# COMP64101 - Lecture 4 - Exercise sheet

**Exercise 1 [Restricted Boltzmann Machine]** The restricted Boltzmann machine (RBM for short) is an undirected graphical model for vectors of binary variables $\mathbf{v} = (v_1, \ldots, v_n)^\top$ and $\mathbf{h} = (h_1, \ldots, h_m)^\top$ with a probability mass function:

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

where the matrix $\mathbf{W} \in \mathbb{R}^{n \times m}$ and vectors $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$ are parameters of this distribution. The variables $v_j$ and $h_i$ both take values in $\{0, 1\}$. The $v_j$ are called the "visible" 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 for each $i$ we have

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

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

**Exercise 2 [Restricted Boltzmann Machine, continued]** For this exercise, we continue considering the RBM from Exercise 1. Use a symmetry argument to show that

(a) The joint conditional $p(\mathbf{v}|\mathbf{h})$ factorises as

$$p(\mathbf{v}|\mathbf{h}) = \prod_{j=1}^{n} p(v_j|\mathbf{h}).$$

(b) For each $j$ we have

$$p(v_j = 1|\mathbf{h}) = \frac{1}{1 + \exp\left(-a_j - \sum_i W_{ji} h_i\right)}.$$

**Exercise 3 [Sampling from a restricted Boltzmann machine]** For this exercise, we continue considering the RBM from Exercise 1. Explain how to use Gibbs sampling to generate samples from the marginal $p(\mathbf{v})$ for any given values of $\mathbf{W}$, $\mathbf{a}$, and $\mathbf{b}$. Notice that the marginal has the form:

$$p(\mathbf{v}) = \frac{\sum_{\mathbf{h}} \exp\left(\mathbf{v}^\top \mathbf{W}\mathbf{h} + \mathbf{a}^\top \mathbf{v} + \mathbf{b}^\top \mathbf{h}\right)}{\sum_{\mathbf{h},\mathbf{v}} \exp\left(\mathbf{v}^\top \mathbf{W}\mathbf{h} + \mathbf{a}^\top \mathbf{v} + \mathbf{b}^\top \mathbf{h}\right)}$$

*Hint:* Feel free to use the identities established in the previous two exercises.