Temporal Difference Variational Auto-encoder

Karol Gregor George Papamakarios Frederic Besse Lars Buesing Thophane Weber
DeepMind

Jan 2019

1. Motivation

The authors propose that it is essential to incorporate the following characteristics in a model for dynamical systems:

1 **Abstraction at the state level:** latent dynamical systems

2 **Belief Representation:** a deterministic coding compressing the observations up to time step t through *LSTMs*

3 **Jumpy State Predictions:** modeling $p(z_{t_2}|z_{t_2}), t_2 > t_1$ in the generative transitions .

1.1. Related Works - factorization of the posterior

In many related works for modeling dynamical systems, the approximate posterior distribution $q(z_{1:T}|x_{1:T})$ is factorized autoregressively as below:

$$q(z_{1:T}|x_{1:T}) = \prod q(z_t|z_{t-1}, \phi(x_{1:t})) \tag{1}$$

Using this autoregressive decomposition, one can formulate the L_{ELBO} as:

$$\log p(x_{1:T}) = \mathop{\mathbb{E}}_{z\sim q(z_{1:T}|x_{1:T})} \big[\sum_{t=1}^{T} \log p(x_t|z_t) + \log p(z_t|z_{t-1}) + \log q(z_t|z_{t-1}, \phi(x_{1:t})) \big] \tag{2}$$

The authors claim that this is an issue since the uncertainty of the marginal posterior of z_t could be leaked in the sample z_{t-1}. What I believe this means is, that basing the estimation of the next state z_t on one sample of the previous state z_{t-1} and hoping to capture all the information about the previous observations is not enough. For an accurate posterior estimation thus, either ancestral sampling is necessary (e.g through *particle filters*) or we use *belief states*.

2. Problem Statement

The goal we are working towards is:

- Estimating the parameters of a State Space Model 1
- Additionally, we want to estimate the posterior probability distribution over states $z_{1:T}$ given just the observations $x_{1:T}$ and (optional) actions $u_{1:T}$.

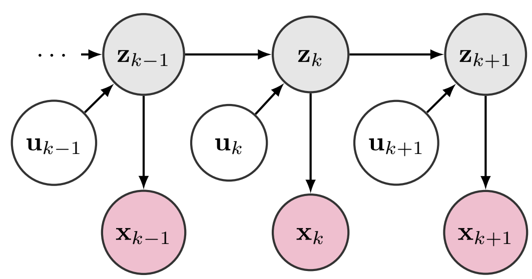

Figure 1. State Space Model

3. A Lowerbound to the Autoregressive Decomposition of $p(x_{1:T})$

Instead of a lower-bound on the likelihood of the complete data sequences, the authors propose a sum of lower-bounds on the factors of an autoregressive decomposition of $p(x_{1:T})$:

$$\begin{aligned}\log p(x_{1:T}) &= \sum_{t=1}^{T} \log p(x_t|x_{<t}) \\ &\geq \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z_{t-1},z_t\sim q(z_{t-1},z_t|x_{\leq t})} \big[\log p(x_t|z_{t-1}, z_t, x_{<t}) \\ &+ \log p(z_{t-1}, z_t|x_{<t}) - \log q(z_{t-1}, z_t|x_{\leq t}) \big]\end{aligned} \tag{3}$$

With the markovian assumptions in place and by choosing the decomposition of the posterior as:

$$q(z_{t-1}, z_t|x_{\leq t}) = \underbrace{q(z_t|x_{\leq t})}_{\text{belief}}\underbrace{q(z_{t-1}|z_t, x_{\leq t})}_{\text{one-step smoothing}} \tag{4}$$

we reformulate 3 as:

$$\begin{aligned} \log p(x_{1:T}) \geq \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z_{t-1}, z_t \sim q(z_{t-1}, z_t|x_{\leq t})} \Big[\log p(x_t|z_t) \\ + \log p(z_{t-1}|x_{\leq t-1}) + \log p(z_t|z_{t-1}) \\ - \log q(z_t|x_{\leq t}) - \log q(z_{t-1}|z_t, x_{\leq t}) \Big] \end{aligned} \tag{5}$$

