Problem 1 Markov Network

We first need to find all the messages entering node a. The message from node d is trivial, $m_{da} = \Phi e_2 = [0.1 \quad 0.9]^T$. We then need to propagate the information from node e all the way to node a. Firstly, $m_{ec} = \Phi e_1 = [0.9 \quad 0.1]^T$. Then

$$\begin{aligned} m_{cb} &= \Psi(\alpha) m_{ec} = \begin{bmatrix} 0.8\alpha + 0.1 \\ -0.8\alpha + 0.9 \end{bmatrix} \\ m_{ba} &= \Psi(\alpha) m_{cb} = \begin{bmatrix} 1.6\alpha^2 - 1.6\alpha + 0.9 \\ -1.6\alpha^2 + 1.6\alpha + 0.1 \end{bmatrix} \end{aligned}$$

Now we can find the marginal distribution as the element wise product of the incoming messages, which gives

$$P(\alpha) = m_{ba} \odot m_{da} = \begin{bmatrix} 0.16\alpha^2 - 0.16\alpha + 0.09 \\ -1.44\alpha^2 + 1.44\alpha + 0.09 \end{bmatrix} \tag{1}$$

Finally, we can then calculate $P(0.99) = [0.088416 \quad 0.104256]^T$, which is normalized to $[0.458894 \quad 0.541106]^T$.

And $P(0.6) = [0.0515 \quad 0.4356]^T$, which is normalized to $[0.105911 \quad 0.894089]^T$.

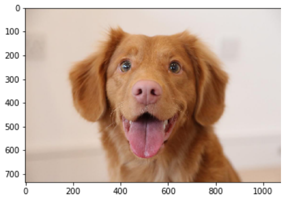
In the first case, all the nodes prefer to be like its neighbours (weighted 0.9 and 0.99), such that the nodes between d and e is about equally distributed as d is in state 1 and e is in state 0, i.e. they "compete" with each other. Node a is slightly favoured for state 1, as it is closest to d.

For the second case, in the links between a and b, and b and c are almost equal for both states, such that the state of e is almost indifferent for the state of node a. Thus node a will mostly prefer to be in the state of d.

Problem 2

See the notebook for code and implementation details.

(a) Flower

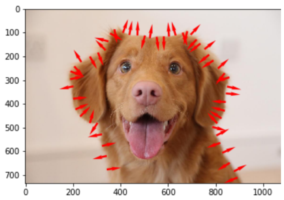


(b) Dog

(c) Pedestrian

Figure 1: Images for problem a

(a) Flower

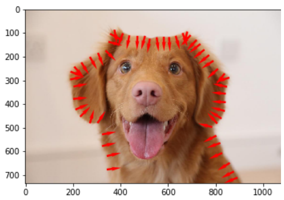


(b) Dog

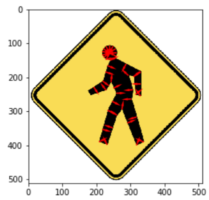
(c) Pedestrian

Figure 2: Images for problem b

(a) Flower



(b) Dog



(c) Pedestrian

Figure 3: Images for problem c

Problem 3 Restricted Boltzmann Machines

Using the law of conditional probability we have

$$P(h|v) = \frac{P(h,v)}{P(v)} = \frac{P(h,v)}{\sum_h P(h,v)} \tag{2}$$

We can now cancel out $\exp(-\sum a_i v_i)/Z$ which gives

$$\begin{aligned} P(h|v) &= \frac{exp(-\sum_j h_j b_j + \sum_i w_{ij} v_i h_j)}{\sum_h \exp(-\sum_j h_j v_j + \sum_i w_{ij} v_i h_j)} \quad (3) \\ &= \frac{\prod_j exp(-h_j(b_j + \sum_i w_{ij} v_i))}{f(v)} \quad (4) \end{aligned}$$

where $f(v)$ is some normalization constant given as a function of v. To show that this has the desired sigmoidal form, we can compute

$$\begin{aligned} P(h_j = 1|v) &= \frac{P(h_j = 1|v)}{P(h_j = 1|v) + P(h_j = 0|v)} \quad (5) \\ &= \frac{exp(b_j + \sum_i w_{ij} v_i)}{1 + exp(b_j + \sum_i w_{ij} v_i)} \quad (6) \\ &= \sigma(b_j + \sum_i w_{ij} v_i) \quad (7) \end{aligned}$$

since h_j only has the two states 0 and 1 and the probability of these must sum to 1.

b

The equations for $P(v|h)$ is the same, but with $a, b = b, a$, $h, v = v, h$ and $j, i = i, j$.

$$P(v|h) = \prod_i (v_i|h) \tag{8}$$

$$P(v_i = 1|h) = \sigma(a_i + \sum_j w_{ij} h_j) \tag{9}$$

c

Yes it does. Since the total probability is just the product of the conditionals for the individual nodes, this indicated that a given hidden node is independent from the other hidden nodes, and a given visible node is independent from the other visible nodes.

For me, it is not obvious from equation (1) that h and v have this element wise independence, i.e. h_j is independent from h_i for $i \neq j$, and similar for v. However, this is indicated by the figure.