We present an algorithm to predict gene regulatory network structure, where gene expression time series data are regarded as vector trajectories of the sample obtained from a stochastic dynamical system defined on a graph.

We use a Gaussian Bayesian network for modeling a gene regulatory network structure. Here a node represents a gene, and a link between 2 nodes represents the relationship between the genes. The output of Gaussian Bayesian network is assumed to be a normal distribution. Since gene expression data may not follow a normal distribution, a certain transformation is critical for the performance of the prediction algorithm. Here, we utilize Box-Cox transformation to the gene expression data.

Our algorithm is built on a Bayesian scheme for learning the graph. Recent studies on the topological structure of different networks, including gene regulatory networks, have revealed the presence of the Zipf law, i.e., a power law. Therefore, the graph topology is assumed to follow a Zipf distribution. In the Bayesian framework, this assumption can be incorporated into the prior distribution.

In our algorithm, the appropriateness of the graph must be evaluated using the posterior mean. However, the computational complexity of the Bayesian learning method increases exponentially with the number of nodes. Here, we applied the Markov Chain Monte Carlo method to calculate the posterior mean. By using the Markov Chain Monte Carlo method, we can avoid the complex integration
which is necessary to evaluate the graph posterior mean. In addition, we applied the Markov Chain Monte Carlo method to predict the parameter of a Zipf distribution.

The performance of our algorithm is validated using synthesized data, and the algorithm appears functional. We also applied our algorithm to actual gene expression data. We have made several remarks on the feasibility of the predicted network from a biological viewpoint.

References and Literature for Further Reading


