*Exercises for the tutorials: 2 and 4.*

*The other exercises are for self-study and exam preparation. All material is examinable unless otherwise mentioned.*

## 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.

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

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

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

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

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

Consider the undirected graphical model defined by the graph in Figure 1.

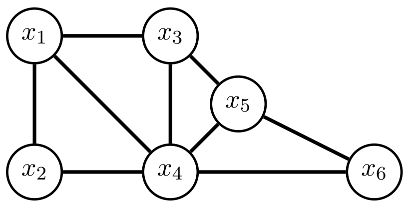



Figure 1: Graph for Exercise 2

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

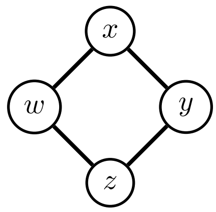
(b) Let $p$ be a pdf that factorises according to the graph. Does $p(x_3|x_2, x_4) = p(x_3|x_4)$ hold?

(c) Explain why $x_2 \perp\!\!\!\perp x_5 \mid x_1, x_3, x_4, x_6$ holds for all distributions that factorise over the graph.

(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 $x_1, \ldots, x_5$?

**Exercise 3.** ***Factorisation and independencies for undirected graphical models***

Consider the undirected graphical model defined by the following graph, sometimes called a diamond configuration.




(a) How do the pdfs/pmfs of the undirected graphical model factorise?

(b) List all independencies that hold for the undirected graphical model.

**Exercise 4.** ***Factorisation from the Markov blankets I***

Assume you know the following Markov blankets for all variables $x_1, \ldots, x_4, y_1, \ldots y_4$ of a pdf or pmf $p(x_1, \ldots, x_4, y_1, \ldots, y_4)$.

$$\begin{aligned} &\mathrm{MB}(x_1) = \{x_2, y_1\} \quad \mathrm{MB}(x_2) = \{x_1, x_3, y_2\} \quad \mathrm{MB}(x_3) = \{x_2, x_4, y_3\} \quad \mathrm{MB}(x_4) = \{x_3, y_4\} && (1)\\ &\mathrm{MB}(y_1) = \{x_1\} \quad \mathrm{MB}(y_2) = \{x_2\} \quad \mathrm{MB}(y_3) = \{x_3\} \quad \mathrm{MB}(y_4) = \{x_4\} && (2) \end{aligned}$$

Assuming that $p$ is positive for all possible values of its variables, how does $p$ factorise?

**Exercise 5.** ***Factorisation from the Markov blankets II***

We consider the same setup as in Exercise 4 but we now assume that we do not know all Markov blankets but only

$$\mathrm{MB}(x_1) = \{x_2, y_1\} \quad \mathrm{MB}(x_2) = \{x_1, x_3, y_2\} \quad \mathrm{MB}(x_3) = \{x_2, x_4, y_3\} \quad \mathrm{MB}(x_4) = \{x_3, y_4\} \qquad (3)$$

Without inserting more independencies than those specified by the Markov blankets, draw the graph over which $p$ factorises and state the factorisation. (Again assume that $p$ is positive for all possible values of its variables).

**Exercise 6.** ***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

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

Such models are sometimes 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 $\phi_{ij}$ for $i \leq j$?

(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.

**Exercise 7.** ***Restricted Boltzmann machine (based on Barber Exercise 4.4)***

The restricted Boltzmann machine is an undirected graphical model for binary variables $\mathbf{v} = (v_1, \ldots, v_n)^\top$ and $\mathbf{h} = (h_1, \ldots, h_m)^\top$ with a probability mass function equal to

$$p(\mathbf{v}, \mathbf{h}) \propto \exp\left(\mathbf{v}^\top \mathbf{W}\mathbf{h} + \mathbf{a}^\top \mathbf{v} + \mathbf{b}^\top \mathbf{h}\right), \tag{4}$$

where $\mathbf{W}$ is a $n \times m$ matrix. Both the $v_i$ and $h_i$ take values in $\{0, 1\}$. The $v_i$ are called the "visibles" variables since they are assumed to be observed while the $h_i$ are the hidden variables since it is assumed that we cannot measure them.

