

LAB2_AML

Andreas C Charitos [andch552]

9/29/2019

Contents

| | |
|--|----------|
| Assignment 1 | 2 |
| Transition Matrix | 2 |
| Emission Matrix | 2 |
| Plot of Transmission and Emission Matrix | 3 |
| Assignment 2 | 3 |
| Assignment 3 | 3 |
| Assignment 4 | 3 |
| Table with the Accuracies from simulation | 3 |
| Assignment 5 | 4 |
| Table with the accuracies for different simulations | 4 |
| Plot of Accuracies for different simulations | 5 |
| Assignment 6 | 5 |
| Table with the accuracies for different sample sizes | 5 |
| Plot of the Accuracies for different samples | 6 |
| Plot of the Entropies for different samples | 7 |
| Plot of Entropies for Filter and Smoothed | 7 |
| Assignment 7 | 8 |
| Conclusion | 8 |
| Appendix | 9 |
| Code used for Lab | 9 |

Assignment 1

We are asked to model the behavior of a robot that walks around a ring. The ring is divided into 10 sectors. At any given time point, the robot is in one of the sectors and decides with equal probability to stay in that sector or move to the next sector. You do not have direct observation of the robot. However, the robot is equipped with a tracking device that you can access. The device is not very accurate though: If the robot is in the sector i , then the device will report that the robot is in the sectors $[i-2, i+2]$ with equal probability. Start by defining the model which requires the following components :

- The hidden states z_t
- The observed states x_t
- The transition matrix with the probabilities for the hidden states
- The emission matrix with the probabilities for the observed states

The states are :

State_1 State_2 State_3 State_4 State_5 State_6 State_7 State_8 State_9 State_10

The observed states are :

a b c d e f g h i j

Next we report the Transition and emission matrices

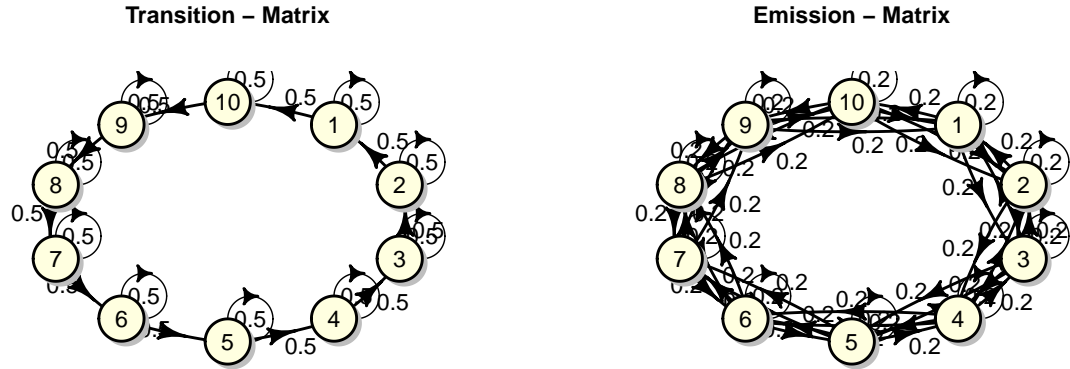
Transition Matrix

| State_1 | State_2 | State_3 | State_4 | State_5 | State_6 | State_7 | State_8 | State_9 | State_10 |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|
| 0.5 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.5 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.5 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 |
| 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.5 |

Emission Matrix

| a | b | c | d | e | f | g | h | i | j |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0.2 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 |
| 0.2 | 0.2 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 |
| 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.0 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.0 |
| 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 |
| 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.2 |

Plot of Transition and Emission Matrix



Assignment 2

Next we simulate from the defined model 100 timesteps

Assignment 3

We then discard the hidden states from the sample obtained above. And use the remaining observations to compute the filtered and smoothed probability distributions for each of the 100 time points. We compute also the most probable path.

Assignment 4

Now we calculate the accuracy for the filter, smooth and most probable path .

Table with the Accuracies from simulation

| | Filter | Smooth | Viterbi |
|----------|--------|--------|---------|
| Accuracy | 0.48 | 0.65 | 0.57 |

We observe that the accuracy of the smooth is better than filter and viterbi. The reasoning behind this is in the way the filtering and smoothing are calculated :

- Filtering : $p(Z^t | x^{0:t}) = \frac{\alpha(Z^t)}{\sum_z t \alpha(z^t)}$
- Smoothing : $p(Z^t | x^{0:T}) = \frac{\alpha(Z^t) \beta(Z^t)}{\sum_z \alpha(z^t) \beta(z^t)}$

The filter is calculating the probability of state given the observations up to this timestep while smooth is calculating the state given all the previous observations.

