Vera Djordjilovic: Searching for a source of difference in Gaussian graphical models

Abstract: We consider a two-sample problem within the framework of Gaussian graphical models. When the global hypothesis of equality of two distributions is rejected, the interest is usually in localizing the source of difference. Motivated by the idea that diseases can be seen as system perturbations, and by the need to distinguish between the origin of perturbation and components affected by the perturbation, we introduce the concept of a minimal seed set, and its graphical counterpart a graphical seed set. We study their connection to the different issues related to high-dimensional inference, such as variable selection and discriminative dimension reduction. We propose a simple and computationally inexpensive testing procedure, linear in the number of nodes, to estimate the graphical seed set from data. We illustrate our approach in the context of gene set analysis, where the aim is to distinguish between the primary and the secondary dysregulation.