Reinforcement Learning Assignment 1

Soumyodeep Dey

26th August 2024

1 Sample Trajectories and Observations

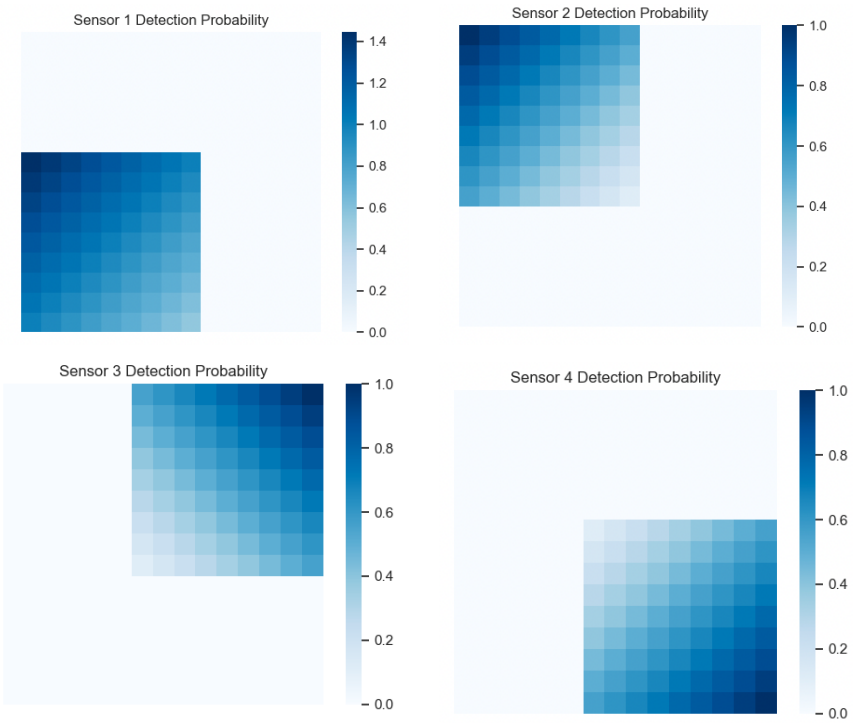


Figure 1: Sensor Detection Probabilities Heatmaps

Following are the 20 plotted trajectories with red being the start state and green being the end state.

