

Robust shape inference from a sparse approximation of the Gaussian trimmed loglikelihood.

Claire Brcheteau, Universit de Nantes

Clment Levrard, Universit Paris-Diderot

Bertrand Michel, Universit de Nantes

Given a noisy sample of points lying around some shape \mathcal{M} , with possibly outliers or clutter noise, we focus on the question of recovering \mathcal{M} , or at least geometric and topological information about \mathcal{M} . Often, such inference is based on the sublevel sets of distance-like functions such as the function distance to \mathcal{M} , the distance-to-measure (DTM) [2] or the k -witnessed distance [4]. In this talk, we firstly widespread the concept of trimmed log-likelihood to probability distributions. This trimmed log-likelihood can be considered as a generalisation of the DTM.

A sparse approximation of the DTM, the m -power distance-to-measure (m -PDTM), has been introduced and studied by Brcheteau and Levrard in 2017 [1]. Its sublevel sets are unions of m balls, with m possibly much smaller than the sample size. By miming the construction of the m -PDTM from the DTM, we propose an approximation of the trimmed log-likelihood associated to the family of Gaussian distributions on \mathbb{R}^d . This approximation is sparse in the sense that its sublevel sets are unions of m ellipsoids.

We provide a Lloyd-type algorithm to compute the centers and covariance matrices associated to the ellipsoids. We improve our algorithm by allowing an additional noise parameter to wipe out some points, just as the trimmed m -means algorithm of Cuesta-Albertos et al. [3]. Our algorithm comes together with a heuristic to select this parameter. Some illustrations on different examples enhance that our algorithm is efficient in wiping out clutter noise, recovering the shape and recovering the homology of \mathcal{M} ; this requiring a storage of only m points and covariance matrices.

Références

- [1] BRCHETEAU, CLAIRE AND LEVRARD, CLMENT, *The k -PDTM : a coresets for robust geometric inference*, preprint, 2017.
- [2] FRDRIK CHAZAL, DAVID COHEN-STEINER AND QUENTIN MRIGOT, *Geometric Inference for Probability Measures*, Foundations of Computational Mathematics, 2011.
- [3] CUESTA-ALBERTOS, JUAN. A. AND GORDALIZA, ALFONSO AND MATRÁN, CARLOS, *Trimmed k -means: an attempt to robustify quantizers*, The Annals of Statistics, 1997.
- [4] GUIBAS, LEONIDAS J. AND MÉRIGOT, QUENTIN AND MOROZOV, DMITRIY, *Witnessed K -distance*, SoCG '11, 2011.