
Tensor Product Kernels: Characteristic Property, Universality*

Zoltán Szabó[†] (CMAP, École Polytechnique)

Abstract

Maximum mean discrepancy (MMD) and Hilbert-Schmidt independence criterion (HSIC) are among the most popular and successful approaches in data science and statistics (here MMD is sometimes referred to as energy distance or N-distance, and HSIC is called distance covariance) to quantify the difference and the independence of random variables, respectively. Thanks to their kernel-based foundations, MMD and HSIC are applicable on a wide variety of domains including documents, images, trees, graphs, time series, dynamical systems, sets, distributions, permutations. Despite their efficiency in numerous areas, quite little is known about when HSIC characterizes independence and MMD with tensor product kernel can discriminate probability distributions, in terms of the underlying kernel components. I am going to present a simple and complete answer to these questions.

- Preprint: <https://arxiv.org/abs/1708.08157>
- Code (Information Theoretical Estimators): <https://bitbucket.org/szzoli/ite-in-python/>

*Hangzhou International Conference on Frontiers of Data Science, Hangzhou, China, May 17-20, 2018; abstract.

[†]Joint work with Bharath K. Sriperumbudur (Department of Statistics, PSU).