EE-613: Homework on Inference

Fall 2017

As we have seen in the course, learning the parameters of an HMM requires the inference of different marginals. The goal of the homework is to use the sum-product framework to derive such inference algorithm. Note that there exist alternative methods for inference.

Settings.

We define an HMM where the data variables are denoted by z_n, with n ranging from 1 to N, and the state variable is denoted by x_n. The HMM is defined by the distributions $p(x_1)$, $p(x_n|x_{n-1})$ for $n = 2 \dots N$, and $p(z_n|x_n)$ for $1 = 2 \dots N$. Furthermore, we will denote by $z_n^\star$ when a data variable is observed (i.e. which value is given, or $z_n = z_n^\star$), and by $x_{1:n}$ the set of variables $\{x_1, ..., x_n\}$.

Questions

1. **Factor graph**.
   We change the directed model to an undirected model, and obtain the factor graph shown in 1. Give the formula of the factors in function of the HMM local distributions, so that the undirected version represents the same distribution.

2. **Observation Messages**.
   We want to run the sum-product algorithm. Using the same definition than in the course for the local message passing scheme, define the following:
   - give the formula of $\mu_{z_n \to f_n^d}(z_n)$;
   - then derive $\mu_{f_n^d \to x_n}(x_n)$
   - assuming that z_n is observed and take the value $z_n^\star$, to what does $\mu_{f_n^d \to x_n}(x_n)$ reduce to? Note this is valid for all values of n.

   For the remaining, we will assume that $z_{1:N} = z_{1:N}^\star$

3. **Left-right state messages**.
   We derive the left-to-right recursion scheme.
   - give the definition of $\mu_{f_1 \to x_1}(x_1)$.
   - derive the recursion existing between $\mu_{f_n \to x_n}(x_n)$ and $\mu_{f_{n-1} \to x_{n-1}}(x_{n-1})$ (don't forget to take into the message from the data).

4. **Right-left state messages**.
   Similarly:

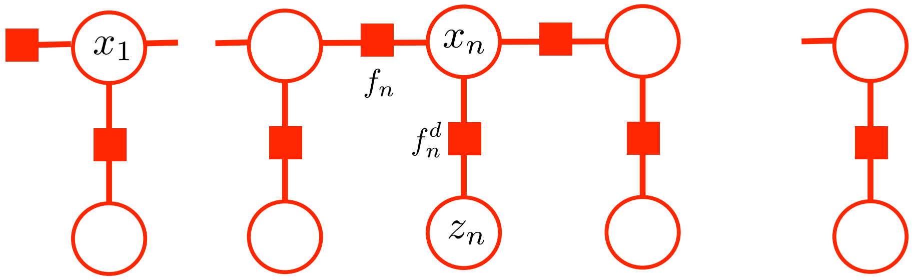

Figure 1: Factor graph of HMM.

- give the formula for $\mu_{x_N \to f_N}(x_N)$ and then $\mu_{f_N \to x_{N-1}}(x_{N-1})$.
- derive the recursion existing between $\mu_{f_{n+1} \to x_n}(x_n)$ and $\mu_{f_n \to x_{n-1}}(x_{n-1})$.

5. **Message passing algorithm**.
Given all the above initialization and recursive equations, write down an algorithm (or pseudo-code) that would compute all messages of interest in the graph. In particular, pay attention to the order in which the messages are computed (i.e. all messages on a right hand-side of an update/formula need to have been computed before it is used !).

6. **Marginals**.
Now we can derive the marginals of interest for the HMM:
   - determine the marginal distributions $p(x_1|z_{1:N} = z^{\star}_{1:N})$, $p(x_N|z_{1:N} = z^{\star}_{1:N})$ and $p(x_n|z_{1:N} = z^{\star}_{1:N})$ (n=2...$N-1$) from the messages.
   - show that:

$$p(x_{n-1}, x_n|z_{1:N} = z^{\star}_{1:N}) \propto \mu_{f_{n-1} \to x_{n-1}}(x_{n-1})\mu_{f_{n+1} \to x_n}(x_n) f_n(x_{n-1}, x_n) f^d_{n-1}(x_{n-1}, z^{\star}_{n-1}) f^d_n(x_n, z^{\star}_n)$$

   Hint: write down the requested term by marginalization (and as a reminder, note that $p(x_n, x_{n-1}|z_{1:N} = z^{\star}_{1:N}) \propto p(x_n, x_{n-1}, z_{1:N} = z^{\star}_{1:N})$).

Sampling Methods

Show that the Gibbs sampling proposal leads to an acceptance ratio of 1 in the Metropolis-Hasting algorithm.