High-Dimensional Non-Linear Variable Selection

Francis Bach INRIA - Ecole Normale Sup'erieure 23, avenue d'Italie 75214 Paris, France francis.bach@ens.fr

Abstract

We consider the problem of high-dimensional non-linear variable selection for supervised learning. Our approach is based on performing linear selection among exponentially many appropriately defined positive definite kernels that characterize non-linear interactions between the original variables. To select efficiently from these many kernels, we use the natural hierarchical structure of the problem to extend the multiple kernel learning framework to kernels that can be embedded in a directed acyclic graph; we show that it is then possible to perform kernel selection through a graph-adapted sparsity-inducing norm, in polynomial time in the number of selected kernels. Moreover, we study the consistency of variable selection in highdimensional settings, showing that under certain assumptions, our regularization framework allows a number of irrelevant variables which is exponential in the number of observations.