Dynamic Bayesian Gaussian Graphical Models for Inferring Evolving Network Structure

Slav Kirov[1], Micol Marchetti-Bowick[2], Willie Neiswanger[2]
[1]Department of Mathematical Sciences and [2]Machine Learning Department, Carnegie Mellon University

ABSTRACT

We propose a method for learning the structure of evolving Gaussian graphical models via Bayesian inference. Gaussian graphical models (GGMs) are often used to model the structure of a network, where they have found applications in computer vision (object pose), finance (relationships among stocks), biology (gene networks), and a number of related fields.

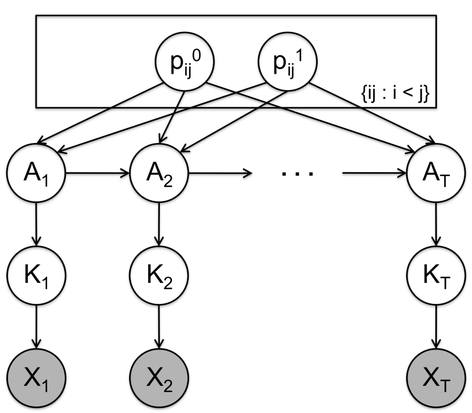
INTRODUCTION

Learning the structure of a network is a problem of interest in a variety of fields. For a biologist, uncovering the structure of a gene network that characterizes functional associations between genes can provide insight into how genes interact during the regulation of a biological process. Similarly, a financial analyst may be interested in understanding how the prices of different stocks relate to one another. In both of these settings, learning a dynamic network that evolves over time can provide even more valuable information about how these relationships change in time.

MODEL

Our generative model is a dynamic Bayesian network (DBN) that incorporates a Gaussian graphical model (GGM) at each time step, where

- $A_t = \{A^t_{ij} : i<j\}$: adjacency matrix for the graph at time t.
- K_t: precision matrix for multivariate Normal distribution.
- X_t: observations / emissions from model.



$$p^0_{ij} \sim \text{Beta}(\alpha^0, \beta^0) \qquad p^1_{ij} \sim \text{Beta}(\alpha^1, \beta^1)$$
$$A^t_{ij} \mid (A^{t-1}_{ij} = 0) \sim \text{Bernoulli}(p^0_{ij}) \qquad A^t_{ij} \mid (A^{t-1}_{ij} = 1) \sim \text{Bernoulli}(p^1_{ij})$$
$$K_t \mid A_t \sim \text{G-Wishart}_{A_t}(b, D) \qquad X_t \mid K_t \sim \text{Normal}(0, K_t^{-1})$$

Where:

$$p^0_{ij} = P(A^t_{ij} = 1 \mid A^{t-1}_{ij} = 0) \qquad 1 - p^0_{ij} = P(A^t_{ij} = 0 \mid A^{t-1}_{ij} = 0)$$
$$p^1_{ij} = P(A^t_{ij} = 0 \mid A^{t-1}_{ij} = 1) \qquad 1 - p^1_{ij} = P(A^t_{ij} = 1 \mid A^{t-1}_{ij} = 1)$$

INFERENCE

We experiment with two different inference methods for estimating the posterior distribution of our dynamic model.

Collapsed Gibbs Sampling

First, we implemented a Gibbs sampler in which we integrate out the precision matrix variables $K_{\{1:T\}}$ and resample each edge at each time point, A^t_{ij}, using the Gibbs proposal distribution.

$$P(A^t_{ij} | rest \setminus K_{1:T}) = \frac{1}{Z} P(X_t|A_t) P(A^t_{ij}|A^{t-1}_{ij}) P(A^{t+1}_{ij}|A^t_{ij}) \quad (1)$$

$$= \frac{1}{Z} P(A^t_{ij}|A^{t-1}_{ij}) P(A^{t+1}_{ij}|A^t_{ij}) \int P(X_t, K_t|A_t) dK_t \quad (2)$$

$$= \frac{1}{Z} P(A^t_{ij}|A^{t-1}_{ij}) P(A^{t+1}_{ij}|A^t_{ij}) \int P(X_t|K_t) P(K_t|A_t) dK_t \quad (3)$$

$$= \frac{1}{Z} P(A^t_{ij}|A^{t-1}_{ij}) P(A^{t+1}_{ij}|A^t_{ij}) (2\pi)^{-\frac{np}{2}} \frac{I_G(b+n, D+S)}{I_G(b, D)} \quad (4)$$

In (4), $I_G(b, D)$ is the normalization constant of the prior G-Wishart distribution $P(K_t)$ and $I_G(b+n, D+S)$ is the normalization constant of the posterior G-Wishart distribution $P(K_t \mid X_t)$. These constants are computed using Monte Carlo simulation.

