Consistency of Orlicz Random Fourier Features*

Zoltán Szabó†

Abstract

Kernel techniques provide highly flexible tools with successful applications at virtually all sub-fields of machine learning and statistics. The random Fourier feature approach (RFF) is probably the most widely-applied and popular idea to combine this representational power of kernels with computational efficiency; it won the 10-year test-of-time award at NIPS-2017. While the RFF technique is typically used in case of tasks expressed via function values (such as kernel ridge regression), in numerous applications taking into account high-order derivatives turns out to be beneficial; examples include nonlinear feature selection or fitting infinite-dimensional exponential family distributions. Despite its practical success, the theoretical understanding of RFFs in case of derivatives is rather limited. In this talk, I will show how a finite α-exponential Orlicz norm assumption allows one to get consistent RFF approximations in case of high-order derivatives, covering for example the popular inverse multiquadric ($\alpha = 1$) or the Gaussian kernel ($\alpha = 2$).

*EPFL, Lausanne, Switzerland, September 23, 2019; abstract.

†Joint work with Linda Chamakh and Emmanuel Gobet.