
Learning with Cross-Kernels and Ideal PCA*

Franz Király

Abstract

We describe how cross-kernel matrices, that is, kernel matrices between the data and a custom chosen set of ‘feature spanning points’ can be used for learning. The main potential of cross-kernels lies in the fact that (a) only one side of the matrix scales with the number of data points, and (b) cross-kernels, as opposed to the usual kernel matrices, can be used to certify for the data manifold. Our theoretical framework, which is based on a duality involving the feature space and vanishing ideals, indicates that cross-kernels have the potential to be used for any kind of kernel learning. We present a novel algorithm, Ideal PCA (IPCA), which cross-kernelizes PCA. We demonstrate on real and synthetic data that IPCA allows to (a) obtain PCA-like features faster and (b) to extract novel and empirically validated features certifying for the data manifold.

*Machine Learning External Seminar, Gatsby Unit, July 30, 2014.