

Variable Selection Using Kullback-Leibler Divergence Loss

Dr. Shibasish Dasgupta, Lead Statistician,
Global Data Insight & Analytics,
Ford Motor Private Limited

The adaptive lasso is a recent technique for simultaneous estimation and variable selection where adaptive weights are used for penalizing different coefficients in the l_1 penalty. In this talk, we propose an alternative approach to the adaptive lasso through the Kullback-Leibler (KL) divergence loss, called the KL adaptive lasso, where we replace the squared error loss in the adaptive lasso set up by the KL divergence loss which is also known as the entropy distance. There are various theoretical reasons to defend the use of Kullback-Leibler distance, ranging from information theory to the relevance of logarithmic scoring rule and the location-scale invariance of the distance. We show that the KL adaptive lasso enjoys the oracle properties; namely, it performs as well as if the true underlying model were given in advance. Furthermore, the KL adaptive lasso can be solved by the same efficient algorithm for solving the lasso. We also discuss the extension of the KL adaptive lasso in generalized linear models (GLMs) and show that the oracle properties still hold under mild regularity conditions.