Assignment 5

Here we simulate 100 different simulated samples with 100 timesteps. The purpose of this procedure is to see if the that the smoothed distribution should be more accurate than the filtered distributions.

Table with the accuracies for different simulations

| | replicate n= 1 | replicate n= 2 | replicate n= 3 | replicate n= 4 | replicate n= 5 | replicate n= 6 | replicate n= 7 | replicate n= 8 | replicate n= 9 | replicate n= 10 |
|------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|
| filter_accuracy | 0.54 | 0.42 | 0.54 | 0.59 | 0.47 | 0.56 | 0.50 | 0.61 | 0.41 | 0.51 |
| smooth_accuracy | 0.75 | 0.58 | 0.77 | 0.66 | 0.67 | 0.70 | 0.59 | 0.64 | 0.60 | 0.59 |
| viterbi_accuracy | 0.50 | 0.50 | 0.63 | 0.47 | 0.67 | 0.54 | 0.51 | 0.51 | 0.59 | 0.51 |

| | replicate n= 11 | replicate n= 12 | replicate n= 13 | replicate n= 14 | replicate n= 15 | replicate n= 16 | replicate n= 17 | replicate n= 18 | replicate n= 19 | replicate n= 20 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.50 | 0.60 | 0.53 | 0.61 | 0.63 | 0.45 | 0.64 | 0.46 | 0.58 | 0.46 |
| smooth_accuracy | 0.69 | 0.73 | 0.69 | 0.72 | 0.64 | 0.67 | 0.77 | 0.59 | 0.71 | 0.73 |
| viterbi_accuracy | 0.48 | 0.44 | 0.64 | 0.66 | 0.51 | 0.39 | 0.42 | 0.62 | 0.48 | 0.65 |

| | replicate n= 21 | replicate n= 22 | replicate n= 23 | replicate n= 24 | replicate n= 25 | replicate n= 26 | replicate n= 27 | replicate n= 28 | replicate n= 29 | replicate n= 30 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.62 | 0.55 | 0.49 | 0.52 | 0.65 | 0.40 | 0.61 | 0.57 | 0.51 | 0.47 |
| smooth_accuracy | 0.70 | 0.69 | 0.61 | 0.61 | 0.69 | 0.65 | 0.79 | 0.71 | 0.59 | 0.63 |
| viterbi_accuracy | 0.44 | 0.49 | 0.51 | 0.42 | 0.52 | 0.44 | 0.52 | 0.44 | 0.44 | 0.47 |

| | replicate n= 31 | replicate n= 32 | replicate n= 33 | replicate n= 34 | replicate n= 35 | replicate n= 36 | replicate n= 37 | replicate n= 38 | replicate n= 39 | replicate n= 40 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.51 | 0.54 | 0.46 | 0.53 | 0.61 | 0.51 | 0.60 | 0.68 | 0.67 | 0.66 |
| smooth_accuracy | 0.71 | 0.65 | 0.54 | 0.62 | 0.72 | 0.76 | 0.57 | 0.69 | 0.81 | 0.73 |
| viterbi_accuracy | 0.53 | 0.51 | 0.33 | 0.58 | 0.40 | 0.59 | 0.34 | 0.51 | 0.60 | 0.44 |

| | replicate n= 41 | replicate n= 42 | replicate n= 43 | replicate n= 44 | replicate n= 45 | replicate n= 46 | replicate n= 47 | replicate n= 48 | replicate n= 49 | replicate n= 50 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.60 | 0.58 | 0.53 | 0.58 | 0.52 | 0.65 | 0.46 | 0.47 | 0.48 | 0.43 |
| smooth_accuracy | 0.71 | 0.77 | 0.72 | 0.70 | 0.76 | 0.77 | 0.65 | 0.77 | 0.63 | 0.64 |
| viterbi_accuracy | 0.48 | 0.41 | 0.46 | 0.63 | 0.49 | 0.49 | 0.49 | 0.38 | 0.34 | 0.50 |

| | replicate n= 51 | replicate n= 52 | replicate n= 53 | replicate n= 54 | replicate n= 55 | replicate n= 56 | replicate n= 57 | replicate n= 58 | replicate n= 59 | replicate n= 60 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.50 | 0.55 | 0.61 | 0.54 | 0.44 | 0.38 | 0.55 | 0.57 | 0.51 | 0.59 |
| smooth_accuracy | 0.67 | 0.73 | 0.75 | 0.66 | 0.70 | 0.58 | 0.64 | 0.73 | 0.66 | 0.64 |
| viterbi_accuracy | 0.47 | 0.64 | 0.66 | 0.43 | 0.43 | 0.54 | 0.53 | 0.49 | 0.44 | 0.46 |

