

k-vectors: An Alternating Minimization Algorithm for Learning Regression Functions

Nir Weinberger
The Viterbi Faculty of Electrical and Computer Engineering
Technion - Israel Institute of Technology
nirwein@technion.ac.il

Meir Feder
School of Electrical Engineering
Tel Aviv University
meir@eng.tau.ac.il

Abstract--- The k-vectors algorithm for learning regression functions proposed here is akin to the well-known k-means algorithm. Both algorithms partition the feature space, but unlike the k-means algorithm, the k-vectors algorithm aims to reconstruct the response rather than the feature. The partitioning rule of the algorithm is based on maximizing the correlation (inner product) of the feature vector with a set of k vectors, and generates polyhedral cells, similar to the ones generated by the nearest-neighbor rule of the k-means algorithm. Similarly to k-means, the learning algorithm alternates between two types of steps. In the first type of steps, k labels are determined via a centroid-type rule (in the response space), which uses a surrogate hinge-type loss function to the mean squared error loss function. In the second type of steps, the k vectors which determine the partition are updated according to a multiclass classification rule, in the spirit of support vector machines. It is proved that both steps of the algorithm only require solving convex optimization problems, and that the algorithm is empirically consistent - as the length of the training sequence increases to infinity, fixed-points of the empirical version of the algorithm tend to fixed points of the population version of the algorithm. Learnability of the predictor class posit by the algorithm is also established.