
Towards Large-Scale Approximation of Tasks with Derivatives – A Kernel Perspective*

Zoltán Szabó[†]

Abstract

Kernel methods form the basis of numerous successful applications in data science thanks to their power in capturing complex relations. Various techniques have been developed to combine this flexibility of kernels with scalability, out of which probably the random Fourier feature (RFF) approach is the most popular and influential. In several applications, taking into account derivatives in addition to function values is highly beneficial; examples include Hermite learning, non-linear variable selection, or density estimation with infinite-dimensional exponential families. We show how to extend the consistency of RFFs to derivatives by an α -exponential Orlicz assumption on the spectral measure associated to the kernel. The Orlicz assumption covers widely-applied kernels such as the inverse multi-quadric ($\alpha = 1$) or the Gaussian kernel ($\alpha = 2$).

*International Conference on Modern Mathematical Methods and High Performance Computing in Science & Technology (M3HPCST), Ghaziabad, India, January 9-11, 2020; abstract.

[†]Joint work with Linda Chamakh and Emmanuel Gobet.