

Bayesian Inference for Switching State Space Models

Nicholas Hoernle

March 22, 2018

1 Abstract

Switching linear dynamical systems (SLDS) are a class of models for time-series data where the parameters controlling a linear dynamical system switch according to a discrete latent process. SLDSs (also termed switching state space models (SSSMs)) combine hidden Markov and state space models to capture *regime* switching in a non-linear time series. The intuition is that a dynamical system evolves over time but may undergo a regime change that informs an intrinsic shift in the system's characteristics. Allowing for discrete points in time where the dynamics change, enhances the power of simple linear dynamical models to capture more complicated dynamics. I propose that the decomposition of the complicated time series into periods of linear dynamics also helps to increase interpretability into the complex system that produced the time series. Following Fox et al. [1], this project implements a version of the *sticky* hierarchical Dirichlet process (HDP) for Hidden Markov Models (HMM) [2] that is tailored for SLDSs. The model is an extension of the HDP-HMM that Teh et al. [5] propose in that it encourages regimes to remain active for longer periods of time (and reduces the tendency of the HDP-HMM to rapidly switch between a larger number of states). Tailoring the HDP-HMM for SLDS is a challenge due to the time dependent observations in each regime. Fox et al. [1] introduce the sticky HDP-HMM for SLDS in general and for vector autoregressive processes (AR) as a useful subset of the general SLDS. I follow Fox's model specification and implement a blocked Gibbs sampler to perform posterior inference in the HDP-AR-HMM and HDP-SLDS models. Classical approaches [4, 3] to performing inference in SLDS assume the known number of models (also termed modes and regimes). These bodies of work appear to be disjoint and thus I will conduct similar experiments to those by both Fox et al. [1] and Ghahramani and Hinton [3] to compare the two approaches in terms of their success in performing inference over the latent switching variable and over the regime specific parameters.

2 Proposal details

2.1 Project timeline and work milestones:

1. Study hierarchical Dirichlet processes [5] and sticky hierarchical Dirichlet processes [2]. Investigate the use of Stan for inference when the *weak limit* approximation is made (see Section 2.3). **Due: Monday 9 April.**
2. Study and **implement** HDP-SLDS and HDP-AR-HMM models[1]. **Due: Monday 23 April.**

3. Design and implement an experiment to compare the inference of the models in (2) vs previous literature for switching state space models (namely the work by Kim [4]). This will involve simulating data from a number of different generative models (e.g. such as the experiments that Fox [1] conducted). **Due:** *Monday 30 April.*
4. Investigate the interpretability of the model for uncovering latent structure from complex time series data. I can use standard data from the field (e.g. bee dancing data) or I can use data from an educational dataset that I have been working with. I am interested in segmenting the timeseries response from an educational simulation into periods that teachers are able to use to create salient reviews of students' work. **Due:** *Friday 4 May.*

2.2 Project deliverables

1. Implementation of sticky hierarchical Dirichlet processes tailored for switching linear dynamical systems and vector autoregressive regimes [1].
2. Experimental design to compare the above implementation and the gaussian merging approach for pseudo-maximum likelihood parameter and state estimation that is implemented in *statsmodels* [4]. This experiment will primarily consist of data that is generated under a number of different settings and I will evaluate how effectively the algorithms are able to recover the regime switches. Ghahramani plots a histogram of the percent agreement between the model and the true data. Fox investigates the Hamming distance between the inferred and the true label assignments.

2.3 Project risks

I am implementing inference algorithms for two unknown models. This will involve programming the Gibbs samplers for both the HDP-SLDS and the HDP-AR-HMM. I will work with synthetic test data such that I am able to effectively evaluate the performance of my implementations (and to debug the samplers). Fox et al. use the *weak limit* of the Dirichlet process by approximating the infinitely many modes with a finite mode limit (i.e. an upper bound of L modes in the model). Making this assumption will allow me to prototype the sampling inference scheme in Stan/Edward/PyMC3 for quicker inference implementation. I am still investigating this possibility but I believe it might present a good risk mitigation possibility should I need it.

References

- [1] E. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky. Bayesian nonparametric inference of switching dynamic linear models. *IEEE Transactions on Signal Processing*, 59(4):1569–1585, 2011.
- [2] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky. An hdp-hmm for systems with state persistence. In *Proceedings of the 25th international conference on Machine learning*, pages 312–319. ACM, 2008.
- [3] Z. Ghahramani and G. E. Hinton. Variational learning for switching state-space models. *Neural computation*, 12(4):831–864, 2000.
- [4] C.-J. Kim, C. R. Nelson, et al. State-space models with regime switching: classical and gibbs-sampling approaches with applications. *MIT Press Books*, 1, 1999.
- [5] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Sharing clusters among related groups: Hierarchical dirichlet processes. In *Advances in neural information processing systems*, pages 1385–1392, 2005.