CS 228 Problem Set 1

Hugh Zhang

February 19, 2017

Problem 1

Let the structure be a two node binary Bayes net with one edge from A to B. Let $P(A = 1) = 0.51, P(B = 1 \mid A = 1) = .5, P(B = 1 \mid A = 0) = .99$

Then, $P(B = 1) = 0.51 * .5 + 0.49 * .99 = .7401$, so the most likely outcome based on marginals without looking at conditionals is $A = 1, B = 1$. However, the $P(A = 1, B = 1) = .51 * .5 = .255$. This is not as likely as the most likely outcome, A=0,B=1 which has $P(A = 0, B = 1) = .49 * .99 = .4851$.

Problem 2

Basic message passing is nd^2 for a chain. Labeling the nodes from A to Y WLOG, you can factor

$$P(A \dots Y) = \phi(A, B) * \phi(B, C) \dots \phi(X, Y)$$

Z, which is the normalizing constant is just

$$\sum_A \cdots \sum_Y P(A \dots Y) = \sum_A \sum_B \phi(A, B) * \sum_C \phi(B, C) \cdots \sum_Y \phi(X, Y)$$

Z takes nd^2 time to calculate because this is just message passing across the entire array. Note that when we try calculating the marginals for a pair of variables, the variables cut out three parts of the chain. WLOG call them x_i and x_j with $i < j$. For all x_k with $k < i$ or x_m with $m > j$, then you can do basic message propogation to calculate each chain for nd^2 using basic message passing. Then, to get the middle part, you can choose to extend the left side (WLOG) and calculate the middle chain for each fixed value of x_i, and then unite it with the right chain, then normalize to have a complete probability. Thus, since you do message passing d times, it takes nd^3.

Let r(x) be the right chain and l(x) be the left chain.

Then

$$P(x) = \frac{l(x)\phi(X_i, x_{i+1}) \dots \phi(x_{j-1}, X_j) * r(x)}{l(x) \sum_{x_i} \phi(x_i, x_{i+1}) \cdots \sum_{x_j} \phi(x_j - 1, x_j) * r(x)}$$

$$= \frac{\phi(X_i, x_{i+1}) \dots \phi(x_{j-1}, X_j)}{\sum_{x_i} \phi(x_i, x_{i+1}) \cdots \sum_{x_j} \phi(x_j - 1, x_j)}$$

2.2

Note that X_i and X_j are independent conditioned on X_{j-1}

$$\sum_{x_{j-1}} P(X_i, X_{j-1}) P(X_j \mid X_{j-1}) = \sum_{x_{j-1}} P(X_i, X_{j-1}, X_j) = P(X_i, X_j)$$

Problem 3

If you change a factor (by changing its value, or by adding an edge) in a clique, check all its neighbors and see if they are affected, AKA the seperation set contains the factor that has been changed. If they are affected, then everything in that entire subtree needs to be updated, since they all depend on that factor in the message passing.

3.2

If you only want the marginal over a single variable, you only need to pass the message along the single path to the variable you want to calculate. In addition, you don't need to update the entire tree, just for every node that you are d-seperated on since you marginalize the probability out otherwise.

EQUATION FOR MESSAGE DOESN"T DEPEND ON THE OLD GUY. Calibrated only if all the messages that are coming in are ok.

Problem 4

4.2.1

Inference is exponential. No

Yes, forward prop