Non-stationary continuous dynamic Bayesian networks

**Introduction:** Dynamic Bayesian networks have been applied widely to reconstruct the structure of regulatory processes from time series data. The standard approach is based on the assumption of a homogeneous Markov chain, which is not valid in many real-world scenarios. We propose a non-stationary dynamic Bayesian network for continuous data, in which parameters are allowed to vary among segments, and in which a common network structure provides essential information sharing across segments. Our approach is based on a Bayesian multiple change point process. Conditional on the change points, the parameters can be integrated out in closed form. As in Nobile and Nobile (2007) the number and location of the change points are then sampled from the posterior distribution with reversible jump (RJ) Markov chain Monte Carlo (MCMC).

The proposed cpBGe model:

\[
P(\text{graph, } K, V, \text{data}) = P(\text{graph}) \cdot P(K) \cdot P(V | K) \cdot \prod_{n=1}^{N} P(\text{data}(n) | \text{graph})
\]

- \( K \): number of mixture components
- \( V \): vector of latent allocation variables
- \( \text{data}(n) \): subset of data points allocated to component \( k \)

The Bayesian Gaussian Mixture (BGM) Bayesian network model:

\[
P(\text{graph, } K, V, \text{data}) = P(\text{graph}) \cdot P(K) \cdot P(V | K) \cdot \prod_{n=1}^{N} P(\text{data}(n) | \text{graph})
\]

- \( K \): number of components for \( X_n \)
- \( V_n \): allocation vector for \( X_n \)
- \( \text{data}(n) \): data relevant for the \( n \)-th local score \( P(\text{data}(n)) \)

**Network reconstruction accuracy for synthetic non-linear network data in terms of AUC scores:**

**Application:** circadian genes in *Arabidopsis thaliana*

**Related models:**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>Marginal Likelihood</td>
<td>Marginal Likelihood</td>
<td>Marginal Likelihood</td>
<td>Marginal Likelihood</td>
<td>BIC</td>
</tr>
<tr>
<td>Node-specific change points</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Time constant</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Data format</td>
<td>Continuous</td>
<td>Discrete</td>
<td>Continuous</td>
<td>Continuous</td>
<td>Discrete</td>
</tr>
<tr>
<td>Latent variable format</td>
<td>Change-point process</td>
<td>Change-point process</td>
<td>Change-point process</td>
<td>Change-point process</td>
<td>Discrete allocation</td>
</tr>
</tbody>
</table>

**References:**


**C: Empircische Modellbildung**