CS 510/610 Topics on probabilistic graphical models

Homework 1

Portland State U, Winter 2021 — *01/16/21*

Lecturer: Fang Song — *Due: 01/28/21*

Instructions. This problem set contains 6 pages (including this cover page) and 3 questions. A random subset of problems will be graded.

- Your solutions will be graded on *correctness* and *clarity*. You should only submit work that you believe to be correct, and you will get significantly more partial credit if you clearly identify the gap(s) in your solution. It is good practice to start any long solution with an informal (but accurate) summary that describes the main idea. You may opt for the "I take 15%" option.
- You need to submit a PDF file before the deadline. Either a clear scan of you handwriting or a typeset document is accepted. You will get 5 bonus points for typing in LaTeX (Download and use the accompany TeX file).
- You may collaborate with others on this problem set. However, you must **write up your own solutions** and **list your collaborators and any external sources** for each problem. Be ready to explain your solutions orally to a course staff if asked.

1. (Probability)

   (a) (5 points) Let A, B, C be random variables, and suppose that the joint distribution is positive. Prove or disprove the following statement. If $A \perp B|C$ and $A \perp C|B$, then $A \perp B$ and $A \perp C$.

   (b) (10 points) Let X be a set of random variables with joint distribution p. Let A, B, C be three disjoint subsets of variables such that $X = A \cup B \cup C$. Prove that $A \perp B|C$ **iff.** we can write p in the form $p(X) = \phi_1(A, C)\phi_2(B, C)$ for some non-negative functions ϕ_1 and ϕ_2.

   (c) (10 points) Let X, Y, Z be random variables. Show that $p(X|Y) = \sum_z p(X, z|Y)$. (Hint: chain rule)

2. (10 points) (Detecting cycle in a graph) Describe an algorithm to determine if a given directed graph G contains a cycle. Show correctness and running time of your algorithm.

3. (Bayesian networks) For each of the following statements, state True or False, and briefly justify your answers.

   (a) (6 points) Let $\mathcal{G}$ be the BN shown in Fig. 1.

$$\begin{tikzcd}
A \arrow[r] & B \arrow[r] & C \\
D \arrow[ur] \arrow[r] & E &
\end{tikzcd}$$

Figure 1: A Bayesian network

   (1) $E \perp C|B$.
   (2) $A \perp E|C$.

(b) (9 points) Read the diagram in Fig. 2. Recall the definitions of local and global independencies of a BN $\mathcal{G}$ and independencies of a distribution p.

$$\begin{aligned}
I_\ell(\mathcal{G}) =& \{X \perp \text{NonDescendants}(X) | \text{Pa}(X)\} \\
I(\mathcal{G}) =& \{X \perp Y | Z : \text{d-Sep}(X, Y | Z)\} \\
I(p) =& \{X \perp Y | Z : p(X, Y | Z) = p(X|Z)p(Y|Z)\}
\end{aligned}$$

$$\begin{tikzcd}
p \text{ factorizes over } \mathcal{G} \arrow[r, Rightarrow, "a"] & I(\mathcal{G}) \subseteq I(p) \arrow[r, Rightarrow, "b"] & I_\ell(\mathcal{G}) \subseteq I(p) \arrow[ll, Rightarrow, bend left, "c"]
\end{tikzcd}$$

Figure 2: Some relations in Bayesian networks.

(1) Relation a is true.
(2) Relation b is true.
(3) Relation c is true.

(c) (5 points) If $\mathcal{G}_1$ is an I-map of distribution p, and $\mathcal{G}_1$ has fewer edges than $\mathcal{G}_2$, then $\mathcal{G}_2$ is not a minimal I-map of p.