Bayesian Modelling of Cross-study Discrepancies in Gene Networks

Xiaodan Fan

Department of Statistics, the Chinese University of Hong Kong
Shatin, N.T., Hong Kong, *xfan@sta.cuhk.edu.hk*

Abstract. There are often multiple studies performed to investigate a same biological system from similar or related angles due to its high complexity. Many meta-analyses over these studies suggested that there are an excess of genes showing discordant gene expression across similar studies compared to what would have been predicted by chance alone. Scharpf et al. (2009) introduced a hierarchical Bayesian model for detecting differential gene expression in multiple data sets while allowing for cross-study discrepancy. Fan et al. (2009) and Fan et al. (2010) used a Bayesian approach to integrate cell-cycle microarray data sets and showed that the discrepancy about the cell-cycle regulated genes exists between individual laboratories and across synchronization techniques. In this paper, instead of dealing with the discrepancy at gene level as in Scharpf et al. (2009), we introduced a Bayesian approach to model the discrepancy at gene network level. The fundamental conjecture is that the gene expression discrepancy is resulted from the dynamics of the gene regulatory networks. Starting with different parameter settings, the network dynamics may show multiple steady states. Therefore, a gene can be highly expressed in one phenotype than the other in some studies, while the opposite is observed in other studies. Similarly, in cell-cycle experiments, a gene's expression can be highly periodic in some studies, while aperiodic in other studies. This phenomenon also exists in some stress response studies, where the lists of differential expressed genes for the same stimulus vary significantly across different study. The new Bayesian approach is applied on the time-series microarray data sets from fission yeast cell-cycle experiments. A gene network is inferred from the combined data. Its dynamics is simulated under the inferred parameter setting as well as other settings as an effort to explain the discrepancy observed in the cell cycle studies.

Keywords: gene network, meta-analysis, Bayesian computing, cell cycle

References

FAN, X. and LIU, J.S. (2009): Comment on "A Bayesian Model for Cross-Study Differential Gene Expression" by Scharpf, Tjelmeland, Parmigiani, and Nobel, *Journal of the American Statistical Association 104 (488), 1314-1318.*

FAN, X., PYNE, S. and LIU, J.S. (2010): Bayesian Meta-Analysis for Identifying Periodically Expressed Genes in Fission Yeast Cell Cycle. To appear in *Annals of Applied Statistics.*

SCHARPF, R.B., TJELMELAND, H., PARMIGIANI, G. and NOBEL, A. (2009): A Bayesian model for cross-study differential gene expression. *Journal of the American Statistical Association 104 (488), 1295-1310.*