

A State Space Representation of VAR Models with Sparse Learning for Dynamic Gene Networks

Kaname Kojima¹, Rui Yamaguchi¹, Seiya Imoto¹, Mai Yamauchi¹, Masao Nagasaki¹, RyoYoshida², Teppei Shimamura¹, Kazuko Ueno¹, Tomoyuki Higuchi², Noriko Gotoh¹, Satoru Miyano¹

¹Human Genome Center, Institute of Medical Science, University of Tokyo
²Institute of Statistical Mathematics

Recently, vector autoregressive (VAR) models are widely used for estimating gene regulatory networks. However, in VAR models, observation noise and system noise cannot be distinguished.

In addition, time course data used must be measured in equal time interval due to the model assumption of VAR. However, almost all time course microarray data are not measured in equal time interval. This study introduces a state space representation for considering observation noise and system noise separately and for dealing with unequal time interval data. While existing approaches using state space model assume modularity of gene regulatory networks to address high dimensional parameter estimation problem, the newly proposed model can deal with gene-gene interaction directly by sparse learning based on L1 regularization. However, in a simple implementation the proposed model requires the calculation of large inverse matrices in a large number of times during parameter estimation process based on EM algorithm. This limits the applicability of the proposed method to a relatively small gene set. We thus introduce a new calculation technique for EM algorithm that does not require the calculation of inverse matrices. The proposed method is applied to time course microarray data of human lung cells stimulated by EGF and/or anticancer drug.