Machine learning 2
Exercise sheet 9

FLEISCHMANN Kay, Matrnr: 352247
ROHRMANN Till, Matrnr: 343756

June 16, 2013

1 Hidden Markov Models

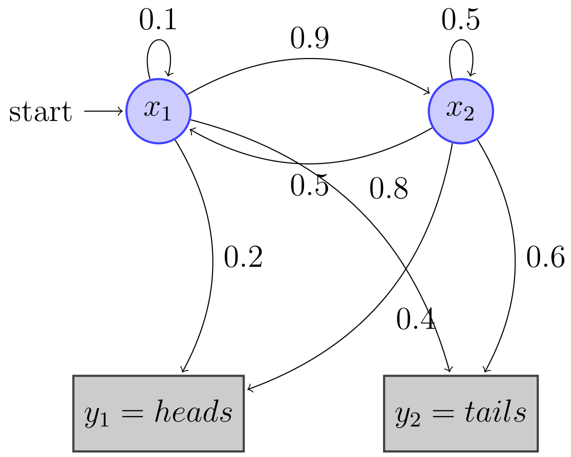
Let $A_{i,j}$ the transition matrix between hidden states x_i and x_j. $B_{i,j}$ is the the probability, beeing in state i to overserve y_j. The following matrices A and B describe two hidden states and two possible observations.

$$A = \begin{pmatrix} 0.1 & 0.9 \\ 0.5 & 0.5 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{pmatrix}$$

a. Draw the graph of the model

The Hidden-Markov-chain for A and B looks like the following4



Because of $\pi = (1|0)$ all state-sequences start with the initial state x_1.

b.

We interpret the above as a model for an experiment with two hidden (unfair) coins and two visible coins. Describe such an experiment which can be modelled by the Markov model given above. Here, heads should correspond to the first indices in A, B, π, heads to the second.

A guy is standing behind an curtain and is throwing two unfair coins, just one each time but starts always with the first one. A second guy in front of the curtain just get the information about the observed heads or tails as a sequence. The task of the second guy is to find the most likely sequence of the thrown coins.

c.

Given the bayes rule:

$$P(x|y_1 = tails, y_2 = tails) = \frac{P(y_1 = tails, y_2 = tails|x)P(x)}{P(y_1 = tails, y_2 = tails)} \tag{1}$$

Because of the statistical independence, we can write

$$P(y_1 = tail, y_2 = tail|x_1, x_2) = \pi_1 \cdot P(y_1|x_1) \cdot P(x_1) \cdot P(y_2|x_2) \cdot P(x_2) \tag{2}$$

In general, for the first two states which should produce the observation $(tails, tails)$ we can write

$$P(y_1 = tail, y_2 = tail|x_1, x_2) = \pi_i \cdot P(y_1|x_i) \cdot P(x_i) \cdot P(y_2|x_i) \cdot P(x_i) \tag{3}$$

All possible sequences with length 2 are

	State transitions	$P(x)$	$P(y_1 = tails, y_2 = tails\|x)$
1	$x_1 \to x_1$	$P(x_1, x_1) = 1 \cdot 0.1 = 0.1$	$1 \cdot 0.4 \cdot 0.1 \cdot 0.4 = 0,016$
2	$x_1 \to x_2$	$P(x_1, x_2) = 1 \cdot 0.9 = 0.9$	$1 \cdot 0.4 \cdot 0.9 \cdot 0.6 = 0,216$
3	$x_2 \to x_1$	$P(x_2, x_1) = 0 \cdot 0.5 = 0$	0
4	$x_2 \to x_2$	$P(x_2, x_2) = 0 \cdot 0.5 = 0$	0

$$\begin{align} P(y_1, y_2) &= \sum_x P(y_1, y_2|x)P(x) \tag{4} \\ &= \sum_x P(y_1 = tails, y_2 = tails|x)P(x) \tag{5} \\ &= 0.016 \cdot 0.1 + 0.216 \cdot 0.9 \tag{6} \\ &= 0.196 \tag{7} \end{align}$$

The result for $P(x|y_1 = tails, y_2 = tails)$ follows as

	State-sequence	$P(x\|y_1 = tails, y_2 = tails)$
1	$x = x_1, x_1$	$\frac{0.016 \cdot 0.1}{0.196} = 0.0081$
2	$x = x_1, x_2$	$\frac{0.216 \cdot 0.9}{0.196} = 0.9918$
3	$x = x_2, x_1$	$\frac{0 \cdot 0.45}{0.196} = 0$
4	$x = x_2, x_2$	$\frac{0 \cdot 0.45}{0.196} = 0$

f.

Prove in the scenario of (e) and the assumptions of the exercise: if $N \to \infty$, then the sample observation frequency of heads/tails at fixed position n converges, in probability and entry-wise, to the vector $\pi \cdot A^n \cdot B$.