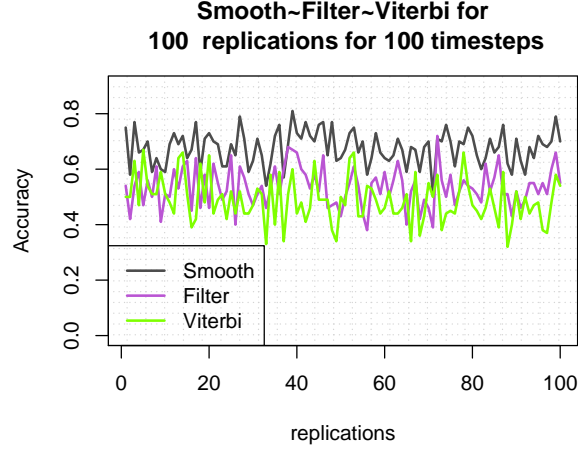
| | replicate n= 61 | replicate n= 62 | replicate n= 63 | replicate n= 64 | replicate n= 65 | replicate n= 66 | replicate n= 67 | replicate n= 68 | replicate n= 69 | replicate n= 70 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.50 | 0.56 | 0.63 | 0.56 | 0.40 | 0.52 | 0.56 | 0.42 | 0.49 | 0.46 |
| smooth_accuracy | 0.63 | 0.65 | 0.71 | 0.67 | 0.59 | 0.68 | 0.67 | 0.59 | 0.68 | 0.70 |
| viterbi_accuracy | 0.52 | 0.44 | 0.44 | 0.46 | 0.51 | 0.34 | 0.59 | 0.36 | 0.43 | 0.55 |

| | replicate n= 71 | replicate n= 72 | replicate n= 73 | replicate n= 74 | replicate n= 75 | replicate n= 76 | replicate n= 77 | replicate n= 78 | replicate n= 79 | replicate n= 80 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.39 | 0.72 | 0.56 | 0.50 | 0.58 | 0.47 | 0.52 | 0.56 | 0.54 | 0.53 |
| smooth_accuracy | 0.51 | 0.71 | 0.70 | 0.76 | 0.70 | 0.61 | 0.70 | 0.69 | 0.75 | 0.72 |
| viterbi_accuracy | 0.50 | 0.58 | 0.38 | 0.44 | 0.45 | 0.44 | 0.52 | 0.66 | 0.54 | 0.47 |

| | replicate n= 81 | replicate n= 82 | replicate n= 83 | replicate n= 84 | replicate n= 85 | replicate n= 86 | replicate n= 87 | replicate n= 88 | replicate n= 89 | replicate n= 90 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| filter_accuracy | 0.51 | 0.48 | 0.62 | 0.52 | 0.57 | 0.65 | 0.51 | 0.51 | 0.43 | 0.52 |
| smooth_accuracy | 0.65 | 0.60 | 0.64 | 0.71 | 0.66 | 0.68 | 0.76 | 0.62 | 0.58 | 0.71 |
| viterbi_accuracy | 0.45 | 0.42 | 0.46 | 0.53 | 0.46 | 0.39 | 0.59 | 0.32 | 0.40 | 0.52 |

| | replicate n= 91 | replicate n= 92 | replicate n= 93 | replicate n= 94 | replicate n= 95 | replicate n= 96 | replicate n= 97 | replicate n= 98 | replicate n= 99 | replicate n= 100 |
|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|
| filter_accuracy | 0.46 | 0.49 | 0.55 | 0.55 | 0.51 | 0.55 | 0.51 | 0.60 | 0.66 | 0.55 |
| smooth_accuracy | 0.63 | 0.58 | 0.68 | 0.64 | 0.72 | 0.69 | 0.68 | 0.70 | 0.79 | 0.70 |
| viterbi_accuracy | 0.42 | 0.50 | 0.44 | 0.47 | 0.48 | 0.38 | 0.37 | 0.48 | 0.58 | 0.54 |

Plot of Accuracies for diffrent simulations



As we can see in general the performance of the smooth is better than filter and viterbi.

Assignment 6

Next, we generate samples with different sizes in order to investigate the claim that having more timesteps we can find better where the robot is. We report the table of the accuracies and the entropy for some samples.

