doivent être estimées.

Ce stage pourra déboucher sur une thèse, dirigée par les mêmes encadrantes, portant sur des modèles de censure et des modèles de mélanges en statistiques directionnelles.

Références

- [Jammalamadaka and SenGupta, 2001] Jammalamadaka, S. R. and SenGupta, A. (2001). *Topics in circular statistics*, volume 5 of *Series on Multivariate Analysis*. World Scientific Publishing Co., Inc., River Edge, NJ.
- [Lacour and Pham Ngoc, 2022] Lacour, C. and Pham Ngoc, T. M. (2022). Semiparametric inference for mixtures of circular data. *Electronic Journal of Statistics*, 16(1) :3482–3522.
- [Ley and Verdebout, 2017] Ley, C. and Verdebout, T. (2017). *Modern directional statistics*. Chapman & Hall/CRC Interdisciplinary Statistics Series. CRC Press, Boca Raton, FL.
- [Mardia and Jupp, 2000] Mardia, K. V. and Jupp, P. E. (2000). *Directional statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester.
- [Pewsey, 2004] Pewsey, A. (2004). Testing for circular reflective symmetry about a known median axis. *Journal of Applied Statistics*, 31 :575–585.