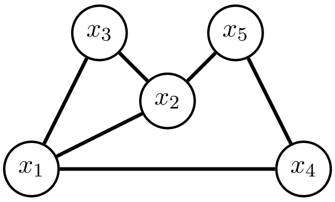
Exercise 1. ***Visualising and analysing Gibbs distributions via undirected graphs***

We here consider the Gibbs distribution

$$p(x_1, \ldots, x_5) \propto \phi_{12}(x_1, x_2)\phi_{13}(x_1, x_3)\phi_{14}(x_1, x_4)\phi_{23}(x_2, x_3)\phi_{25}(x_2, x_5)\phi_{45}(x_4, x_5)$$

(a) *Visualise it as an undirected graph.*

Solution. We draw a node for each random variable x_i. There is an edge between two nodes if the corresponding variables co-occur in a factor.



(b) *What are the neighbours of x_3 in the graph?*

Solution. The neighbours are all the nodes for which there is a single connecting edge. Thus: $\text{ne}(x_3) = \{x_1, x_2\}$. (Note that sometimes, we may denote $\text{ne}(x_3)$ by ne_3.)

(c) *Do we have $x_3 \perp\!\!\!\perp x_4 \mid x_1, x_2$?*

Solution. Yes. The conditioning set $\{x_1, x_2\}$ equals ne_3, which is also the Markov blanket of x_3. This means that x_3 is conditionally independent of all the other variables given $\{x_1, x_2\}$, i.e. $x_3 \perp\!\!\!\perp x_4, x_5 \mid x_1, x_2$, which implies that $x_3 \perp\!\!\!\perp x_4 \mid x_1, x_2$. (One can also use graph separation to answer the question.)

(d) *What is the Markov blanket of x_4?*

Solution. The Markov blanket of a node in a undirected graphical model equals the set of its neighbours: $\text{MB}(x_4) = \text{ne}(x_4) = \text{ne}_4 = \{x_1, x_5\}$. This implies, for example, that $x_4 \perp\!\!\!\perp x_2, x_3 \mid x_1, x_5$.

(e) *On which minimal set of variables A do we need to condition to have $x_1 \perp\!\!\!\perp x_5 \mid A$?*

Solution. We first identify all trails from x_1 to x_5. There are three such trails: (x_1, x_2, x_5), (x_1, x_3, x_2, x_5), and (x_1, x_4, x_5). Conditioning on x_2 blocks the first two trails, conditioning on x_4 blocks the last. We thus have: $x_1 \perp\!\!\!\perp x_5 \mid x_2, x_4$, so that $A = \{x_2, x_4\}$.

Exercise 2. ***Factorisation and independencies for undirected graphical models***

We here consider the graph in Figure 1.

(a) *What is the set of Gibbs distributions that are induced by the graph?*

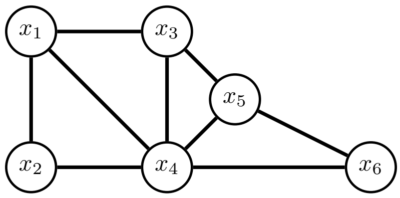


Figure 1: Graph for Exercise 2

Solution. The graph in Figure 1 has four maximal cliques:

$$(x_1, x_2, x_4) \quad (x_1, x_3, x_4) \quad (x_3, x_4, x_5) \quad (x_4, x_5, x_6)$$

The Gibbs distributions are thus

$$p(x_1, \ldots, x_6) \propto \phi_1(x_1, x_2, x_4)\phi_2(x_1, x_3, x_4)\phi_3(x_3, x_4, x_5)\phi_4(x_4, x_5, x_6)$$

(b) *Let p be a pdf that factorises according to the graph. Can we expect that $p(x_3|x_2, x_4) = p(x_3|x_4)$?*

Solution. $p(x_3|x_2, x_4) = p(x_3|x_4)$ means that $x_3 \perp\!\!\!\perp x_2 \mid x_4$. We can use the graph to check whether this generally holds for pdfs that factorise according to the graph. There are multiple trails from x_3 to x_2, including the trail (x_3, x_1, x_2), which is not blocked by x_4. From the graph, we thus cannot conclude that $x_3 \perp\!\!\!\perp x_2 \mid x_4$, and $p(x_3|x_2, x_4) = p(x_3|x_2)$ will generally not hold (the relation may hold for some carefully defined factors ϕ_i).

