Bayesian regularization of non-homogeneous dynamic Bayesian networks by globally coupling interaction parameters

Marco Grzegorczyk  
Department of Statistics, TU Dortmund, Germany  
grzegorczyk@statistik.tu-dortmund.de

Dirk Husmeier  
School of Mathematics and Statistics, Glasgow University, UK  
dirk.husmeier@glasgow.ac.uk

Abstract: To relax the homogeneity assumption of dynamic Bayesian networks (DBNs), various recent studies have combined DBNs with multiple changepoint processes. The underlying assumption is that the parameters associated with time series segments delimited by multiple changepoints are prior independent. Under weak regularity conditions, the parameters can be integrated out in the likelihood. However, the assumption of prior independence is unrealistic in many real-world applications. We therefore propose a Bayesian coupling scheme to introduce information sharing among the segment-specific interaction parameters. We investigate the effect of this model improvement has on the network reconstruction accuracy in a reverse engineering context, where the objective is to learn the structure of a gene regulatory network.

MOTIVATION: Non-homogeneous (NH) dynamic Bayesian networks (DBNs) combine a node-specific multiple changepoint process with a conventional DBN. Model parameters, and possibly network structures, are allowed to vary between time series segments.


Shortcoming of these methods: no information sharing among time series segments, leading to large posterior uncertainty (Bayesian) or overfitting (classical) when data are sparse.

Objective: Incorporate prior knowledge that changes are gradual, while maintaining the changepoint structure for analytically tractability.

BACKGROUND: We formulate the NH-DBN in terms of a Bayesian piecewise linear regression model.

Time-shifted target: \( y_{h,k} \) node (gene) time segment change points

Design matrix: \( X_{p,h,k} = (x_{p, \tau(h,k) + 1,}, \ldots, x_{p, \tau(h,k) + h}) \)

Parent nodes time segment change points

Noise model:

\[ P(y_{h,k}, X_{p,h,k}, w_{h,k}, \sigma^2_{\epsilon,h,k}) = N(X^T_{p,h,k} w_{h,k}, \sigma_{\epsilon,h,k}^2) \]

Interaction parameter prior:

\[ P(w_{h,k}, m_{p,h,k}, \sigma^2_{g,h,k}, \delta_{g,h,k}) = N(w_{h,k}, m_{p,h,k}, \sigma^2_{g,h,k}, C_{g,h,k}) \]

The computation of the marginal likelihood is analytically tractable:

\[ P(y_{h,k}, X_{p,h,k}, \sigma^2_{\epsilon,h,k}) = \int P(y_{h,k}, X_{p,h,k}, w_{h,k}, \sigma^2_{\epsilon,h,k}) \]

RESULTS: The results show that for low perturbations \((c = 0)\), the proposed method outperforms uncoupled NH-DBNs, while for large perturbations \((c = 1)\) it outperforms conventional homogeneous DBNs.

Figure: Network reconstruction accuracy (average AUC-ROC scores) for 3 signal-to-noise ratios (SNRs). The global parameter vector of amplitude \(\nu\) was perturbed with random perturbations of amplitude \(\epsilon\) (abscissa).

The top row shows the total AUC-ROC scores.

The centre row and the bottom row show the AUC-ROC differences.

The results above were obtained with a specific choice of the highest-level hyperparameters: \(B_0 \sim \text{Gam}(1,200)\) and \(B_0 \sim \text{Gam}(200,1000)\).

The proposed method is robust with respect to a variation of the level-3 hyperparameters as long as the prior on the scale of the coupling strengths \(\delta^2\) is not too vague.

METHODOLGY: The proposed model can be illustrated with a probabilistic graphical model and plate diagram:

Various coupling schemes can be distinguished:

Overview of the coupling schemes (S1-SB) for the noise variance hyperparameters.

No coupling: The noise variance hyperparameters are d-separated, i.e., they have separate level-2 hyperparameters which are fixed.

Weak coupling: The noise variance hyperparameters are not d-separated, i.e., they share a set of common level-2 hyperparameters which are flexible.

Hard coupling: There are common noise variance hyperparameters (with fixed level-2 hyperparameters).

SYNTHETIC BIOLOGY IN YEAST: Cantone et al. (2009) synthetically designed a gene network in Saccharomyces cerevisiae (yeast).

For the Saccharomyces cerevisiae time series we assume the segmentation to be unknown so that the changepoints have to be inferred from the data. To this end, we impose a negative binomial distribution \(\text{NBBin}(p,k)\) with hyperparameters \(p\) and \(k\) onto the distance between neighbouring changepoints.

We vary the hyperparameter \(p\) to obtain different segmentations (i.e. numbers of changepoints per gene).

As a consequence of the proposed coupling scheme, we obtain increased robustness with respect to a variation of the hyperparameters of the changepoint process.

Figure: Network reconstruction accuracy of the coupled and the uncoupled NH-DBN plotted against the mean number of changepoints per gene.

Simulated data were generated from an NH-DBN with segment-specific parameters of the following form: For each node \(g\):

\[ w_{g,h,k} \sim N(0,1), \quad w_{g,h,k} = w_{g,h,k}^2/2 \]

and then for each node-specific segment \(h\):

\[ w_{g,h,k} \sim N(0,1), \quad w_{g,h,k} = w_{g,h,k}^2/2 \]

Time series of length \(m=41\) with 4 equidistant time segments \((h=1,2,3,4)\) were generated. Observational white Gaussian noise at a target signal-to-noise ratio (SNR) was added.

As the underlying network:

Figures: The RAF protein signaling pathway as reported in Sachs et al. (Science, 2005). The RAf network consists of 11 nodes (proteins) and 20 directed edges.