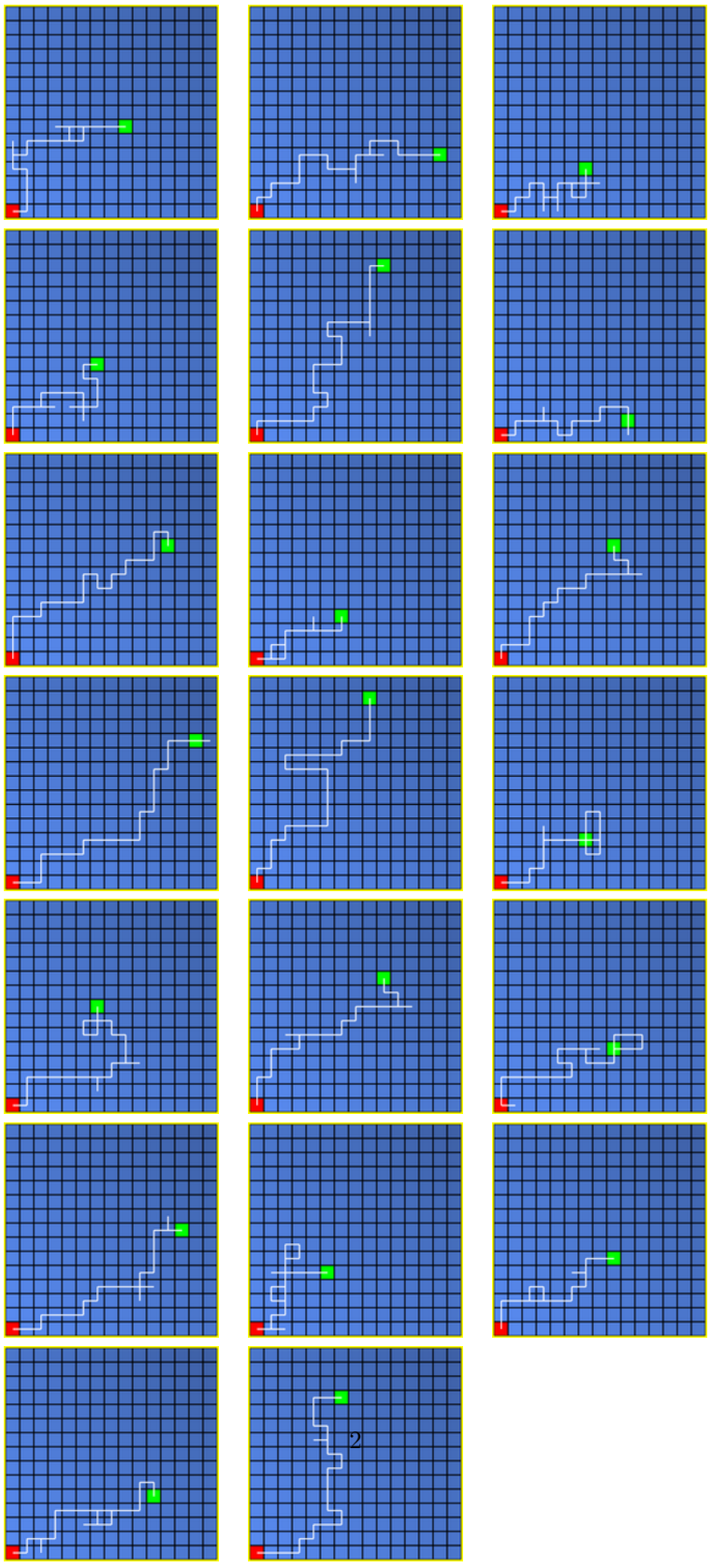


Figure 2: Plotted Trajectories

2 Likelihood Estimation

The following are the likelihood values that I am getting after implementing the Forward algorithm on the given observation sequence along with the hidden states.

```
Observation Likelihood for trajectory 1: 4.408736923098082e-17
Observation Likelihood for trajectory 2: 2.548511045460333e-09
Observation Likelihood for trajectory 3: 8.646043455800283e-13
Observation Likelihood for trajectory 4: 1.088682372419606e-06
Observation Likelihood for trajectory 5: 1.1472125349901736e-13
Observation Likelihood for trajectory 6: 6.154505972155134e-14
Observation Likelihood for trajectory 7: 2.9119266102546677e-12
Observation Likelihood for trajectory 8: 5.0538844038763685e-08
Observation Likelihood for trajectory 9: 5.397031896092145e-07
Observation Likelihood for trajectory 10: 6.717314426744151e-06
Observation Likelihood for trajectory 11: 1.378403733912653e-10
Observation Likelihood for trajectory 12: 3.106257634927332e-12
Observation Likelihood for trajectory 13: 1.1158728570155114e-09
Observation Likelihood for trajectory 14: 1.6219564644453562e-11
Observation Likelihood for trajectory 15: 6.332481837876784e-08
Observation Likelihood for trajectory 16: 1.575199141878402e-09
Observation Likelihood for trajectory 17: 5.6062314858740645e-11
Observation Likelihood for trajectory 18: 1.1754020070290825e-14
Observation Likelihood for trajectory 19: 1.3208466523662103e-11
Observation Likelihood for trajectory 20: 3.400736279908848e-08
```

The forward algorithm computes the joint probability of the sequence of states and observations. Due to the multiplication of many small probabilities (especially in long sequences, here 30), the resulting likelihood is very small, approaching 0. However, it can never be negative or greater than 1. Small likelihood values (approaching 0) is more common due to the multiplication of many small probabilities over a sequence of states and observations. The likelihood of a trajectory in an HMM decreases with increasing trajectory length due to the multiplication of probabilities less than 1 over a longer sequence of states and observations.

3 Decoding

Original trajectories and the decoded trajectories using the Viterbi algorithm are plotted here. The original trajectories are in white and the decoded trajectories are in pink colour.

