Homework 1

CMU 10-708: Probabilistic Graphical Models (Fall 2020)

September 15, 2020

Instructions:

- **Collaboration policy:** Collaboration on solving the homework is allowed, after you have thought about the problems on your own. It is also OK to get clarification (but not solutions) from books or online resources, again after you have thought about the problems on your own. There are two requirements: first, cite your collaborators fully and completely (e.g., "Bob explained to me what is asked in Question 4.3"). Second, write your solution *independently*: close the book and all of your notes, and send collaborators out of the room, so that the solution comes from you only.
- **Submitting your work:** Assignments should be submitted as PDFs using Gradescope unless explicitly stated otherwise. Each derivation/proof should be completed on a separate page. Submissions can be handwritten, but should be labeled and clearly legible. Else, submission can be written in LaTeX. Upon submission, label each question using the tempate provided by Gradescope.

1 Markov Properties for UGMs [20 points] (Chih-Kuan)

1. Prove the following properties, by using the following equivalence:

$$X \perp Y \mid Z \equiv P(X, Y, Z) = \phi_1(X, Z)\, \phi_2(Y, Z),$$

for some factors $\phi_1(\cdot), \phi_2(\cdot)$.

   - **[2 points]** If $A \perp (B, D) \mid C$ then $A \perp B \mid C$.
   - **[3 points]** If $A \perp (B, D) \mid C$ then $A \perp B \mid (C, D)$ and $A \perp D \mid (B, C)$.
   - **[5 points]** For strictly positive distributions, if $A \perp B \mid (C, D)$ and $A \perp C \mid (B, D)$ then $A \perp (B, C) \mid D$.

2. Let $I_p(G), I_\ell(G), I(G)$ denote the set of pairwise, local, and global Markov properties entailed by an undirected graph G. Let $I(P)$ denote the set of all conditional independences satisfied by some distribution P.

   - **[2 points]** Show that $I_p(G) \subseteq I_\ell(G) \subseteq I(G)$.
   - **[3 points]** Use the above to show that: {distributions P that satisfy global Markov prop.} $\subseteq$ {distributions P that satisfy local Markov prop.} $\subseteq$ {distributions P that satisfy pairwise Markov prop.}.
   - **[5 points]** In class we showed that any distribution that factors according to G satisfies the global Markov properties associated with G. Suppose we know that all positive distributions P that satisfy pairwise Markov properties associated with G also factor according to the graph G.
     Show that this entails that: {positive distributions P that satisfy pairwise Markov prop.} $\subseteq$ {positive distributions P that satisfy local Markov prop.} $\subseteq$ {positive distributions P that satisfy global Markov prop.}.

2 Independence and Correlation on UGMs [30 points] (Chang)

Let $\mathbf{X} = (X_1, \dots, X_d)$ be a random vector (not necessarily Gaussian) with mean $\boldsymbol{\mu}$ and covariance matrix Σ.

The partial correlation matrix R of $\mathbf{X}$ is a $d \times d$ matrix where each entry $R_{ij} = \rho(X_i, X_j | \mathbf{X}_{-ij})$ is the partial correlation between X_i and X_j given the $d-2$ remaining variables $\mathbf{X}_{-ij}$. Let $\Theta = \Sigma^{-1}$ be the inverse covariance matrix of $\mathbf{X}$.

We will prove the relation between R and Θ, and furthermore how Θ characterizes conditional independence in Gaussian graphical models.

1. **[10 points]** Show that

$$\begin{pmatrix} \Theta_{ii} & \Theta_{ij} \\ \Theta_{ji} & \Theta_{jj} \end{pmatrix} = \begin{pmatrix} \mathrm{Var}[e_i] & \mathrm{Cov}[e_i, e_j] \\ \mathrm{Cov}[e_i, e_j] & \mathrm{Var}[e_j] \end{pmatrix}^{-1} \tag{1}$$

   for any $i, j \in [d]$, $i \neq j$. Here e_i is the residual resulting from the linear regression of $\mathbf{X}_{-ij}$ to X_i, and similarly e_j is the residual resulting from the linear regression of $\mathbf{X}_{-ij}$ to X_j.