Table with the accuracies for different sample sizes

| | nSample= 10 | nSample= 15 | nSample= 20 | nSample= 25 | nSample= 30 | nSample= 35 | nSample= 40 | nSample= 45 | nSample= 50 | nSample= 55 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| filter_accuracy | 0.3 | 0.4000000 | 0.65 | 0.48 | 0.6333333 | 0.6000000 | 0.400 | 0.5111111 | 0.46 | 0.6000000 |
| smooth_accuracy | 0.6 | 0.5333333 | 0.75 | 0.52 | 0.8333333 | 0.6285714 | 0.625 | 0.7555556 | 0.74 | 0.7090909 |
| viterbi_accuracy | 0.7 | 0.2000000 | 0.45 | 0.28 | 0.6333333 | 0.3714286 | 0.550 | 0.4444444 | 0.54 | 0.3090909 |

| | nSample= 60 | nSample= 65 | nSample= 70 | nSample= 75 | nSample= 80 | nSample= 85 | nSample= 90 | nSample= 95 | nSample= 100 | nSample= 105 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|
| filter_accuracy | 0.6000000 | 0.4307692 | 0.6000000 | 0.44 | 0.575 | 0.6470588 | 0.5333333 | 0.4736842 | 0.52 | 0.4000000 |
| smooth_accuracy | 0.7000000 | 0.6615385 | 0.6571429 | 0.48 | 0.750 | 0.6941176 | 0.6444444 | 0.7157895 | 0.57 | 0.5714286 |
| viterbi_accuracy | 0.4833333 | 0.3384615 | 0.4857143 | 0.40 | 0.600 | 0.3764706 | 0.4222222 | 0.4631579 | 0.39 | 0.4000000 |

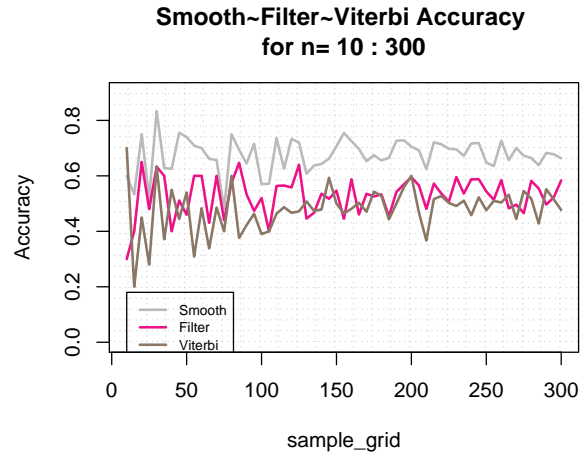
| | nSample= 110 | nSample= 115 | nSample= 120 | nSample= 125 | nSample= 130 | nSample= 135 | nSample= 140 | nSample= 145 | nSample= 150 | nSample= 155 |
|------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| filter_accuracy | 0.5636364 | 0.5652174 | 0.5583333 | 0.640 | 0.4461538 | 0.4666667 | 0.5357143 | 0.5172414 | 0.5466667 | 0.4451613 |
| smooth_accuracy | 0.7363636 | 0.6260870 | 0.7333333 | 0.720 | 0.6076923 | 0.6370370 | 0.6428571 | 0.6620690 | 0.7066667 | 0.7548387 |
| viterbi_accuracy | 0.4636364 | 0.4869565 | 0.4666667 | 0.472 | 0.5076923 | 0.4740741 | 0.4785714 | 0.5931034 | 0.5000000 | 0.4645161 |

| | nSample= 160 | nSample= 165 | nSample= 170 | nSample= 175 | nSample= 180 | nSample= 185 | nSample= 190 | nSample= 195 | nSample= 200 | nSample= 205 |
|------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| filter_accuracy | 0.58750 | 0.4606061 | 0.5352941 | 0.5257143 | 0.5333333 | 0.4540541 | 0.5421053 | 0.5692308 | 0.595 | 0.5658537 |
| smooth_accuracy | 0.72500 | 0.6969697 | 0.6529412 | 0.6742857 | 0.6555556 | 0.6648649 | 0.7263158 | 0.7282051 | 0.705 | 0.6926829 |
| viterbi_accuracy | 0.48125 | 0.5030303 | 0.4705882 | 0.5428571 | 0.5277778 | 0.4432432 | 0.5000000 | 0.5589744 | 0.600 | 0.4682927 |

