

Multi-task Learning: Theory and Practice

Massimiliano Pontil (University College London)

We discuss the problem of estimating a structured matrix with a large number of elements. A key motivation for this problem occurs in multi-task learning. In this case, the columns of the matrix correspond to the parameters of different regression or classification tasks, and there is structure due to relations between the tasks. We present a general method to learn the tasks' parameters as well as their structure. Our approach is based on solving a convex optimization problem, involving a data term and a penalty term. We highlight different types of penalty terms which are of practical and theoretical importance. They implement structural relations between the tasks and achieve a sparse representations of parameters. We address computational issues as well as the predictive performance of the method. Finally we discuss how these ideas can be extended to learn non-linear task functions by means of reproducing kernels.