   (Hint: Use the formula for block matrix inversion.)

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & * \\ * & * \end{pmatrix}$$

2. **[10 points]** Show that

$$R_{ij} = -\frac{\Theta_{ij}}{\sqrt{\Theta_{ii}}\sqrt{\Theta_{jj}}} \tag{2}$$

3. **[10 points]** From the above result and the relation between independence and correlation, we know $\Theta_{ij} = 0 \iff R_{ij} = 0 \impliedby X_i \perp X_j \mid \mathbf{X}_{-ij}$. Note the last implication only holds in one direction.

   Now suppose $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$ is jointly Gaussian. Show that $R_{ij} = 0 \implies X_i \perp X_j \mid \mathbf{X}_{-ij}$.

   (Hint: You can directly use the fact that the conditional distribution $p(X_i, X_j \mid \mathbf{X}_{-ij})$ is Gaussian when $\mathbf{X}$ is jointly Gaussian.)

3 DGMs, d-Separation and Moralization [30 points] (Chirag)

Figure 1 shows three different DAGs over the random variables X, Y, Z, T, A.

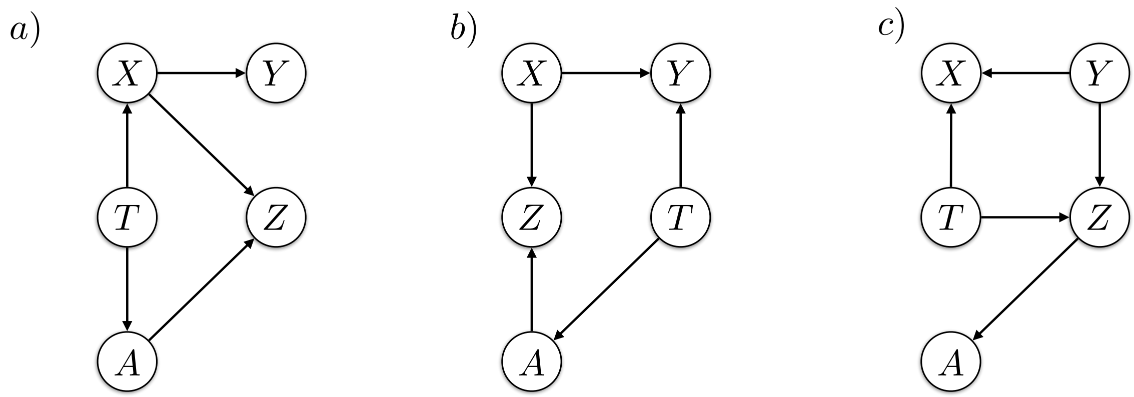

Figure 1: Directed Graphical Models

1. [**6 points**] For each of the DAGs in Figure 1, show how the joint distributions of the corresponding DGMs factorize.
2. For each of the DAGs in Figure 1, using rules of d-Separation to answer the following questions.

   (For the cases which are not independent, explicitly show at least one trail that is not blocked.)

   (i) [**3 points**] Is $Y \perp Z$?

   (ii) [**3 points**] Is $Y \perp T$?

   (iii) [**3 points**] Is $Y \perp A$?

   (iv) [**3 points**] Is $Y \perp Z \mid X$?

   (v) [**3 points**] Is $Y \perp T \mid A$?

   (vi) [**3 points**] Is $Y \perp A \mid T$?

3. Consider now the consequences of moralizing the DAGs in Figure 1.

   (i) [**3 points**] Show the moralized undirected graphs corresponding to each DAG in Fig. 1.

   (ii) [**3 points**] For each moralized graph, show how the joint distributions of the corresponding UGMs factorize over the corresponding maximal cliques.

***