

Probabilistic Graphical Models, Spring 2007

Homework 1

Due at the beginning of class on 10/08/07

Instructions

There are six questions in this homework. The last question involves programming which should be done in MATLAB. Do *not* attach your code to the writeup. Make a tarball of your code and output, name it `<userid>.tgz` where your userid is your CS or Andrew id, and copy it to `/afs/cs/academic/class/10708-f07/hw1/<userid>`.

If you are not submitting this from an SCS machine, you might need to authenticate yourself first. See http://www.cs.cmu.edu/~help/afs/cross_realm.html for instructions. If you are not a CMU student and don't have a CS or Andrew id, email your submission to `10708-07-instr@cs.cmu.edu`.

You are allowed to use any of (and only) the material distributed in class for the homeworks. This includes the slides and the handouts given in the class¹. Refer to the web page for policies regarding collaboration, due dates, and extensions.

1 [6 pts] Representation

Consider $N+1$ binary random variables $X_1 \dots X_N, Y$ that have the following conditional independencies: $\forall i, j, X_i \perp X_j | Y$

1. Suppose you wish to store the joint probability distribution of these $N+1$ variables as a single table. How many parameters will you need, to represent this table?
2. Draw a graphical model over these $N+1$ variables that encodes the conditional independencies above. How many parameters will you need to completely describe the distribution using this representation?

2 [10 pts] Gaussians

1. Show that for any joint density function $P(X, Y)$, if we have $X \perp Y$, then $Cov(X, Y) = 0$.
2. Show that if $P(X, Y)$ is a Gaussian distribution and $Cov(X, Y) = 0$, then $X \perp Y$.

3 [15 pts] Undirected Graphical Models

A BN \mathcal{G} is said to preserve the dependencies of a Markov Network \mathcal{H} if whenever \mathcal{G} implies $X \perp Y | Z$, \mathcal{H} also implies it. Provide an example of a Markov network, which, when transformed to a dependency-preserving Bayesian network, grows exponentially, ie the size of the largest clique in \mathcal{H} is constant, yet any BN I-map is exponentially large² in the number of nodes n .

¹Please contact Monica Hopes(meh@cs) if you need a copy of a handout.

²It will therefore suffice to show an example \mathcal{H} , where any BN that is an I-map of \mathcal{H} has atleast c^{n^k} parameters for any real constants $c > 1, k > 0$.

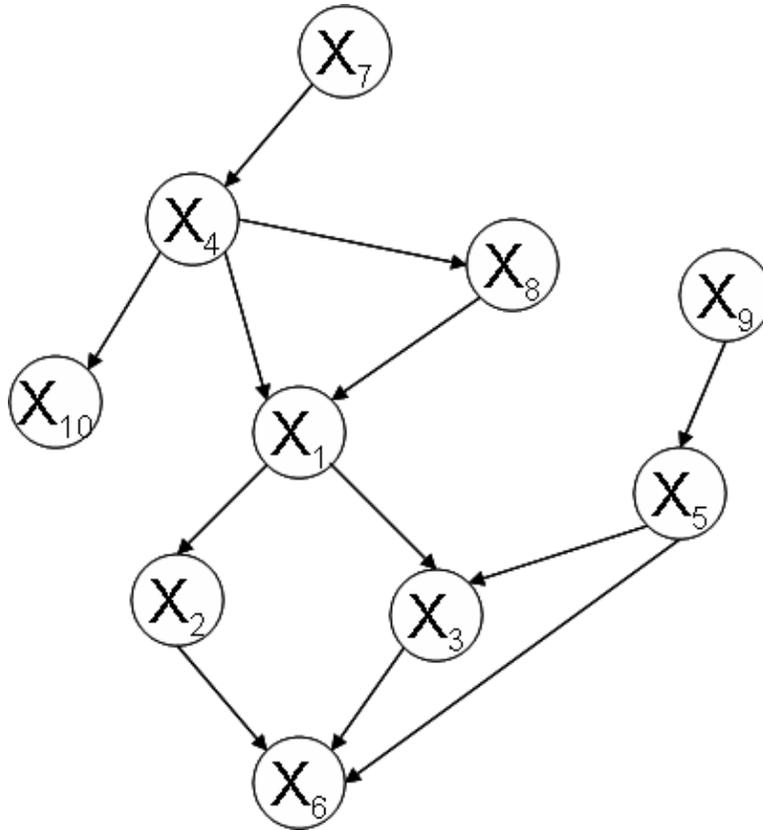


Figure 1: Bayesian network for Problem 4

4 [15 pts] Elimination

Consider the Bayesian Network given in Fig 1.