(c) *Explain why $x_2 \perp\!\!\!\perp x_5 \mid x_1, x_3, x_4, x_6$ holds.*

Solution. The distribution that factorises according to the graph satisfies the pairwise Markov property. Since x_2 and x_5 are not neighbours, and x_1, x_3, x_4, x_6 are the remaining nodes in the graph, the independence relation follows from the pairwise Markov property.

(d) *Assume you would like to approximate $\mathbb{E}(x_1 x_2 x_5 \mid x_3, x_4)$, i.e. the expected value of the product of x_1, x_2, and x_5 given x_3 and x_4, with a sample average. Do you need to have joint observations for all five variables $x_1, \ldots, x_5$?*

Solution. In the graph, all trails from $\{x_1, x_2\}$ to x_5 are blocked by $\{x_3, x_4\}$, so that $x_1, x_2 \perp\!\!\!\perp x_5 \mid x_3, x_4$. We thus have

$$\mathbb{E}(x_1 x_2 x_5 \mid x_3, x_4) = \mathbb{E}(x_1 x_2 \mid x_3, x_4)\mathbb{E}(x_5 \mid x_3, x_4).$$

Hence, we only need joint observations of (x_1, x_2, x_3, x_4) and (x_3, x_4, x_5). Variables (x_1, x_2) and x_5 do not need to be jointly measured.

Exercise 3. ***Undirected graphical model with pairwise potentials***

We here consider Gibbs distributions where the factors only depend on two variables at a time. The probability density or mass functions over d random variables $x_1, \ldots, x_d$ then take the form

$$p(x_1, \ldots, x_d) \propto \prod_{i \leq j} \phi_{ij}(x_i, x_j)$$

These models are typically called pairwise Markov networks.

(a) *Let* $p(x_1, \ldots, x_d) \propto \exp\left(-\frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}\right)$ *where* $\mathbf{A}$ *is symmetric and* $\mathbf{x} = (x_1, \ldots, x_d)^\top$. *What are the corresponding factors* ϕ_{ij} *for* $i \leq j$*?*

Solution. Denote the (i, j)-th element of $\mathbf{A}$ by a_{ij}. We have

$$\mathbf{x}^\top \mathbf{A}\mathbf{x} = \sum_{ij} a_{ij} x_i x_j \tag{S.1}$$

$$= \sum_{i<j} 2a_{ij} x_i x_j + \sum_i a_{ii} x_i^2 \tag{S.2}$$

where the second line follows from $\mathbf{A}^\top = \mathbf{A}$. Hence,

$$-\frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x} = -\frac{1}{2}\sum_{i<j} 2a_{ij} x_i x_j - \frac{1}{2}\sum_i a_{ii} x_i^2 - \sum_i b_i x_i \tag{S.3}$$

so that

$$\phi_{ij}(x_i, x_j) = \begin{cases} \exp\left(-a_{ij} x_i x_j\right) & \text{if } i < j \\ \exp\left(-\frac{1}{2} a_{ii} x_i^2 - b_i x_i\right) & \text{if } i = j \end{cases} \tag{S.4}$$

For $\mathbf{x} \in \mathbb{R}^d$, the distribution is a Gaussian with $\mathbf{A}$ equal to the inverse covariance matrix. For binary $\mathbf{x}$, the model is known as Ising model or Boltzmann machine. For $x_i \in \{-1, 1\}$, $x_i^2 = 1$ for all i, so that the a_{ii} are constants that can be absorbed into the normalisation constant. This means that for $x_i \in \{-1, 1\}$, we can work with matrices $\mathbf{A}$ that have zeros on the diagonal.

(b) *For* $p(x_1, \ldots, x_d) \propto \exp\left(-\frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x}\right)$, *show that* $x_i \perp\!\!\!\perp x_j \mid \{x_1, \ldots, x_d\} \setminus \{x_i, x_j\}$ *if the* (i, j)*-th element of* $\mathbf{A}$ *is zero.*

Solution. The previous question showed that we can write $p(x_1, \ldots, x_d) \propto \prod_{i \leq j} \phi_{ij}(x_i, x_j)$ with potentials as in Equation (S.4). Consider two variables x_i and x_j for fixed (i, j). They only appear in the factorisation via the potential ϕ_{ij}. If $a_{ij} = 0$, the factor ϕ_{ij} becomes a constant, and no other factor contains x_i and x_j, which means that there is no edge between x_i and x_j if $a_{ij} = 0$. By the pairwise Markov property it then follows that $x_i \perp\!\!\!\perp x_j \mid \{x_1, \ldots, x_d\} \setminus \{x_i, x_j\}$.