Community detection on probabilistic graphical models with group-based penalties

A new strategy of probabilistic graphical modeling is developed that draws parallels from social network analysis. Probabilistic graphical modeling summarizes the information coming from multivariate data in a graphical format where nodes, corresponding to random variables, are linked by edges that indicate dependence relations between the nodes. The purpose is to estimate the structure of the graph (which nodes connect to which other nodes) when data at the nodes are available. On the opposite side of the spectrum, social network analysis considers the graph as the observed data. Given thus the graph where connections between nodes are observed rather than estimated, social network analysis estimates models that represent well an underlying mechanism which has generated the observed graph.

We propose a new method that exploits the strong points of each framework as it estimates jointly an undirected graph and communities of homogenous nodes, such that the structure of the communities is taken into account when estimating the graph and conversely, the structure of the graph is accounted for when estimating homogeneous communities of nodes. The procedure uses a joint group graphical lasso approach with community detection-based grouping, such that some groups of edges co-occur in the estimated graph. The grouping structure is unknown and is estimated based on community detection algorithms.

Theoretical derivations regarding graph convergence and sparsistency, as well as accuracy of community recovery are included, while the method’s empirical performance is illustrated in an fMRI context, as well as with simulated examples. Joint work with Gerda Claeskens.

Refreshments will be served following the seminar in 1181 Comstock Hall.