(a) Use graph separation to show that the joint conditional $p(\mathbf{h}|\mathbf{v})$ factorises as

$$p(\mathbf{h}|\mathbf{v}) = \prod_{i=1}^{m} p(h_i|\mathbf{v}).$$

(b) Show that

$$p(h_i = 1|\mathbf{v}) = \frac{1}{1 + \exp\left(-b_i - \sum_j W_{ji} v_j\right)} \tag{5}$$

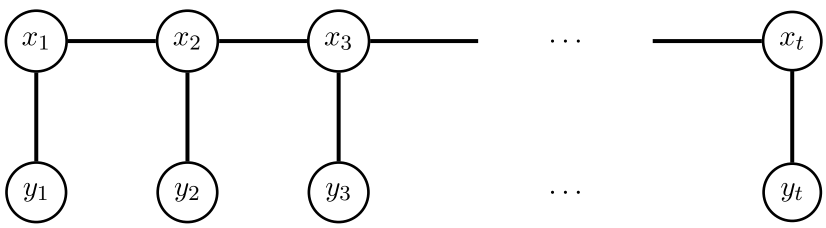
where $W_{ji}$ is the $(ji)$-th element of $\mathbf{W}$, so that $\sum_j W_{ji} v_j$ is the inner product (scalar product) between the $i$-th column of $\mathbf{W}$ and $\mathbf{v}$.

(c) Use a symmetry argument to show that

$$p(\mathbf{v}|\mathbf{h}) = \prod_i p(v_i|\mathbf{h}) \quad \text{and} \quad p(v_i = 1|\mathbf{h}) = \frac{1}{1 + \exp\left(-a_i - \sum_j W_{ij} h_j\right)}$$

**Exercise 8.** ***Hidden Markov models and change of measure***

Consider the following undirected graph for a hidden Markov model where the $y_i$ correspond to observed (visible) variables and the $x_i$ to unobserved (hidden/latent) variables.




The graph implies the following factorisation

$$p(x_1, \ldots, x_t, y_1, \ldots, y_t) \propto \phi_1^y(x_1, y_1) \prod_{i=2}^{t} \phi_i^x(x_{i-1}, x_i) \phi_i^y(x_i, y_i), \tag{6}$$

where the $\phi_i^x$ and $\phi_i^y$ are non-negative factors.

Let us consider the situation where $\prod_{i=2}^t \phi_i^x(x_{i-1}, x_i)$ equals

$$f(\mathbf{x}) = \prod_{i=2}^{t} \phi_i^x(x_{i-1}, x_i) = f_1(x_1) \prod_{i=2}^{t} f_i(x_i|x_{i-1}), \tag{7}$$

with $\mathbf{x} = (x_1, \ldots, x_t)$ and where the $f_i$ are (conditional) pdfs. We thus have

$$p(x_1, \ldots, x_t, y_1, \ldots, y_t) \propto f_1(x_1) \prod_{i=2}^{t} f_i(x_i|x_{i-1}) \prod_{i=1}^{t} \phi_i^y(x_i, y_i). \tag{8}$$

(a) Provide a factorised expression for $p(x_1, \ldots, x_t|y_1, \ldots, y_t)$

(b) Draw the undirected graph for $p(x_1, \ldots, x_t|y_1, \ldots, y_t)$

(c) Show that if $\phi_i^y(x_i, y_i)$ equals the conditional pdf of $y_i$ given $x_i$, i.e. $p(y_i|x_i)$, the marginal $p(x_1, \ldots, x_t)$, obtained by integrating out $y_1, \ldots, y_t$ from (8), equals $f(\mathbf{x})$.

(d) Compute the normalising constant for $p(x_1, \ldots, x_t|y_1, \ldots, y_t)$ and express it as an expectation over $f(\mathbf{x})$.

(e) Express the expectation of a test function $h(\mathbf{x})$ with respect to $p(x_1, \ldots, x_t|y_1, \ldots, y_t)$ as a reweighted expectation with respect to $f(\mathbf{x})$.