Taking belief as $b_t = \underbrace{f_\theta(x_t, b_{t-1})}_{\text{hidden state of an RNN}}$

$$\begin{aligned} \log p(x_{1:T}) \geq \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\substack{z_t \sim p_B(z_t|b_t) \\ z_{t-1} \sim q(z_{t-1}|z_t, b_t, b_{t-1})}} \Big[\log p(x_t|z_t) \\ + \log p_B(z_{t-1}|b_{t-1}) + \log p(z_t|z_{t-1}) \\ - \log p_B(z_t|b_t) - \log q(z_{t-1}|z_t, b_t, b_{t-1}) \Big] \end{aligned} \tag{6}$$

TODO: Check if these equations 3, 5, 6 are correct

3.1. Jumpy version

As per the motivation in point 3 in Section 1, the term wise loss is re-framed as:

$$\begin{aligned} -\mathcal{L}_{t_1,t_2} \geq \mathop{\mathbb{E}}_{\substack{z_{t_2} \sim p_B(z_{t_2}|b_{t_2}) \\ z_{t_1} \sim q(z_{t_1}|z_{t_2}, b_{t_1}, b_{t_1})}} \Big[\log p(x_{t_2}|z_{t_2}) \\ + \log p_B(z_{t_1}|b_{t_1}) + \log p(z_{t_2}|z_{t_1}) \\ - \log p_B(z_{t_2}|b_{t_2}) - \log q(z_{t_1}|z_{t_2}, b_{t_2}, b_{t_1}) \Big] \end{aligned} \tag{7}$$

While it seems that the quantity $\mathcal{L}_{t_1,t_2}$ refers to $-\log p(x_t|x_{\leq t})$, this is not completely clear.

For a schematic of this architecture, see Figure 2
The total loss is written by the authors as:

$$\begin{aligned} -\mathcal{L} &= \mathbb{E}_S \big[\log p(x_{t_{1:N}}) \big] \\ &\geq \mathbb{E}_S \Big[\sum_i \mathop{\mathbb{E}}_{\substack{z_{t_i} \sim p_B(z_{t_i}|b_{t_i}) \\ z_{t_{i-1}} \sim q(z_{t_{i-1}}|z_{t_i}, b_{t_{i-1}}, b_{t_{i-1}})}} \big[\log p(x_{t_i}|z_{t_i}) \\ &+ \log p_B(z_{t_{i-1}}|b_{t_{i-1}}) + \log p(z_{t_i}|z_{t_{i-1}}) \\ &- \log p_B(z_{t_i}|b_{t_i}) - \log q(z_{t_{i-1}}|z_{t_i}, b_{t_i}, b_{t_{i-1}}) \big] \Big] \end{aligned} \tag{8}$$

where S is the sampling scheme for pairs.

4. Discussion and Questions

1. **Possibly incorrect decomposition of term:** Earlier we decomposed the quantity $p(z_{t-1}, z_t|x_{<t})$ as $p_B(z_{t-1}|b_{t-1})$ and $p(z_{t-1}|z_t)$ following the chain rule decomposition and the markov assumption. The jumpy loss follows the same pattern by decomposing $p(z_{t_1}, z_{t_2}|x_{<t_2})$ as $p_B(z_{t_1}|b_{t_1})$ and $p(z_{t_2}|z_{t_1})$. But in the second term of the decomposition, we cannot remove $x_{<t_2}$ from the conditional!! If we are following the graphical model in Figure 1, the markov assumptions do not dictate that.

2. **The possible effects of including the smoothing factor:** The authors argue that the smoothing factor $q(z_{t_1}|z_{t_2}, b_{t_2}, b_{t_1})$ is of particular significance in learning correct transitions.

   They give the example of a closed box with either item A or B in it. At time t_1 the box is unopened, and at time t_2 the box has been opened and is known to have B in it. A good transition model of the box would postulate that its contents are unchanged. But sampling from just the belief at time t_1 wold give us item A 50% of the time, which would tell us that in our transition the content of the box changes from A to B. Therefore, it's imperative to sample from the smoothing factor for the content at a previous time-step.

