

# EQ2340 Pattern Recognition

## Exercise 1

Kalogiras Vasileios  
vaskal@kth.se

Tatarakis Nikolaos  
nta@kth.se

September 26, 2016

### 1

Your theoretically calculated  $P(\text{St} = j)$  for the first infinite-duration HMM, and your corresponding measured relative frequencies.

From Markov chains theory we know for an ergodic square matrix  $A$  that  $\lim_{n \rightarrow \infty} A^n$  will converge to the overall distributions of our states... If we do this for our matrix we converge to  $P(x = 1) = 0.75$  and  $P(x = 2) = 0.25$ . In figure 1 we see the number of times we picked each state on a 10000 long sequence. We can observe that the actual counts produced from our Markov chain closely approach the true probabilities calculated previously.

### 2

Your theoretically calculated  $E[X_t]$  and  $\text{var}[X_t]$ , and your corresponding measured results.

For  $E[X_t]$  we get from the pdf that is calculated as the mean of the submeans of our two distributions. That is

$$E[X_t] = E[B_1]P(S_1) + E[B_2]P(S_2) = 0 \times 0.75 + 3 \times 0.25 = 0.75$$

As for the variance we get:

$$\begin{aligned} \text{var}[X] &= E_Z[\text{var}_X[X|Z]] + \text{var}_Z[E_X[X|Z]] \\ &= P(S_1)\text{Var}(S_1) + P(S_2)\text{Var}(S_2) + E[E_X[X|S]^2] - E[E_X[X|S]]^2 \\ &= 0.75 \times 1 + 0.25 \times 4 + 0.75 \times 0 + 0.25 \times 9 - 0.75 \times 0.75 \\ &= 3.4375 \end{aligned}$$

Now all we have to do is generate a sequence of size 10000 and check whether its mean and variance are similar. For such a sequence  $S_t$  using matlab we got

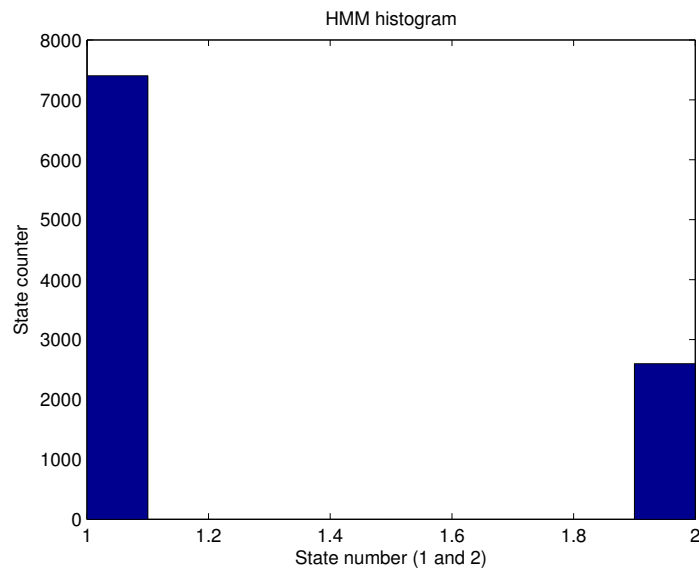


Figure 1: histogram of a 10000 state sequence

$E[S_t] = 0.67$  and  $Var[X_t] = 3.2297$  which are close to the actual values more or less. In general however we always got some fluctuation over the runs.

### 3

A plot of 500 contiguous values randomized from the first infinite duration HMM, with a description of typical output behaviour.

In figure 2 we can see five different observation sequences produced from our HMM. It is clear that the observations vary around the values 0 and 3 which are the means of our two distributions and we expected them to be that way. In figure 3 we can see the trend clearer for a single observation. What is more we can almost confidently say from which distribution each value came from and which state accordingly.

