Sparse Regulation Networks

by
Ji Zhu
University of Michigan
439 West Hall
Ann Arbor, MI 48109
jizhu@umich.edu

Abstract

A significant amount of research in bioinformatics has recently concentrated on the problem of estimating "transcription regulation networks (TRN)." It is postulated that, in an organism such as E. Coli, the expression levels of each gene are controlled by the activation levels of known "transcription factors (TF)." The goal in estimating a TRN is to simultaneously compute the network topology and the strength of the relationships. In practice the values of the transcription factors are rarely observed so these must also be estimated based on observed values of the gene expression arrays over multiple experiments. In this paper, we take an approach that treats estimations of both the connections and the transcription factors as a variable selection problem. In this context our data has an extremely large number of variables, i.e. potential connections, but is sparse in terms of the number of "true" variables, i.e. connections that actually exist. Specifically, our! method utilizes L1 penalties on the connection strengths as well as the transcription factors. Similarly as an earlier Bayesian approach, our method also utilizes prior knowledge of the network structure. However, our approach is more computationally efficient which provides increased flexibility in determining the final network topology.