Unsupervised Learning with Optimal Kernels

Sepp Hochreiter and Klaus Obermayer
Technische Universität Berlin

Abstract

We investigate the optimal kernel for sample-based model selection in unsupervised learning if maximum likelihood approaches are intractable. Given a set of training data and a set of data generated by the model, two kernel density estimators are constructed. A model is selected through gradient descent w.r.t. the model parameters on the integrated squared difference between the density estimators. Firstly we prove that convergence is optimal, i.e. that the cost function has only one global minimum w.r.t. the locations of the model samples, if and only if the kernel in the reparametrized cost function is a Coulomb kernel. As a consequence, Gaussian kernels commonly used for density estimators are suboptimal. Secondly we show that the absolute value of the difference between model and reference density convergences at least with $1/t$. Finally, we apply the new methods to distribution free ICA and to nonlinear ICA.