Inference and Statistics: Homework 2

eh1885

September 2019

1 Problem 1

1.1 a

We list the independence relations below, **without repeating independence relations we have mentioned previously**.

$$1 \perp \{2, 3, 5, 7, 8, 9, 10\} \tag{1}$$
$$2 \perp \{7, 8\} \tag{2}$$
$$3 \perp \{7, 8\} \tag{3}$$
$$4 \perp \{8\} \tag{4}$$
$$5 \perp \{6\} \tag{5}$$
$$6 \perp \{7, 8\} \tag{6}$$
$$7 \perp \{8, 10\} \tag{7}$$
$$8 \perp \{10\} \tag{8}$$

1.2 b

Given $\{X_2, X_9\}$, the independence relations for X_1 do not actually change. Without knowing any other node values, 1 is dependent only on 6 and 4. Knowing 2 changes the distribution of 6 and 4, but this change does not interact with the relation between 1 and 6 or 4. Likewise, knowing 9 reveals information on 7, but the contribution of 7 to 4 is independent of the contribution of 1 to 4, so the independence relations between 1 and $\{4, 6\}$ persist.

2 Problem 2

We can see that $P(X = x) = P(X = x|Y) = P(X = x|Z)$ for an arbitrary labeling X, Y, Z of the three i.i.d. random variables.

The only dependence graph which has $X \perp Y$, $X \perp Z$, and $Y \perp Z$ is the trivial graph with no edges in this graph.

However, this graph also does not capture the dependencies between X, Y, Z, since $P(X = 1|Y = 1, Z = 1) = 2/3, P(X = 1|Y = 1, Z = 0) = 1/3$, and $P(X = 1) = 1/2$. Thus X, Y, and Z are not actually independent as the trivial graph would suggest.

Thus there is no directed acyclic graph G whose dependence relations capture this probability distribution.

3   Problem 3

Note that $\{-1, 1\}$ can be mapped to $\{0, 1\}$ using the function $f(x) = \dfrac{x' + 1}{2}$. If we perform this change of variables, we get

$$\frac{1}{Z} \exp \Big(\sum_{i,j} w_{i,j}(1/2x_i' + 1/2)(1/2x_j' + 1/2) - \sum_i \mu_i(1/2x_i' + 1) \Big) \tag{9}$$

$$= \frac{1}{Z} \exp \Big(\sum_{i,j} 1/2w_{i,j}(1/2x_ix_j - x_i - x_j + 1) - \sum_i (1/2\mu_ix_i + \mu_i) \Big) \tag{10}$$

$$= \frac{1}{Z} \exp \Big(\sum_{i,j}(1/4w_{i,j}x_ix_j) + \sum_i (1/2\mu_i - \sum_j w_{i,j})x_i) + \sum_{i,j} w_{i,j} + \sum_i \mu_i \tag{11}$$

$$= \frac{1}{Z} \exp \Big(\sum_{i,j}(1/4w_{i,j}x_ix_j) + \sum_i (1/2\mu_i - \sum_j w_{i,j})x_i) + C \tag{12}$$

$$= \frac{1}{Z_2} \exp \Big(\sum_{i,j}(1/4w_{i,j}x_ix_j) + \sum_i (1/2\mu_i - \sum_j w_{i,j})x_i) \tag{13}$$

where the last step squished the constant into Z to form Z_2. We can see that the parameters w' become $4w$ and μ' become $2\mu - w \sum_{i,j} w_{i,j}$ respectively.

4   Problem 4

In the case of trees, the undirected graphical model and the directed graphical models imply the exact same dependence relations, since the Markov blankets for each node are the same in the undirected and directed cases. Since trees guarantee a node has a single parent and that there are no cycles, the Markov blanket is the node's single parent, and any potential children of that node.

We proceed using an inductive proof using the factorization implied by directed graphical models, noting that we can back out a directed graphical model by arbitrarily picking a node in the undirected graphical model and using the strategy in the hint given.

For the base case where the tree is comprised of two nodes with one edge between them, we can arbitrarily label one x_i and the other x_j. WLOG we pick

x_i to be the base, making it the parent of x_j. Then the directed model factorizes to be $p(x_i)p(x_j|x_i) = p(x_i)\dfrac{p(x_i, x_j)}{p(x_i)} = \dfrac{p(x_i, x_j)}{p(x_i)p(x_j)}p(x_i)p(x_j)$ as required.

Now for the inductive case, suppose we already have such a factorization. We can add any node x_j to our tree so long as the node only has one parent x_i, and the parent x_i is already a member of our tree. Then we have additional terms in the factorization $\dfrac{p(x_i, x_j)}{p(x_i)p(x_j)}$ from the new edge between x_i, x_j and an additional term $p(x_j)$ from the new node x_j. Multiplying these together, we get $\dfrac{p(x_i, x_j)}{p(x_i)p(x_j)}p(x_j) = \dfrac{p(x_i, x_j)}{p(x_i)} = p(x_j|x_i)$.

But this is just the additional dependence relation implied by adding this node to the directed graphical model. Again, since the directed graphical model's independence relations are the same as its undirected counterpart, we can see that these two factorizations are indeed the same.