Sequential Monte Carlo

Second, we implemented a version of Sequential Monte Carlo (SMC) that builds on the Birth-Death MCMC inference method proposed by Mohammadi & Wit (2012). We chose SMC because it performs inference in an online fashion and is thus well-suited to time-varying models. At a high level, SMC proceeds with the following steps:

SMC Algorithm:

1. At the first time step, run BD-MCMC using X_1 as input and get N samples from the partial posterior $P(A_1, K_1 \mid X_1)$ as output
2. At time t-1, we have N particles (samples of $A_{1:t-1}, K_{1:t-1}$) that are approximately distributed according to $P(A_{1:t-1}, K_{1:t-1} \mid X_{1:t-1})$
3. At time t, run a modified version of BD-MCMC to sample A_t and K_t from proposal distribution Q, where:

$$Q(A_t, K_t \mid A_{1:t-1}, X_{1:t}) = P(A_t, K_t \mid A_{t-1}, X_t)$$

4. Extend the particles with these new values
5. Compute the importance weight of each particle
6. Resample the N particles according to a categorical distribution made up of the normalized importance weights
7. Output N samples from the full posterior $P(A_{1:T}, K_{1:T} \mid X_{1:T})$

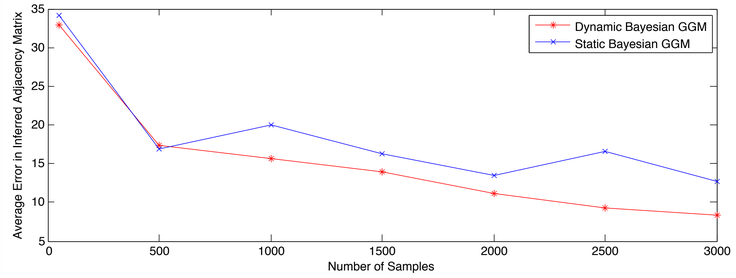
In step (3), we sample from the proposal distribution $P(A_t, K_t \mid A_{t-1}, X_t)$ by modifying BD-MCMC to incorporate the prior over the graph structure $P(A_t \mid A_{t-1})$. BD-MCMC traverses the space of hidden variables by executing birth moves (adding an edge) and death moves (removing an edge), and re-sampling the precision matrix after each move.

RESULTS

We carry out several experiments on both synthetic and real data to evaluate the success of our method.

Synthetic Data Experiment

We performed SMC inference on a synthetic time-varying network with 12 nodes over 10 time-steps, in order to judge performance. We compared its algorithm error with that of running BD-MCMC independently at each time step. The figure below shows that the dynamic model was better able to recover the correct network structure.



Finance and Genetic Data Experiments

We perform inference on two real-world datasets. In the first, we infer the structure of a GGM for a set of 20 companies over a sequence of 10 weeks based on their stock prices, and in the second, we infer a network of 25 genes as the network varies over a sequence of 9 cell types.

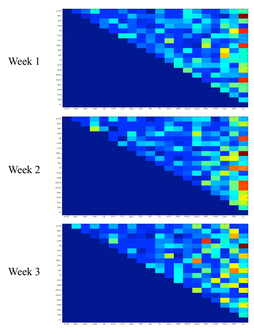


Figure 1: Heatmaps showing posterior probability of edges in stock network over a sequence of 3 weeks.

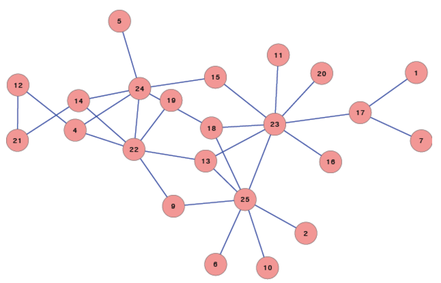


Figure 2: Inferred genetic network for a single cell type.

CONCLUSIONS

We have derived a time-varying, Bayesian model for inferring the structure of a sequence of Gaussian graphical models. By running this model on synthetic data (for which there exists ground-truth) we have shown that our model outperforms a static equivalent. We have also demonstrated our model on multiple real-life datasets in order to show its ability to infer network structure in the fields of both finance and genetics. We hope that pursuing Gaussian graphical model learning in a Bayesian setting will allow for flexible specification of dependencies, especially of the dynamics of time-varying graphs