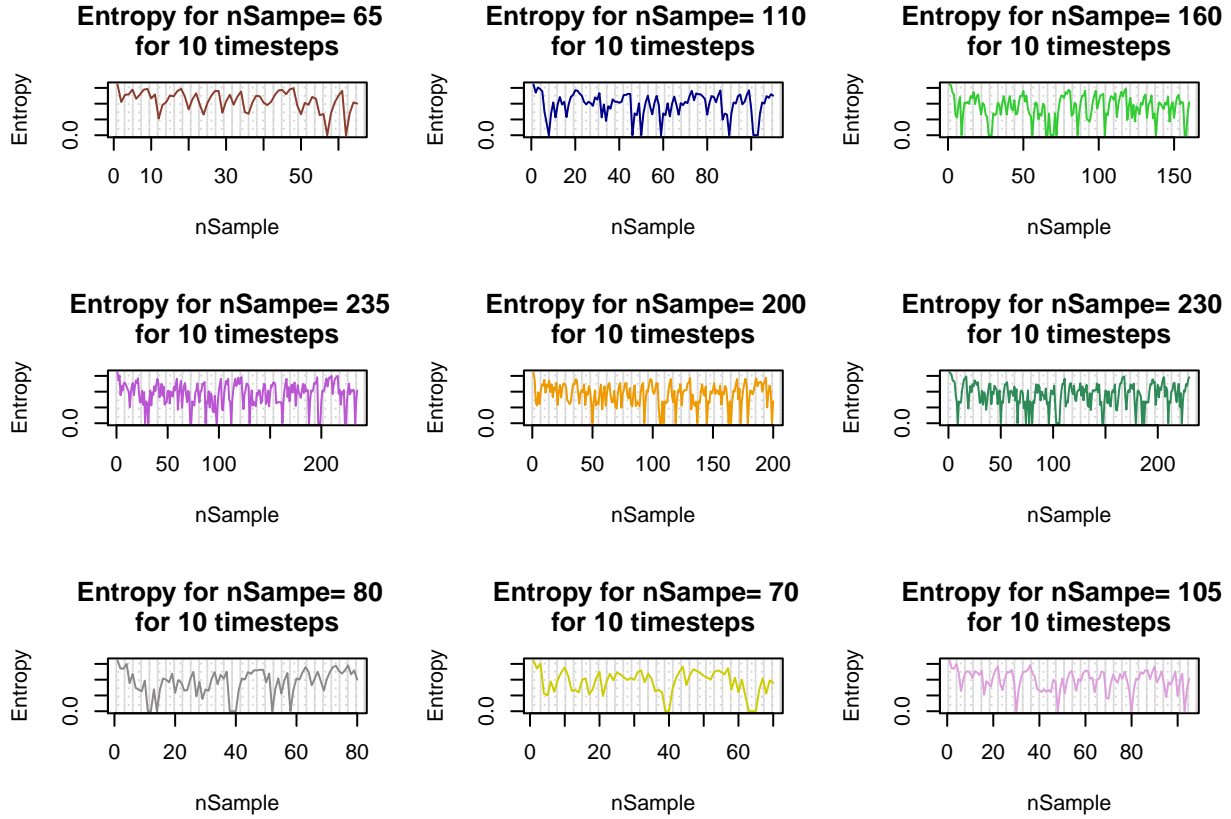
| | nSample= 210 | nSample= 215 | nSample= 220 | nSample= 225 | nSample= 230 | nSample= 235 | nSample= 240 | nSample= 245 | nSample= 250 | nSample= 255 |
|------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| filter_accuracy | 0.4809524 | 0.5720930 | 0.5363636 | 0.5066667 | 0.5956522 | 0.5361702 | 0.5875000 | 0.5877551 | 0.544 | 0.5137255 |
| smooth_accuracy | 0.6238095 | 0.7209302 | 0.7136364 | 0.6977778 | 0.6956522 | 0.6723404 | 0.7166667 | 0.7183673 | 0.648 | 0.6352941 |
| viterbi_accuracy | 0.3666667 | 0.5162791 | 0.5272727 | 0.5022222 | 0.4913043 | 0.5106383 | 0.4583333 | 0.5224490 | 0.476 | 0.5098039 |

| | nSample= 260 | nSample= 265 | nSample= 270 | nSample= 275 | nSample= 280 | nSample= 285 | nSample= 290 | nSample= 295 | nSample= 300 |
|------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| filter_accuracy | 0.5846154 | 0.4830189 | 0.4962963 | 0.4654545 | 0.5821429 | 0.5543860 | 0.4965517 | 0.5220339 | 0.5833333 |
| smooth_accuracy | 0.7269231 | 0.6566038 | 0.7000000 | 0.6727273 | 0.6642857 | 0.6385965 | 0.6827586 | 0.6779661 | 0.6633333 |
| viterbi_accuracy | 0.5038462 | 0.5320755 | 0.4444444 | 0.5454545 | 0.5178571 | 0.4280702 | 0.5517241 | 0.5152542 | 0.4766667 |

