Scalable Kernel Embedding of Latent Variable Models*

Le Song

Abstract

Kernel embedding of distributions maps distributions to the reproducing kernel Hilbert space (RKHS) of a kernel function, such that subsequent manipulations of distributions can be achieved via RKHS distances, linear and multilinear transformations, and spectral analysis. This framework has led to simple and effective nonparametric algorithms in various machine learning problems, such as feature selection, two-sample test, time-series modeling and belief propagation. In this talk, I will focus on kernel embedding of latent variable models where the components in the models can have nonparametric form. The presence of latent variables in a model induces a sophisticated low rank structure in its kernel embedding, and is exploited for designing kernel algorithms for learning the latent parameters. While the method can adapt to the increasing complexity of the data as their volume grow, it is not scalable to large datasets. I will also introduce an approach called doubly stochastic functional gradients to scale up the methods and present some empirical results.

*Machine Learning External Seminar, Gatsby Unit, Oct. 21, 2014.