MIT 6.882 Final Project

(Jeremiah) Zhe Liu , Will Townes

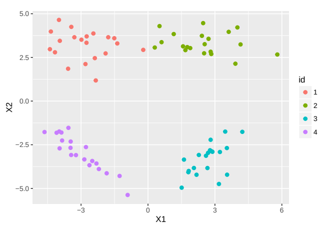
April 14, 2016

Contents

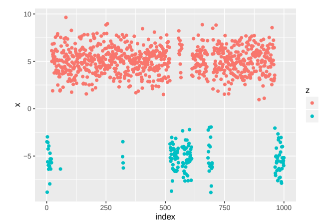
1 Current Progress

1.1 Dirichlet Process Sampler

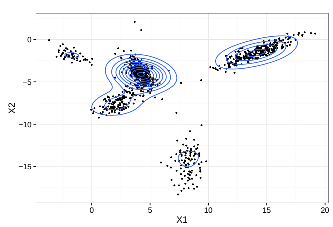
We have completed the code for sampling from the DP prior. We have also created the simulated data (Figure 1a, 1c, 2b). We are about half-way through the Gibbs Sampler for sampling from the DP posterior. One obstacle here was we weren't sure whether to use the collapsed sampler (Chinese Restaurant Process) or a blocked sampler. It seems that Fox et al [1] used the latter with a truncated stick-breaking process. Will worked on this part.



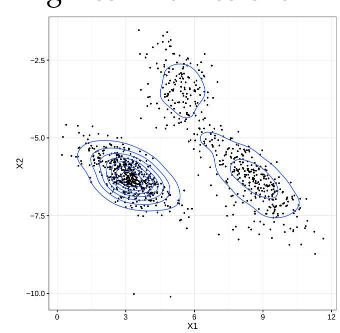
(a) Test Data for Dirichlet Process Mixture Model: A four-component Gaussian Mixture Model



(b) Test Data for Hidden Markov Model: two hidden states with gaussian emissions



(c) Samples from DP mixture of Gaussians Prior



(d) Samples from DP mixture of Gaussians Prior

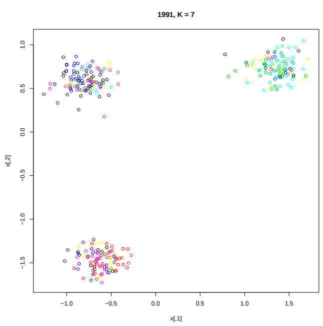
Figure 1: Test Data/Generated Samples from MDP and HMM models

1.2 Hierarchical Dirichlet Process Sampler

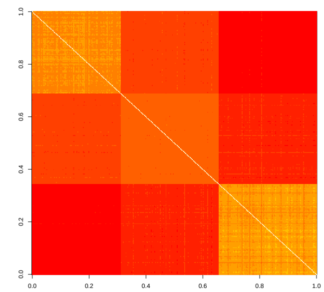
A modularized implementation of HDP sampler (using the direct assignment approach as described in Teh (2006)[2]) has been completed and available at GitHub project page. We have performed preliminary trials on clustered Gaussian data (Figure) with promising results (Figure). However, our current implementation suffers some classical problems with DP sampler, such as slow mixing rate for the number of underlying modes, and over-estimation of the number of clusters. We plan to further strengthen our implementation by adding:

1. Block Sampler:
   Improve the mixing rate by using a finite-dimensional approximate of DP prior as discussed in Ishwaran and James (2001). [3]

2. Sampling concentration parameters:
   Since model performance depend largely on the specification of γ and α, it is advisable to sample the two concentration parameters as well as discussed in Teh (2006) [2].



(a) Estimated Cluster Assignment of Gaussian Clustering Data in Iteration 1991/2000, with 7 clusters



(b) Estimated Correlation Matrix based on subject cluster assignment. Lighter color indicates stronger correlation.

1.3 Linear Dynamical System

We didn't start on this part yet, although we have done some reading in the Murphy book to gain familiarity.

1.4 Hidden Markov Model

We have completed the code for generating data from HMMs. We are about half-way through the Gibbs Sampler for sampling from HMM posterior. The sampler we are using is the same as in the Fox et al paper, where we compute the backward messages and then sample from the forward conditional distributions. Will worked on this.

2 Upcoming Progress

The hierarchical dirichlet process sampler has emerged as the most difficult component of the project. Fox et al. glossed over many complicated details for sampling from the full conditionals which we needed to dig out of supplements to other papers. However, Jeremiah has spent a lot of time working through the theory part and we anticipate making faster progress going forward. Sampling from the matrix-normal inverse wishart distribution and the linear dynamical system components are the major remaining pieces. The automatic relevance determination prior does not seem too problematic since it is just a conditionally conjugate prior. We probably will not try the Firefly MCMC approach since our data is low-dimensional.

3 Plan and Deadlines

Key internal deadlines are listed in parentheses.

1. (3/21)
   Code and visualize a dirichlet process sampler. Visualize with a simple two-dimensional clustering simulation.
2. (3/28)
   Code and visualize a"sticky" hierarchical dirichlet process sampler. Visualize via a two-dimensional clusters-within-clusters simulation.
3. (4/4)
   Write a sampler for an ordinary (non-switching) linear dynamical system and test with simulation.
4. (4/11)
   Write a sampler for a simple hidden markov model with a fixed number of modes.
5. (4/18)
   Finish sampler from DP and HMM (Will). Finish sampler for HDP (Jeremiah). Write sampler for generative model of LDS (Will).
6. (4/25)
   Integrate HDP with HMM (Jeremiah). Integrate HMM with LDS (Will).
7. (5/2)
   Combine everything together (HDP-HMM-LDS). Prepare draft of final report.

References

[1] Emily Beth Fox. *Bayesian nonparametric learning of complex dynamical phenomena*. Thesis, Massachusetts Institute of Technology, 2009.

[2] Yee Whye Teh, Michael I. Jordan, Matthew J. Beal, and David M. Blei. Hierarchical Dirichlet Processes. *Journal of the American Statistical Association*, 101(476):1566–1581, 2006.

[3] Hemant Ishwaran and Lancelot F. James. Gibbs Sampling Methods for Stick-Breaking Priors. *Journal of the American Statistical Association*, 96(453):161–173, March 2001.