Seminaire Point de vue.
Peter Bühlmann.

Sparse graphs and high-dimensional causal inference

Understanding cause-effect relationships between variables is of interest in many fields of science. Inferring such causal relations from data is entirely different from estimating regression effects. We will discuss high-dimensional causal inference using sparse graphical modeling.

More details :
It is desirable to obtain causal information from observational data obtained by observing a system of interest without subjecting it to interventions (randomized experiments) ; or from a combination of observational and interventional (perturbation) data. When assuming no or little information about (causal) influence diagrams, the problem in its full generality is ill-posed. However, we will show how graphical modeling and intervention calculus can be used for quantifying useful bounds for causal effects or for inferring certain identifiable causal effects, even for the high-dimensional, sparse case where the number of variables can greatly exceed sample size.

The statistical methodology and theory is very different from more "main-stream" high-dimensional methods and analysis, mainly due to severe non-convexity inherent in estimation for (Markov equivalence classes of) directed acyclic graphs.

We validate the statistical causal inference method with gene intervention experiments in yeast and arabidopsis : strong effects can be detected and further used for prioritization of experiments (a much less ambitious but perhaps more realistic goal than inferring a causal regulatory structure of an entire network).