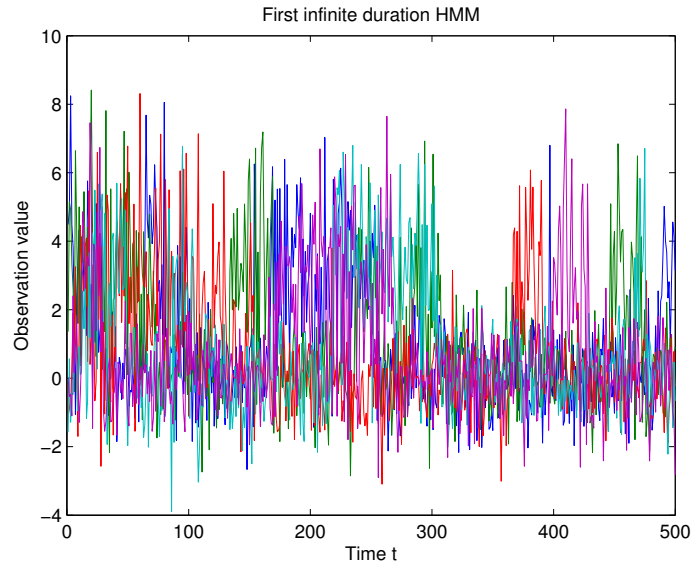


Figure 2: 5 different observations of length 500 each

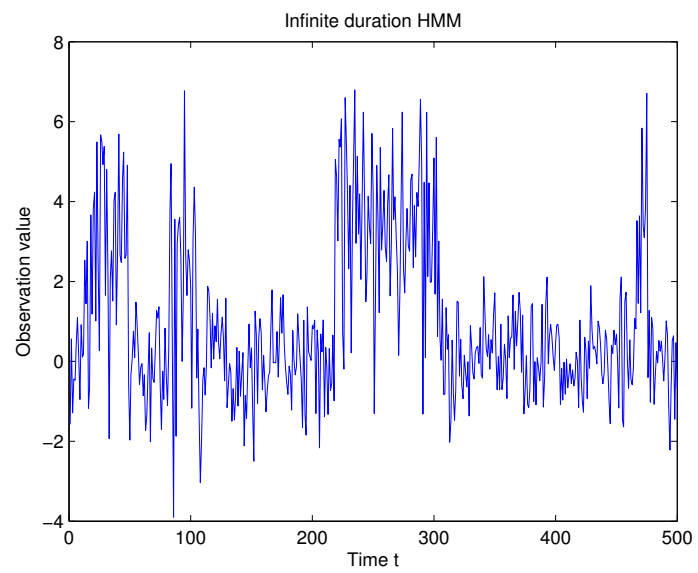


Figure 3: 1 observation of length 500

## 4

A discussion of the output behaviour of the second infinite-duration HMM, with answers to the associated questions.

In figure 4 an observation using Gaussians with zero mean is plotted. In this case it is much more difficult to guess which Gaussian/state is responsible for each observation value. We can probably guess that values deviating much from the mean are coming from the second one but for values close to zero now it is impossible to guess.

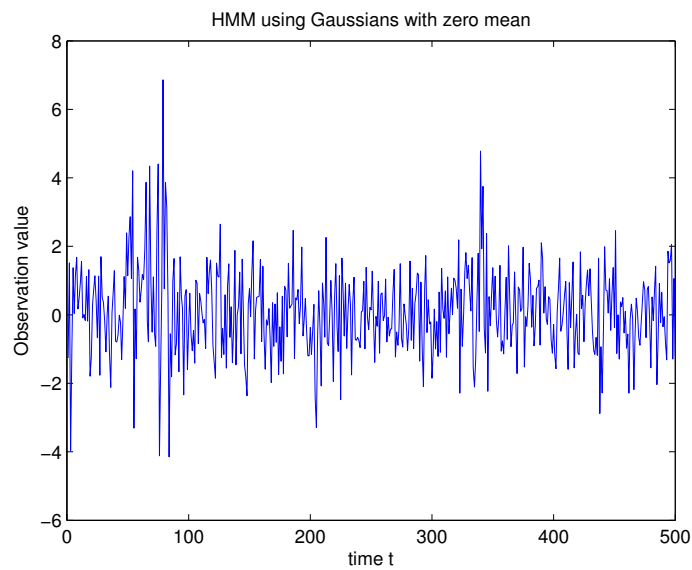


Figure 4: 1 observation of length 500

Table 1: State matrix of a finite markov chain

0.7000	0.2800	0.0100	0.0100
0.1500	0.8100	0.0300	0.0100
0.1000	0.0800	0.8000	0.0200
0	0	0	1.0000

## 5

The definition of your finite-duration test HMM, together with the lengths of some test sequences you obtained, and relevant code. Discuss briefly why you think those lengths are reasonable

When having an absorbing Markov chain you can easily calculate which is the expected size of a state sequence [Wik16] before you reach the end state... For example, let's say we have the following state matrix  $S$  like in table 1

with  $Q$  being the  $3 \times 3$  state matrix without the end state,  $eye$  a diagonal matrix of size 3 and  $ones$  a vector of size 3 with only ones inside. It will hold that  $((eye - Q)^{-1}) \times ones$  will give as the average size of a state sequence. In our case that size is  $\approx 90$ . Now, in order to check our Markov chain implementation we generated 10000 state sequences using the aforementioned state matrix and calculated their mean size. In figure 5 is the histogram of the sequences; their mean size turned out to be 89,5258 which is very close to what we previously calculated.

