On the Intrinsic Dimensionality of Multi-View Regression

Dean Foster
University of Pennsylvania

In the multi-view regression problem, we have a regression problem where the input variable can be partitioned into two different views, where it is assumed that either view of the example would be sufficient for learning — this is essentially the co-training assumption for the regression problem. For example, the task might be to identify a person, and the two views might be a video stream of the person and an audio stream of the person.

We show how Canonical Correlation Analysis, CCA, (related to PCA for two random variables) implies a ridge regression algorithm, where we can characterize the intrinsic dimensionality of this regression problem by the correlation of the two views. An interesting aspect of our analysis is that the norm used by the ridge regression algorithm is derived from the CCA — no norm or Hilbert space is assumed apriori (unlike in kernel methods).