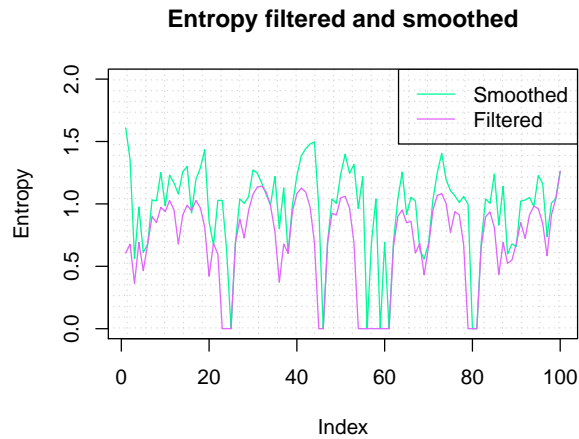
Plot of the Accuracies for diffrent samples



Plot of the Entropies for different samples



Plot of Entropies for Filter and Smoothed



So at first we plotted the entropies for different timesteps and then for a sampled simulation with 100 timesteps. As we can see there is no evidence that supports the claim that having more samples helps up to predict the state of the robot better.

Assignment 7

Finally, we take a random sample from the 100th timestep samples we created in Assignment 5 and calculate the probabilities in the 101 timestep for the hidden states.

| | prob_step101 |
|----------|--------------|
| State_1 | 0.0000000 |
| State_2 | 0.0000000 |
| State_3 | 0.0000000 |
| State_4 | 0.0000000 |
| State_5 | 0.0000000 |
| State_6 | 0.0000000 |
| State_7 | 0.0636856 |
| State_8 | 0.2682927 |
| State_9 | 0.4363144 |
| State_10 | 0.2317073 |

Conclusion

In conclusion we can see that on general the smooth will report better accuracy compared with the 2 other methods and the number of the samples does not provide any more information that can help us identify the robot position better.

Appendix

Code used for Lab

```
1  # Libraries
2  # -----
3  library(HMM)
4  library(seqHMM) #another library for HMMs
5  library(diagram)
6  library(entropy)
7  library(dplyr)
8  # -----
9  # Question 1
10 # -----
11 names = 1:10
12 States = c() # initialize the hidden states z
13 for (i in 1:10) {
14   States = c(States, paste("State_", names[i], sep = ""))
15 }
16 States # print the vector with the States x
17 Symbols = letters[1:10] # initialie the observed states
18 # transmission matrix
19 transProbs = matrix(0, 10, 10) # transmissison matrix-hidden
20 diag(transProbs) = 0.5
21 diag(transProbs[, -1]) = 0.5
22 transProbs[10, 1] = 0.5
23 emissionProbs = toeplitz(c(0.2, 0.2, 0.2, rep(0, 5), 0.2,
24   0.2)) # emission matrix-observable
25 # print transition probabilities
26 TP = as.data.frame(transProbs)
27 colnames(TP) = States
28 knitr::kable(TP)
29 # print emission probabilities
30 EP = as.data.frame(emissionProbs)
31 colnames(EP) = Symbols
32 knitr::kable(EP)
33 # initialize HMM
34 hmm = initHMM(States, Symbols, transProbs = transProbs,
35   emissionProbs = emissionProbs)
36 # using the diagram package
37 plotmat(transProbs, box.size = 0.05, box.col = "lightyellow",
38   arr.length = 0.4, main = "Transition - Matrix") # plot the transition-matrix
39 plotmat(emissionProbs, box.size = 0.05, box.col = "lightyellow",
40   arr.length = 0.4, main = "Emission - Matrix") # plot the emission-matrix
41 # -----
42 # Question 2
43 # -----
44 nSim = 100
45 sim = simHMM(hmm, nSim) # simulate 100 timesteps
46 # -----
47 # Question 3
48 # -----
49 sim.obs = sim$observation # discard the hidden states
```

```

50 # filter
51 f = forward(hmm, sim.obs)
52 f_tab = prop.table(exp(f), margin = 2) # marginalize over the columns so thats why we have 2 to prop t
53 # smooth
54 b = backward(hmm, sim.obs)
55 # b_tab=prop.table(exp(b),2)
56 smooth_ = posterior(hmm, sim.obs) # or we can use the prop.table(exp(f)*exp(b),2)
57 # viterbi algorithm for most probable path
58 vit = viterbi(hmm, sim.obs)
59 # -----
60 # Question 4
61 # -----
62 # compute accuracy for filter
63 t.index = apply(f_tab, 2, which.max) # marginalize over the columns
64 t = sapply(t.index, function(y) {
65     States[y]
66 })
67 # accuracy for filter
68 ac1 = sum(t == sim$states)/length(sim$states)
69 # compute accuracy for smooth
70 t1.index = apply(smooth_, 2, which.max)
71 t1 = sapply(t1.index, function(y) {
72     States[y]
73 })
74 # accuracy for smooth
75 ac2 = sum(t1 == sim$states)/length(sim$states)
76 # compute accuracy for viterbi
77 ac3 = sum(vit == sim$states)/length(sim$states)
78 accuracy_tab = cbind(ac1, ac2, ac3)
79 colnames(accuracy_tab) = c("Filter", "Smooth", "Viterbi")
80 rownames(accuracy_tab) = "Accuracy"
81 accuracy_tab
82 knitr::kable(as.data.frame(accuracy_tab))
83 # -----
84 # Question 5
85 # -----
86 accuracy_func = function(user_hmm, number_sim) {
87     # --function that returns the filter ,smooth,viterbi
88     # accuracy for a hmm , and number of time steps --
89     simHMM = simHMM(user_hmm, number_sim)
90     sim_hmm_obs = simHMM$observation # discard the hidden states
91     # filter
92     forward_pass = forward(user_hmm, sim_hmm_obs)
93     forward_table = prop.table(exp(forward_pass), 2) # marginalize over the cols Margin=2 to prop tabl
94     # smooth
95     back_pass = backward(user_hmm, sim_hmm_obs)
96     # back_table=prop.table(exp(back_pass),2)
97     smooth_ = posterior(user_hmm, sim_hmm_obs) # or prop.table(exp(forward_pass)*exp(back_pass),2)
98     # viterbi algorithm for most probable path
99     viterbi_res = viterbi(user_hmm, sim_hmm_obs)
100    # compute accuracy for filter
101    t.index = apply(forward_table, 2, which.max) # marginalize over the columns
102    t = sapply(t.index, function(y) {