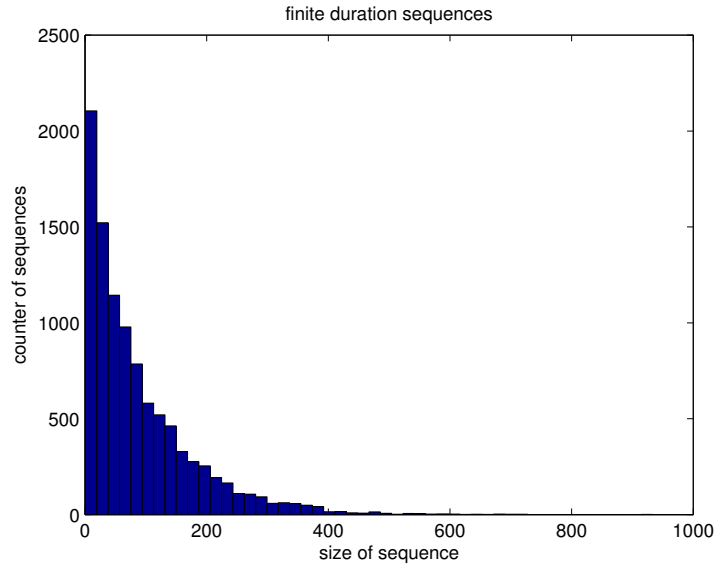


Figure 5: size histogram of 10000 finite duration state sequences

## 6

The definition of your vector-valued test HMM, and the code you used to verify that vector output distributions work with your implementation.

In order to check the vector-valued HMM implemetation we used the this state matrix:

$$S = \begin{bmatrix} .95 & .03 & .02 \\ .03 & .95 & .02 \\ 03 & 0.02 & 0.95 \end{bmatrix}$$

And the following 2d Gaussians:

$$\begin{aligned} a &= \mathcal{N}\left(\begin{bmatrix} 10 & 8 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 100 \end{bmatrix}\right), \\ b &= \mathcal{N}\left(\begin{bmatrix} -10 & -15 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 3 & 8 \end{bmatrix}\right), \\ c &= \mathcal{N}\left(\begin{bmatrix} 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}\right) \end{aligned}$$

which we deliberately put their means further apart in order for them to be easily distinguishable... Afterwards, by sampling 10000 observations we got the values plotted in figure 6

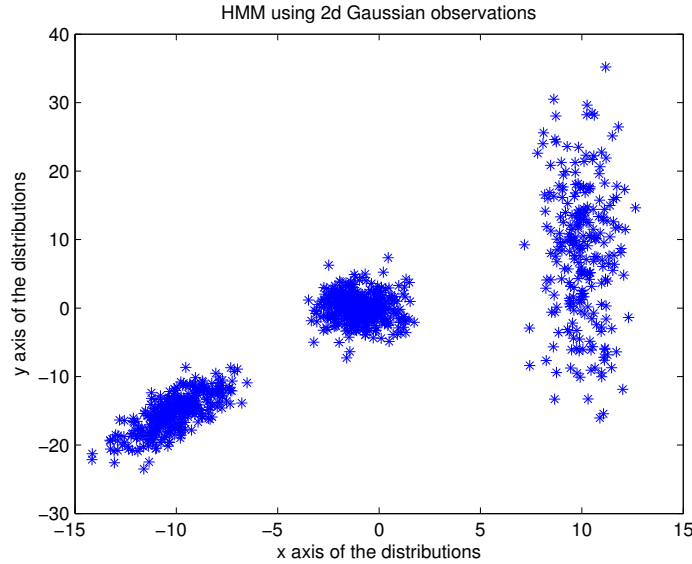


Figure 6: Plot of 10000 observations sampled from 2d Gaussians

From, there we can clearly distinguish the three Gaussian clusters which approach the true means of our three distributions too.

## References

- [Wik16] Wikipedia. *Absorbing Markov chain* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 24-September-2016]. 2016. URL: %5Curl%7Bhttps://en.wikipedia.org/w/index.php?title=Absorbing\_Markov\_chain&oldid=721021760%7D.