

Learning Bayesian networks from microarray data using multivariate linear splines

Iosifina Pournara* & Lorenz Wernisch

Birkbeck College, University of London

Abstract

Bayesian networks have been used successfully in several studies in the reconstruction of gene regulatory networks from microarray data (Friedman et al, 2000). It has been observed frequently that modelling gene expression data on continuous domains gives better results than modelling them on discrete domains which requires arbitrary cutoffs and implies a loss of information. The disadvantage of current probabilistic scoring schemes for continuous domains is their assumption of a linear dependency of a child node on its parents, which is unrealistic since most regulatory relationships between genes are highly nonlinear. Nonlinear regression networks have been suggested where dependencies are modelled by nonlinear additive models using splines (Imoto et al, 2002) or Gaussian Processes (Nachman and Friedman, 2000). Multivariate linear splines have been suggested by Holmes and Mallick (2001) for nonlinear Bayesian regression. We developed a fully Bayesian framework for learning networks from multivariate data that goes beyond additivity and that models true interactions by multivariate linear splines. An MCMC approach allows us to calculate the posterior probability of arbitrary features of the network. We evaluate the algorithm on simulated as well as biological data.

References

- N. Friedman, M. Linial, I. Nachman, and D. Pe'er. Using Bayesian networks to analyze expression data. *J. Comput. Biol.*, 7:601–620, 2000.
- C. C. Holmes and B. K. Mallick. Bayesian regression with multivariate linear splines. *J. Roy. Statist. Soc. B*, 63:3–18, 2001.
- S. Imoto, T. Goto, and S. Miyano. Estimation of genetic networks and functional structures between genes by using Bayesian networks and nonparametric regression. In *Pac. Symp. Biocomput.*, 175–186, 2002.
- I. Nachman and N. Friedman. Gaussian process networks. In *Proc. Sixteenth Conf. on Uncertainty in Artificial Intelligence (UAI)*, 2000.