Machine Learning and Bayesian Inference

Problem Sheet IV: Bayesian networks and hidden Markov models

Sean B. Holden © 2010-17

1 Bayesian networks

1. Continuing with the running example of the roof-climber alarm...

   The porter in lodge 1 has left and been replaced by a somewhat more relaxed sort of chap, who doesn't really care about roof-climbers and therefore acts according to the probabilities

$$\begin{array}{cc} \Pr(l1|a) = 0.3 & \Pr(\neg l1|a) = 0.7 \\ \Pr(l1|\neg a) = 0.001 & \Pr(\neg l1|\neg a) = 0.999 \end{array}.$$

   Your intrepid roof-climbing buddy is on the roof. What is the probability that lodge 1 will report him? Use the variable elimination algorithm to obtain the relevant probability. Do you learn anything interesting about the variable $L2$ in the process?

2. In designing a Bayesian network you wish to include a node representing the value reported by a sensor. The quantity being sensed is real-valued, and if the sensor is working correctly it provides a value close to the correct value, but with some noise present. The correct value is provided by its first parent. A second parent is a Boolean random variable that indicates whether the sensor is faulty. When faulty, the sensor flips between providing the correct value, although with increased noise, and a known, fixed incorrect value, again with some added noise. Suggest a conditional distribution that could be used for this node.

2 Hidden Markov models

1. Derive the equation

$$b_{t+1:T} = \mathbf{S}\mathbf{E}_{t+1} b_{t+2:T}$$

   for the backward message in a hidden Markov model.

2. Explain why the backward message update should be initialized with the vector $(1, \ldots, 1)$.

3. Establish how the prior $\Pr(S_0)$ should be included in the derivation of the Viterbi algorithm. (This is mentioned in the lectures, but no detail is given.)

4. A hidden Markov model has transition matrix $S_{ij} = \Pr(S_{t+1} = s_j | S_t = s_i)$ where

$$\mathbf{S} = \begin{pmatrix} 0.2 & 0.4 & 0.4 \\ 0.1 & 0.6 & 0.3 \\ 0.8 & 0.1 & 0.1 \end{pmatrix}.$$

In any state we observe one of the symbols $\triangle$, $\bigtriangledown$, $\bigcirc$, $\square$ with the following probabilities:

	$\triangle$	$\bigtriangledown$	$\bigcirc$	$\square$
s_1	0.7	0.1	0.1	0.1
s_2	0.3	0.2	0.4	0.1
s_3	0.4	0.2	0.2	0.2

Prior probabilities for the states are Pr $(s_1) = 0.3$, Pr $(s_2) = 0.3$ and Pr $(s_3) = 0.4$. We observe the sequence of symbols

$$\bigcirc \bigcirc \square \triangle \triangle \square \bigtriangledown \square.$$

Use the Viterbi algorithm to infer the most probable sequence of states generating this sequence.

3 Old exam questions

On Bayesian Networks:

1. 2005, paper 8, question 2.
2. 2006, paper 8, question 9.
3. 2009, paper 8, question 1.
4. 2014, paper 7, question 2.
5. 2016, paper 7, question 3.

On hidden Markov models:

1. 2005, paper 9, question 8.
2. 2008, paper 9, question 5.
3. 2010, paper 7, question 4.
4. 2013, paper 7, question 2.