A Mutual Information Characterization of Sparse Signal Processing

Venkatesh Saligrama, George Atia

Suppose among a set of covariates $X_1, X_2, \ldots, X_d$ there is a subset $S$ of covariates that are salient for predicting outcomes $Y$. Specifically, we suppose $Y$ when conditioned on $\{X_k\}_{k \in S}$ is independent of the other covariates. Our goal is to identify the subset $S$ from data samples of the covariates and associated outcomes. We present a precise mutual information expression that characterizes the sample complexity for accurately identifying the subset $S$ for many interesting scenarios and derive precise sample complexity bounds for these cases.