1. For each of the following assertions of (conditional) independence, state if they are True or False with justification.
 - (a) $X_7 \perp X_9$
 - (b) $X_{10} \perp X_7 | X_4$
 - (c) $X_4 \perp X_8 | X_1$
 - (d) $X_4 \perp X_3 | X_1$
 - (e) $X_1 \perp X_5 | X_3$
 - (f) $X_1 \perp X_5 | X_6$
 - (g) $X_9 \perp X_7 | X_6$
2. For the elimination order $X_7, X_{10}, X_4, X_9, X_5, X_8, X_1, X_2, X_3$, draw the moralized graph and the fill-in edges created in the process of elimination.
3. For the elimination order $X_6, X_9, X_4, X_8, X_7, X_{10}, X_1, X_2, X_5$, draw the resulting junction tree.

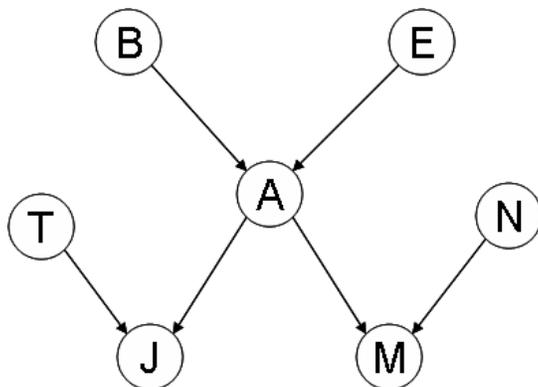


Figure 2: Bayesian network for Problem 5

5 [20 pts] Arc Reversal

1. Consider the Bayesian Network in Fig 2. Suppose we wish to reverse arc $B \rightarrow A$. What additional modifications to the structure of the network are needed to ensure that the new network is an I-map of the original distribution? Your network should not reverse any additional edges, and should only differ minimally from the original in terms of the number of edge additions or deletions. Justify your response.
2. Now, consider a Bayesian Network \mathcal{G} . Assume that $X \rightarrow Y$ is the only directed trail from X to Y . Define a general procedure to reverse the arc $X \rightarrow Y$, i.e., for constructing a graph \mathcal{G}' that is an I-map for the original distribution but that contains an arc $Y \rightarrow X$ and otherwise differs minimally from \mathcal{G} in the same sense as above. Once again, you must provide a justification for your response.
3. Suppose that we use the above procedure to transform \mathcal{G} into \mathcal{G}' and then, reverse the arc back to its original direction by repeating the above transformation to obtain graph \mathcal{G}'' . Is $\mathcal{G} = \mathcal{G}''$ always?

6 [30 pts] Belief Propagation

Cecilia and Slackernerney, two graduate students in Markovford, have gotten into an argument over the weather. Cecilia thinks summer has ended and that we are into Autumn(Fall), while Slackernerney (who has been spending all his time in a room without windows) thinks that this is summer. Since they weren't able to resolve it themselves, they have decided to ask the fine students at CMU to help them decide.

A bit about Markovford: there are three seasons in Markovford – Summer(S), Autumn(A) and Winter(W). Given the season the previous day, the season on a day is conditionally independent of the season on all previous days. The weather is either Hot(H), Rainy(R) or Freezing(F). Given the season on any given day, the weather that day is independent of all other variables.

More formally, if we let C^i denote the season on the i^{th} day (taking values S,A,W) and O^i denote the observed weather pattern(one of H,R,F); we have $\forall j < i - 1 C^i \perp C^j | C^{i-1}$ and $\forall X, O^i \perp X | C^i$ where X is any random variable other than C^i, O^i .

1. Draw a graphical model over $C^1 \dots C^N, O^1 \dots O^N$ that satisfies the conditional independencies listed above.
2. Implement sum-product and max-product algorithms in MATLAB, for this graphical model. In order to fully specify the graphical model, you will need to fill the values for the CPT. Some of these are

listed in `params.txt`. Assume a uniform distribution for any additional CPTs that you think you need to fully specify your graphical model (Hint: You will need one additional CPT).

3. Cecilia has made 20 observations of the weather over the last few months (ie $O^1 \dots O^N$). This data is in `weather.txt`.
 - (a) Apply inference, to compute the most likely season for each of these days and submit the output in a textfile called `hw1-a.txt`
 - (b) Apply inference to determine the most likely sequence of $C^1 \dots C^N$ that generated this observed sequence. Submit your output in a textfile called `hw1-b.txt`
 - (c) Cecilia made the last observation today. Who do you think was correct? What is the probability that it is currently winter in Markovford?

Submit your code and a `readme` describing briefly what each file does, along with the output listed above.