CSE 291: Unsupervised learning

Homework 8

This homework is due on Saturday June 10 at 11.59pm.

- All homeworks must be typewritten and uploaded to Gradescope.
- No late homeworks will be accepted.

1. A *restricted Boltzmann machine* specifies a probability distribution over a set of *visible units* $v_1, \ldots, v_n \in \{0, 1\}$ and *hidden units* $h_1, \ldots, h_m \in \{0, 1\}$. The distribution is of the form

$$P(v, h) = \frac{1}{Z} \exp\left(\sum_{i,j} w_{ij} v_i h_j + \sum_i \alpha_i v_i + \sum_j \beta_j h_j \right),$$

where the matrix $W = (w_{ij}) \in \mathbb{R}^{n \times m}$ and the vectors $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}^m$ are parameters of the model.

   (a) Draw the minimal undirected graph for this family of distributions.

   (b) It follows from the graph that the conditional distribution over hidden units given visible units, $P(h|v)$, can be written as

$$P(h|v) = \prod_j P(h_j|v).$$

   Derive the functional form of $P(h_j|v)$.

   (c) A similar factorization applies to $P(v|h)$. What is the functional form of $P(v_i|h)$?

   (d) Write down an algorithm for generating a sample v from this distribution. *Hint:* The easiest solution is probably to use a Gibbs sampler.

2. Consider a random walk on state space $\mathcal{X} = \{1, 2, 3\}$ given by the transition matrix

$$M = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/3 & 2/3 \\ 2/3 & 0 & 1/3 \end{pmatrix}.$$

(Recall that M_{ij} is the probability of transitioning from i to j in a single step of the walk.) Find all stationary distributions of this random walk.

3. *Random walk on regular undirected graph.* Let $\mathcal{X}$ be any finite set and let $G = (\mathcal{X}, E)$ be an undirected graph with nodes $\mathcal{X}$ and edges E. Suppose that

   - G is connected (that is, there is a path between any pair of nodes), and
   - each vertex in G has exactly the same number of edges incident upon it.

We define a random walk on $\mathcal{X}$ as follows: if the current state is $x \in \mathcal{X}$, then the next state x' is either x itself (with probability $1/2$) or a neighbor of x in G (where all neighbors are chosen with equal probability).

(a) Show that this random walk has a unique stationary distribution.

(b) Say what the stationary distribution is, and use the detailed balance condition to prove that your answer is correct.

4. *A Bayesian approach to image denoising.* Download the file `gibbs.zip` from the course website. This contains a Jupyter notebook and two images (`noisy_bear.png` and `original_bear.png`). The former is a corrupted version of the latter (which is provided only for reference).

   Let X be the corrupted image and Y the uncorrupted version that we wish to recover. Both are of the form $\{-1,+1\}^{M \times N}$. In the notebook, a Markov random field joint distribution for (X, Y) is given. You will use this to design a Gibbs sampler for $Y|X$.

   (a) Let p denote a specific pixel in the image. Then Y_p is the (reconstructed) value at that pixel, and $Y_{\setminus p}$ is the value at all remaining pixels. Give a precise expression for $\Pr(Y_p = +1|Y_{\setminus p}, X)$.

   (b) Implement a Gibbs sampler at the appropriate place in the notebook. Initialize Y by setting each pixel to $\{-1,+1\}$ uniformly at random. Then apply the sampler and show the image Y at various stages while the Markov chain is mixing: show any **five** of the images obtained along the way.

   (c) To reconstruct the image, just set the value of each pixel p to $\mathbb{E}[Y_p|X]$. Show this image.