Name: Souvik Seal
Email: souvik.seal@cuanschutz.edu
PI Name: Katerina Kechris
PI email: katerina.kechris@cuanschutz.edu
Presentation preference: Online

Rapid Condition adaptive Fused Graphical Lasso

Souvik Seal[1], Qunhua Li[2], Elle Butler Basner[2], Laura Saba[3] and Katerina Kechris[1]

[1] Department of Biostatistics and Informatics, University of Colorado Anschutz Medical Campus; [2] Department of Statistics, Pennsylvania State University; [3] Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus

Inferring gene co-expression networks is important for understanding gene regulation and pathway activity. These networks are usually undirected graphs where genes are represented as nodes and an edge represents a significant co-expression relationship. When gene-expression data from multiple conditions (e.g., treatments, tissues, strains) are available, a joint estimation of networks harnessing shared information across them can significantly increase the power. Condition adaptive fused graphical lasso (CFGL) is one such method for joint estimation of networks that also accounts for condition specificity i.e., retains condition-specific patterns of co-expression which can provide insights into underlying mechanisms activated in a particular condition. However, the current algorithm is prohibitively slow even for a moderate number of genes and can only be employed for a maximum of three conditions. We propose a faster alternative of CFGL known as rapid condition adaptive fused graphical lasso (RCFGL). Along with being computationally feasible, RCFGL can also jointly analyze more than three conditions. We present a novel screening rule to determine if the full network estimation problem can be broken down into estimation of smaller disjoint sub-networks, improving RCFGL's performance further. We demonstrate the advantages of our method over CFGL and other existing approaches in different simulation scenarios. We also use RCFGL to jointly estimate the gene co-expression networks from three brain regions from a genetically diverse cohort of rats. We find that the network module common across all the regions contains genes relevant for transcription regulation. A C and Python based package implementing RCFGL is provided.