```

```

103     user_hmm$States[y]
104   })
105   # accuracy for filter
106   filter_accuracy = sum(t == simHMM$states)/length(simHMM$states)
107   # compute accuracy for smooth
108   t1.index = apply(smooth_, 2, which.max)
109   t1 = sapply(t1.index, function(y) {
110     user_hmm$States[y]
111   })
112   # accuracy for smooth
113   smooth_accuracy = sum(t1 == simHMM$states)/length(simHMM$states)
114   # compute accuracy for viterbi
115   viterbi_accuracy = sum(viterbi_res == simHMM$states)/length(simHMM$states)
116   # compute entropy for filter
117   filter_entropy = apply(forward_table, 2, FUN = entropy.empirical)
118   return(list(simHMM, accuracies = c(filter_accuracy,
119     smooth_accuracy, viterbi_accuracy), entropy = list(filt_ent = filter_entropy,
120     smooth_ent = smooth_entropy)))
121 }
122 # replicate the results from accuracy function 10 times
123 # with 100 timespters every time
124 nreps = 100
125 rep_res = replicate(nreps, accuracy_func(hmm, 100))
126 # create a matrix with the results from the rep_res
127 # --the second row of rep_res
128 results_mat = matrix(unlist(rep_res[2, ]), nrow = 3, ncol = nreps,
129   byrow = F)
130 # colnames-rownames
131 colnames(results_mat) = paste("replicate n=", 1:nreps)
132 rownames(results_mat) = c("filter_accuracy", "smooth_accuracy",
133   "viterbi_accuracy")
134 results_mat
135 # make table knitr::kable(as.data.frame(results_mat)) #
136 # don't fit page make tables
137 t1 <- kable(as.data.frame(results_mat[, 1:10]), format = "latex",
138   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
139 t2 <- kable(as.data.frame(results_mat[, 11:20]), format = "latex",
140   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
141 t3 <- kable(as.data.frame(results_mat[, 21:30]), format = "latex",
142   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
143 t4 <- kable(as.data.frame(results_mat[, 31:40]), format = "latex",
144   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
145 t5 <- kable(as.data.frame(results_mat[, 41:50]), format = "latex",
146   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
147 t6 <- kable(as.data.frame(results_mat[, 51:60]), format = "latex",
148   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
149 t7 <- kable(as.data.frame(results_mat[, 61:70]), format = "latex",
150   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
151 t8 <- kable(as.data.frame(results_mat[, 71:80]), format = "latex",
152   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
153 t9 <- kable(as.data.frame(results_mat[, 81:90]), format = "latex",
154   booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
155 t10 <- kable(as.data.frame(results_mat[, 91:100]), format = "latex",