3. **No sampling from the actual transition:** z_{t_1} is sampled from the smoothing posterior and z_{t_2} is sampled from the distribution of state given it's belief $p_B(z_{t_2}|b_{t_2})$. No parameters here are shared with the transition prior, which means the prior learns only from the KL term by the virtue of forcing the posterior to be closer to it. DVBF[1] has shown that this isn't enough to ensure that a good transition model is learned. One hypothesis about how this model works is that it relies on the smoothing factor - samples a particular future for z_{t_2} from $p_B(z_{t_2}|b_{t_2})$ and makes z_{t_1} stick to it .

4. Possible *non-iid* sampling when we use S, because the pairs may be related.

5. While the images in the result section look good, they only show that the transition learns to stay within a good manifold of latent space that can give good looking reconstructions, but it doesn't show that the actual learnt transitions are good too, no results in terms of application of controls

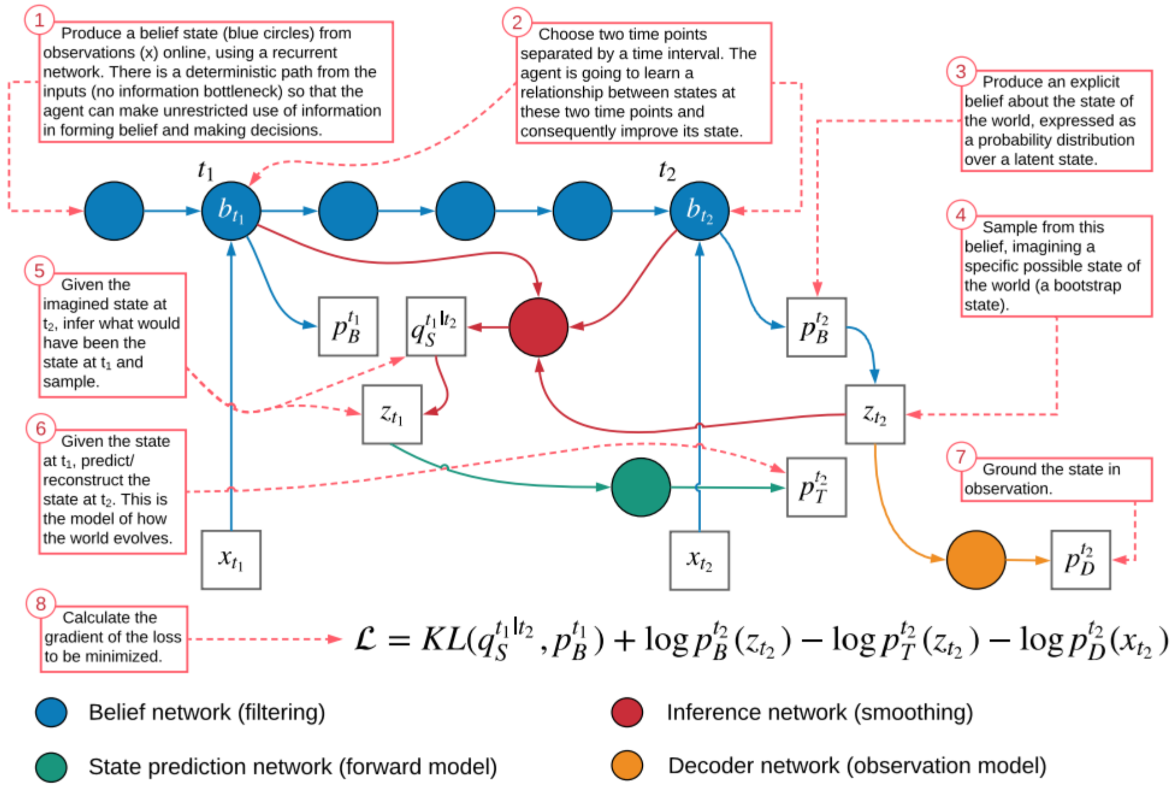

Figure 2. Schematic of TD-VAE's jumpy state model(This is from the original paper)

References

[1] M. Karl, M. Soelch, J. Bayer, and P. van der Smagt. Deep variational bayes filters: Unsupervised learning of state space models from raw data. *arXiv preprint arXiv:1605.06432*, 2016. 2