Homework 4 part 2

Instructions: Please put all answers in a single PDF with your name and NetID and upload to SAKAI before class on the due date (there is a LaTeX template on the course web site for you to use). Definitely consider working in a group; please include the names of the people in your group and write up your solutions separately. If you look at any references (even wikipedia), cite them. If you happen to track the number of hours you spent on the homework, it would be great if you could put that at the top of your homework to give us an indication of how difficult it was.

Problem 4

Revisiting the forward-backward algorithm for HMMs

In class, when we discussed the forward-backward algorithm for the HMM, we focused on the second moment statistics (expected sufficient statistics) $\mathbb{E}[z_t^j z_{t+1}^k]$. Write out the first moment statistics $\mathbb{E}[z_t^k]$ in terms of $\alpha_t(k)$ and $\beta_t(k)$ using the techniques we've seen in class for this; you may want to use Bayes rule, chain rule, or conditional independence, among others. Recall that:

$$\begin{aligned}
\alpha_t(k) &= p(z_t^k | x_{1,\cdots,t}) \\
\beta_t(k) &= p(x_{t+1,\cdots,T} | z_t^k).
\end{aligned}$$

Problem 5

Variable elimination for HMMs

The graph and associated conditional independencies for an HMM can be thought of as a simple tree-structured graphical model. In this problem, you will derive an algorithm for exact inference using variable elimination to obtain an algorithm equivalent to Forward-Backward algorithm.

(a) What is the tree width for the HMM as shown in the figure in Problem 2 of HW4?

(b) Apply variable elimination to derive the posterior probability of a single state: $p(z_t | x_{1:T})$ (this is called *smoothing*). This should be a $K \times T$ matrix, where K is the number of states and T is the number of observations. Hint: think of the optimal ordering of variables, inspired by the forward backward algorithm.

Problem 6

Viterbi algorithm for HMMs

We are often interested in the most probable state $\hat{z}_{1:T} = arg \max_Z p(Z_{1:T} | X_{1:T}, \hat{\theta})$ for an HMM with estimated parameters $\hat{\theta}$. This is also known as the Viterbi algorithm. How would you modify the algorithm you derived in question 5.b? Explain using a few sentences with key equations.