Bayesian Learning and Inference in Recurrent Switching Linear Dynamical Systems

Charbel Abi Younes, Marvyn Bailly,
Bart Boom, Rohin Gilman, Daran Xu

April 29, 2023

UNIVERSITY *of*
WASHINGTON | Applied Mathematics

Table of Contents

Rohin
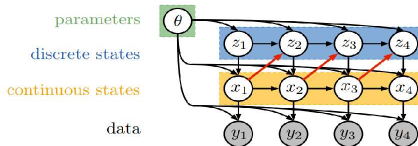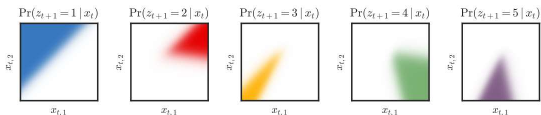
Model

SLDS and rSLDS

- Observation $y_t = Cx_t + d + w_t$, $w_t \overset{\text{iid}}{\sim} \mathcal{N}(0, S)$
- Continous latent state $x_{t+1} = A_{z_{t+1}} x_t + b_{z_{t+1}} + \nu_t$, $\nu_t \overset{\text{iid}}{\sim} \mathcal{N}(0, Q_{z_{t+1}})$
- Discrete latent state $z_t \in \{1, \ldots, K\}$
  - SLDS [1] $z_{t+1}|z_t \sim \pi_{z_t}$
  - rSLDS [2] $z_{t+1}|z_t, x_t \sim \pi_{SB}(\nu_{t+1})$, $\nu_{t+1} = R_{z_t} x_t + r_{z_t}$

Model

Stick Breaking Logitstic Regression [3]

- $p(z|x) \sim \pi_{SB}(\nu), \ \nu = Rx + r$
- Link function: $\pi_{SB}(\nu) = \left(\pi_{SB}^{(1)}(\nu), \ldots, \pi_{SB}^{(K)}(\nu)\right)$ (with $\sigma(x) = \frac{e^x}{1+e^x}$)

$$\pi_{SB}^{(k)}(\nu) = \begin{cases} \sigma(\nu_k) \prod_{j<k} \sigma(-\nu_j), & \text{if } k = 1, \ldots, K-1 \\ \prod_1^K \sigma(-\nu_j), & \text{if } k = K \end{cases}$$

- $p(z|x) \sim \prod_{k=1}^K \sigma(\nu_k)^{\mathbb{1}[z=k]} \sigma(-\nu_k)^{\mathbb{1}[z>k]}$ (Likelihood)

- With a Gaussian Prior $p(x)$, the posterior $p(x|z) \propto p(x) * p(z|x)$ is non-Gaussian

- Bayesian updating is not efficient

  - Gibbs Sampler: sampling $x^{(i+1)} \sim p(x|z = z^{(i)})$, sampling $z^{(i+1)} \sim p(z|x = x^{(i+1)})$

  - Message Passing $m_{t \to (t+1)}(x_{t+1}) =$ $\int \psi(x_t, y_t) \, \psi(x_t, z_{t+1}) \psi(x_t, x_{t+1}, z_{t+1}) \, m_{(t-1) \to t}(x_t) \, \mathrm{d}x_t$

Model

Polya-gamma augmentation [3]

$$\frac{(e^\nu)^a}{(1+e^\nu)^b} = 2^{-b}\int_0^\infty e^{\kappa\nu}e^{-\omega\nu^2/2}p_{\mathsf{PG}}\,(\omega\mid b,0)\mathrm{d}\omega\,\left(\kappa = a - \frac{b}{2}\right) \tag{2.1}$$

$$p(x_t|z_{t+1}) \propto \prod_{k=1}^{K-1}\frac{(e^{\nu_{t+1,k}})^{\mathbb{I}[z_{t+1}=k]}}{(1+e^{\nu_{t+1,k}})^{\mathbb{I}[z_{t+1}\geq k]}} \tag{2.2}$$

$$p(x_t|z_{t+1}) = \int p(x_t,\omega_t|z_{t+1})\mathrm{d}\omega_t = \int p(x_t|z_{t+1},\omega_t)p(\omega_t)\mathrm{d}\omega_t \tag{2.3}$$

$$\omega_{t,k}\mid x_t,z_{t+1} \sim \mathsf{PG}\left(\mathbb{I}\left[z_{t+1}\geq k\right],\nu_{t+1,k}\right) \tag{2.4}$$

- $p(x_t|z_{t+1},\omega_t)$ is Gaussian

- Message Passing $\psi(x_t,z_{t+1},\omega_t) \propto \mathcal{N}(\nu_{t+1}|\Omega_t^{-1}\kappa_{t+1},\Omega_t^{-1})$

- Thus instantiating these auxiliary variables in a Gibbs sampler enables efficient block updates

Rohin

Rohin

This is a frame with a title and a subtitle.

This is a frame with no title.

This is a frame with no header.

This is underlined text.

This is italic text.

This is bold text.

This is `mono-spaced` text.

Lists

This is a bulleted list:

- An item
- Another item
    - A subitem
    - Another subitem

This is a numbered list:

1. An item
2. Another item
    2.1 A subitem
    2.2 Another subitem

Colored Blocks

Block Title

This is a block.

Alert Block Title

This is an alertblock.

Example

This is an example block.

This is a frame the second section.

This is a frame with a...

This is a frame with a...

\pause.

Bibliography

📄 Guy A. Ackerson and King-Sun Fu.

On state estimation in switching environments.

IEEE Transactions on Automatic Control, 15(1):10–17, 1970.

📄 David Barber.

Expectation correction for smoothed inference in switching linear dynamical systems.

Journal of Machine Learning Research, page 7(Nov):2515–2540, 2006.

📄 Scott W Linderman, Matthew J Johnson, and Ryan P Adams.

Dependent multinomial models made easy: Stick-breaking with the polya-gamma augmentation.