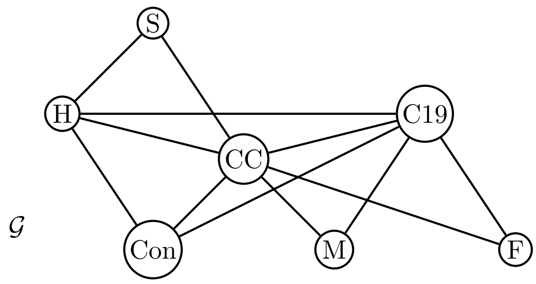
Probabilistic Graphical Models: Problem Set 1

Svante Linusson, Liam Solus
KTH Royal Institute of Technology

14 September 2023

1. (Medical diagnosis.) Using a graph can help us reduce the "size" of a representation of a distribution (i.e., the number of parameters). Consider a distribution $\mathbf{X} = [Con, M, F, H, CC, C19, S]^T$ for a medical diagnosis scenario, where S represents the season (4 states), M represents "muscle pain" (2 states), F represents "fever" (2 states), Con represents "congestion" (two states). These are symptoms that maybe be exhibited when a patient has one of the following diseases CC (a common cold), $C19$ (Covid-19) or H (hayfever). We treat CC, $C19$ and H as binary variables. Suppose we represent our model with the UG:



   (a) Suppose our distribution $\mathbf{X}$ satisfies the global Markov property with respect to $\mathcal{G}$. Do we need to make any assumptions about P in order to ensure that $\mathbf{X}$ factorizes according to $\mathcal{G}$?

   (b) How many parameters are needed to represent our distribution $\mathbf{X}$ if we assume it factorizes according to $\mathcal{G}$?

   (c) If we assume $\mathbf{X}$ satisfies the global Markov property with respect to $\mathcal{G}$, what CI relations $A \perp\!\!\!\perp B \mid C$ do we know hold in $\mathbf{X}$ where $A = \{S\}$ and $B = \{M\}$ are singletons?

   (d) Does this model seem reasonable to you? For instance, do the CI relations you detected in the previous part make intuitive sense? What about the dependencies represented by the edges in $\mathcal{G}$?

2. Suppose $\mathbf{X} = [X_1, \dots, X_m^T$ is a distribution satisfying the intersection axiom and suppose that we are given an undirected graph $\mathcal{G} = ([m], E)$. What is the quickest way to check that $\mathbf{X}$ satisfies the global Markov property with respect to $\mathcal{G}$? That is, what is the fewest number of CI relations that we need to check hold in $\mathbf{X}$ to verify that $\mathbf{X}$ is Markov to $\mathcal{G}$?

3. Let $\mathbf{X} = [X_1, \dots, X_m]^T$ have joint distribution $\mathbf{X}$ that is multivariate normally distributed as $\mathcal{N}(0, \Sigma)$. Suppose also that $\mathbf{X}$ is satisfies the global Markov property with respect to an undirected graph $\mathcal{G} = ([m], E)$. The **concentration matrix** of $\mathbf{X}$ is defined to be $K := \Sigma^{-1}$. For $i, j \in [m]$, show that the entry $K_{i,j}$ of K equals zero if and only if $X_i \perp\!\!\!\perp X_j \mid X_{[p]\setminus\{i,j\}}$. (Hint: How would Cramer take the inverse of a matrix?)

4. Suppose $[X_1, \dots, X_m]^T$ has probability density function $f(\mathbf{x})$, and let $A, B, C \subset [m]$ be disjoint subsets with $A, B \neq \emptyset$. Show that $X_A \perp\!\!\!\perp X_B \mid X_C$ if and only if there exist functions h, k such that

$$f(\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C) = h(\mathbf{x}_A, \mathbf{x}_C)k(\mathbf{x}_B, \mathbf{x}_C).$$

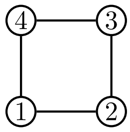
5. In this problem, our goal is to observe that if a distribution is nonpositive then it can obey the global Markov property for a nonchordal UG $\mathcal{G}$ and, at the same time, fail to factorize according

to $\mathcal{G}$. In other words, we will see that the hypotheses of Theorem 5 or Proposition 3 from class are necessary.

Consider the distribution $\mathbf{X} = [X_1, X_2, X_3, X_4]^T$, for X_i having outcomes $\{0, 1\}$ for all $i = 1, \dots, 4$, that assigns probability $1/8$ to each of the following outcomes and probability zero to all other outcomes:

$$\begin{array}{cccc} (0,0,0,0) & (1,0,0,0) & (1,1,0,0) & (1,1,1,0) \\ (0,0,0,1) & (0,0,1,1) & (0,1,1,1) & (1,1,1,1) \end{array}$$

Let $\mathcal{G}$ denote the UG



(a) Show that $\mathbf{X}$ is satisfies the global Markov property with respect to $\mathcal{G}$.

(b) Show that $\mathbf{X}$ does not factorize according to $\mathcal{G}$. (Hint: Use proof by contradiction)