GRAPHICAL MODEL SELECTION FOR LEARNING FROM BIG DATA OVER NETWORKS

ALEX JUNG

Modern technology is generating heterogeneous datasets at unprecedented speed and scale, i.e., big data. In many applications the observed datasets have an intrinsic network or graph structure. Exploiting the network structure results in efficient learning methods via scalable message passing algorithms. Moreover, exploiting the network-structure allows for accurate learning even in the high-dimensional regime.

While the graph structure is explicit in some applications, e.g., due to physical constraints, most often the underlying graph structure has to be learned from training data. In particular, by interpreting the datasets as realisations of a high-dimensional random processes, a particular graph structure is induced by conditional dependencies between the individual process components.

In this talk, I will present some of our recent work on the sample complexity for learning the conditional independence graph for high-dimensional nonstationary processes. Our results indicate that accurate learning is possible, even in the high-dimensional regime, whenever the underlying graph is sufficiently sparse.