MASCOT-NUM 2022
Clermont Ferrand, France
7-9 June 2022

Gaussian processes indexed by clouds of points: a study

Babacar Sow
Ecole Nationale Supérieure des Mines de Saint-Etienne

Supervisor(s): R. Le Riche (CNRS LIMOS), J. Pelamatti (EDF R&D), S. Zannane (EDF R&D) and M. Keller (EDF R&D)

PhD expected duration: Nov. 2021 - Oct. 2024

Address: Ecole Nationale Supérieure des Mines de Saint-Etienne, LIMOS, France

E-mail: babacar.sow@emse.fr

Abstract:
Optimizing the electrical production of a wind farm starts with finding the number and the optimal positions of the turbines that minimizes the losses caused by wake interactions and the installation cost while satisfying exclusion constraints. The optimization variables are the number and the set of positions of the turbines, and take the form of a cloud of points. These clouds are sets of coordinates with varying finite cardinalities. This is a natural encoding for variables that are invariant under permutation. Simulating in a realistic fashion a wind farm's yield is computationally expensive and all analyses requiring a large number of evaluations are inefficient. In a classic way, this work deals with replacing the time-consuming simulations by a surrogate model, such as a Gaussian process, for a later use in a Bayesian optimization framework.

This study addresses the problem of defining valid and efficient covariance kernels over variable-size clouds of points. This allows capturing the trend and important properties of the functions of interest.

There are several ways to define such kernels. A common way is to compose conditionally negative definite metrics with specific functions such as Radial Basis Functions (RBF). This ensures the kernel validity, in terms of semi-positive definiteness. In [1], for instance, sliced-Wasserstein is shown to comply with these conditions.
Another way consists in associating measures to the clouds of points through a deterministic mapping [2]. Yet, this mapping is not unique: it typically consists in selecting a specific probability density (uniform, Gaussian, etc): for a given cloud, one can consider several empirical hidden generative distributions. Once these empirical probabilities are chosen, one can directly exploit them as in [2] or embed them into a characteristic Reproducing Kernel Hilbert Space (RKHS) [3]. In the RKHS, the inner product of the images of two clouds is a valid covariance and defines a distance, which can be plugged into a kernel.

We test different kernels built on the above principles. Those that are based on a metric plugged into an RBF comprise the sliced-Wasserstein distance, the Wasserstein distance between empirical Gaussians which can be expressed in a closed form [4] and the Mean Maximum Discrepancy (MMD).
Kernels based on embedding distributions include (the estimations of) the uniform and the normal laws. One should keep in mind that these different methods are not mutually exclusive for characterizing a given kernel.

We also propose an alternative kernel constructed by considering the vector of features relevant to our clouds such as the mean and eigen components of the covariance matrix.

The above mathematically valid kernels are compared using two series of tests.
First, for each kernel, we assess how the covariance between two clouds of points changes with simple geometrical transformations of the clouds (see 1).
Second, the prediction performances of the different kernels are compared. For this, we create a benchmark of relevant test functions over clouds of points, including one that mimics wind farms power production, and a design of experiment problem.
Key words: Point clouds, Kernel, Hilbertian, Maximum Mean Discrepancy, Characteristic RKHS, Wasserstein distance, Gaussian Processes.

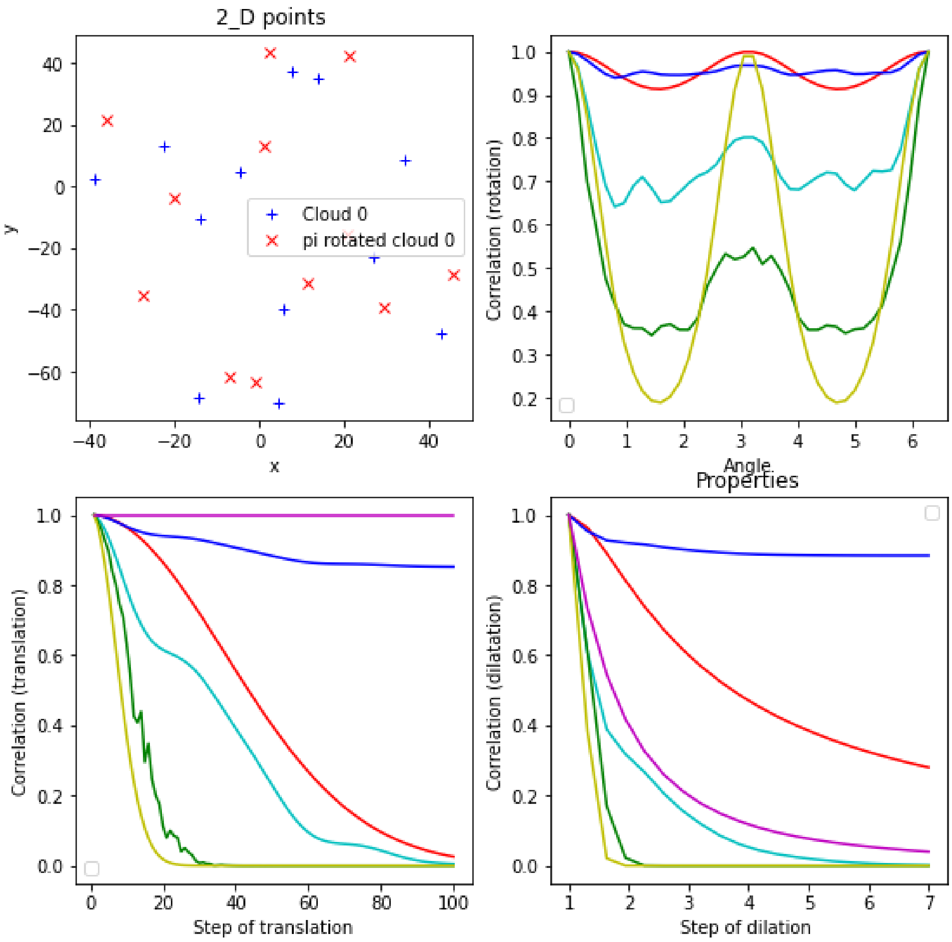

Figure 1: Correlations between clouds as measured by different kernels: Battachayra kernel (red), sliced Wassertein (green), Wassertein between Gaussians (yellow), MMD substitution(blue), uniform embeddings (dark turquoise), and Gaussian embeddings (violet). The clouds are rotated (top right), translated (bottom left) and dilated (bottom right).

References

[1] Mathieu Carriere, Marco Cuturi, and Steve Oudot. Sliced wasserstein kernel for persistence diagrams. In International conference on machine learning, pages 664–673. PMLR, 2017.

[2] Tony Jebara and Risi Kondor. Bhattacharyya and expected likelihood kernels. In Learning theory and kernel machines, pages 57–71. Springer, 2003.

[3] Krikamol Muandet, Kenji Fukumizu, Bharath Sriperumbudur, Bernhard Schölkopf, et al. Kernel mean embedding of distributions: A review and beyond. Foundations and Trends® in Machine Learning, 10(1-2):1–141, 2017.

[4] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport. Center for Research in Economics and Statistics Working Papers, (2017-86), 2017.

Short biography – Babacar Sow has received his Masters degree from Telecom Paris and Institut Polytechnique de Paris in Data Science And Applied Statistics. He currently pursues a PhD with Mines Saint-Etienne in collaboration with EDF. This work is funded by the ANR SAMOURAI project.