```

```

156     booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
157 t1
158 t2
159 t3
160 t4
161 t5
162 t6
163 t7
164 t8
165 t9
166 t10
167 # plot of the accuracies
168 cols = c("gray30", "mediumorchid", "chartreuse")
169 plot(1:nreps, results_mat[2, ], type = "l", ylim = c(0,
170     0.9), col = cols[1], lwd = 2, main = paste("Smooth~Filter~Viterbi for\n ",
171     nreps, " replications for 100 timesteps "), panel.first = grid(25,
172     25), ylab = "Accuracy", xlab = "replications")
173 lines(1:nreps, results_mat[1, ], col = cols[2], lwd = 2)
174 lines(1:nreps, results_mat[3, ], col = cols[3], lwd = 2)
175 legend("bottomleft", legend = c("Smooth", "Filter", "Viterbi"),
176     col = cols, lty = 1, lwd = 2)
177 # -----
178 # Question 6
179 # -----
180 sample_grid = seq(10, 300, 5) # generate grid of timesteps
181 # apply the accuracy function to the grid
182 accuracy_mat = sapply(sample_grid, function(y) {
183     accuracy_func(hmm, y)
184 })
185 # create a matrix with the results from the accuracy_mat
186 # --the second row of accuracy_mat
187 res_acc_mat = matrix(unlist(accuracy_mat[2, ]), nrow = 3,
188     ncol = length(sample_grid), byrow = F)
189 # row-column names
190 colnames(res_acc_mat) = paste("nSample=", sample_grid)
191 rownames(res_acc_mat) = c("filter_accuracy", "smooth_accuracy",
192     "viterbi_accuracy")
193 # print the matrix
194 res_acc_mat
195 # construct a table
196 # knitr::kable(data.frame(res_acc_mat)) # don't fit
197 # page
198 t11 <- kable(as.data.frame(res_acc_mat[, 1:10]), format = "latex",
199     booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
200 t12 <- kable(as.data.frame(res_acc_mat[, 11:20]), format = "latex",
201     booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
202 t13 <- kable(as.data.frame(res_acc_mat[, 21:30]), format = "latex",
203     booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
204 t14 <- kable(as.data.frame(res_acc_mat[, 31:40]), format = "latex",
205     booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
206 t15 <- kable(as.data.frame(res_acc_mat[, 41:50]), format = "latex",
207     booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
208 t16 <- kable(as.data.frame(res_acc_mat[, 51:59]), format = "latex",

```

```

209     booktabs = TRUE) %>% kable_styling(latex_options = "scale_down")
210 t11
211 t12
212 t13
213 t14
214 t15
215 t16
216 # plot of the accuracies
217 cols1 = c("grey73", "deeppink2", "peachpuff4")
218 plot(sample_grid, res_acc_mat[2, ], type = "l", ylim = c(0,
219     0.9), col = cols1[1], lwd = 2, main = paste("Smooth~Filter~Viterbi Accuracy \nfor n=",
220     sample_grid[1], ":", rev(sample_grid)[1]), panel.first = grid(25,
221     25), ylab = "Accuracy")
222 lines(sample_grid, res_acc_mat[1, ], col = cols1[2], lwd = 2)
223 lines(sample_grid, res_acc_mat[3, ], col = cols1[3], lwd = 2)
224 legend(x = 10, y = 0.18, legend = c("Smooth", "Filter",
225     "Viterbi"), col = cols1, lty = 1, lwd = 2)
226 sample_indx <- sample(1:50, 9)
227 sample_indx
228 par(mfrow = c(3, 3))
229 for (i in sample_indx) {
230     plot(accuracy_mat[, i]$entropy, type = "o", col = sample(colors(),
231     1), pch = 20, ylab = "Entropy", xlab = "nSample",
232     panel.first = grid(25, 25), main = paste("Entropy for nSampe=",
233     sample_grid[i], " \nfor 10 timesteps"))
234 }
235 # take a sample with 100 timesteps from 4
236 samp1 = rep_res[, 60]
237 plot(samp1$entropy, $filt_ent, main = "Entropy filtered and smoothed",
238     type = "l", col = "mediumspringgreen", ylab = "Entropy",
239     ylim = c(0, 2), panel.first = grid(25, 25))
240 lines(samp1$entropy, $smooth_ent, col = "mediumorchid1")
241 legend(x = "topright", lty = 1, legend = c("Smoothed", "Filtered"),
242     col = c("mediumspringgreen", "mediumorchid1"))
243 # -----
244 # Question 7
245 # -----
246 rndSample = rep_res[1, sample(1:dim(rep_res)[2], 1)]
247 post = posterior(hmm, rndSample[[1]]$observation)
248 prob_step101 = transProbs %*% post[, 100]
249 df = data.frame(prob_step101)
250 rownames(df) = States
251 knitr::kable(df)
252 # -----
253 # Other
254 # -----
255 par(mfrow = c(5, 5))
256 for (i in 5:14) {
257     plot.ts(f_tab[, i])
258     pch1 = as.character(1:10)
259     points(f_tab[, i], pch = pch1, cex = 1.25, font = 4,
260         col = 1:10)
261     legend("topleft", legend = States, col = 1:10, pch = pch1)

```