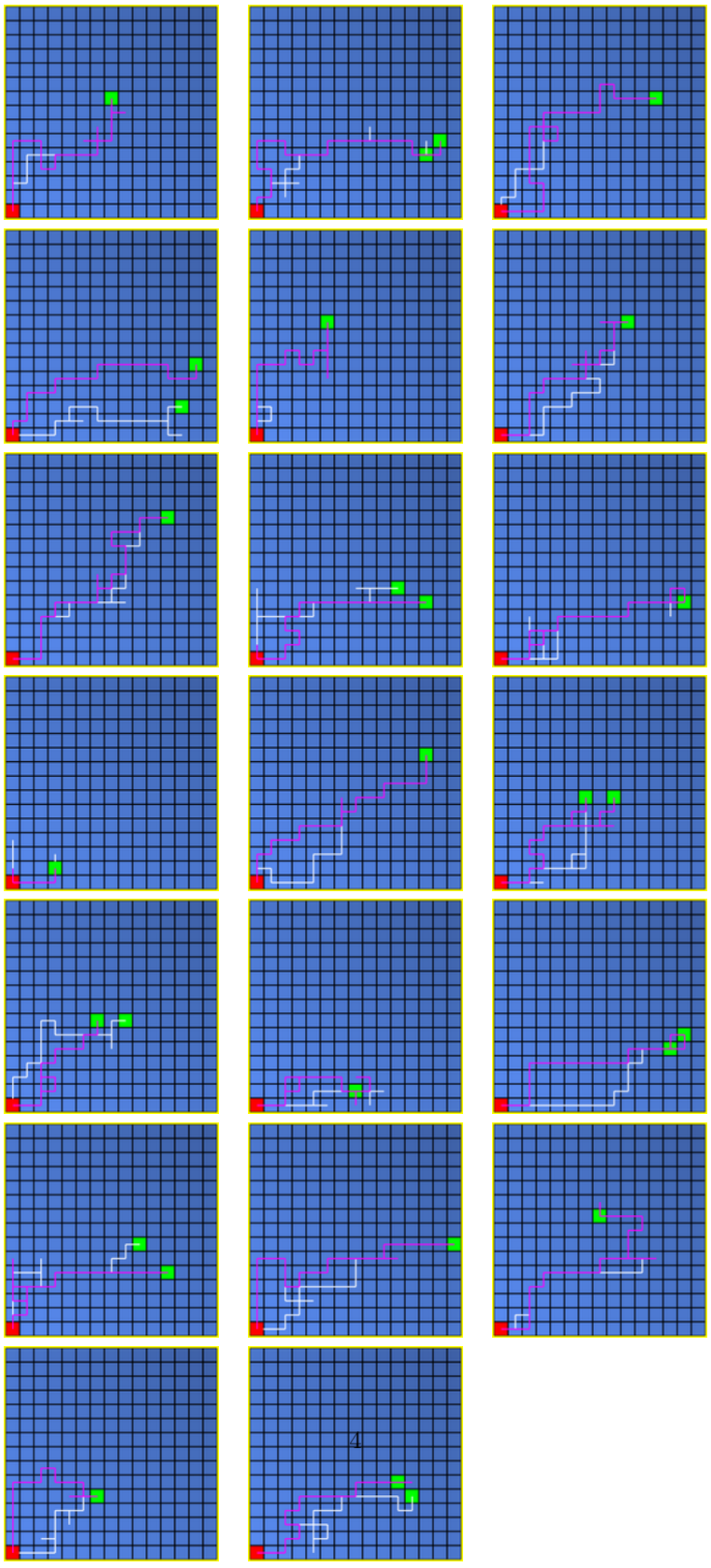
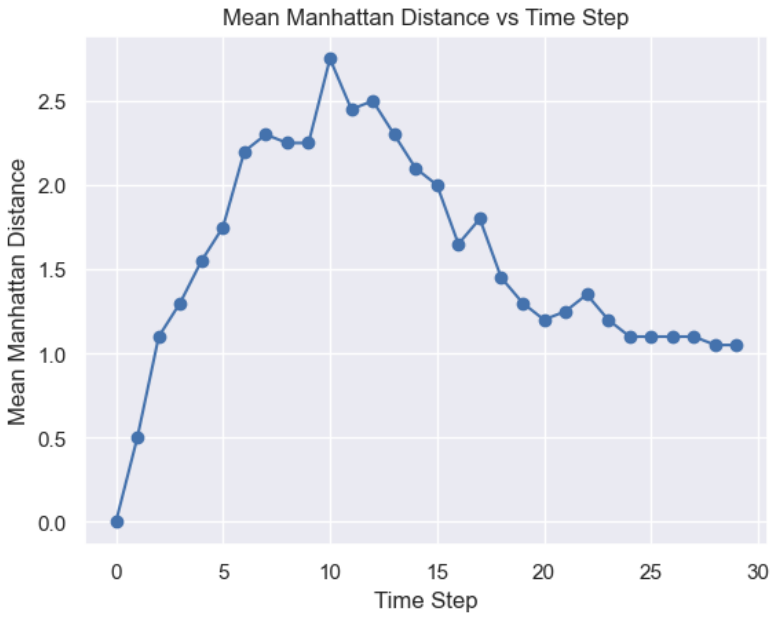


Figure 3: Original and Decoded trajectories



4 Baum Welch

I have done the first two parts of the Baum Welch algorithm. I have used the same code for both the parts except in the first part I have simply not updated the observation probability matrix after the M-step. I could observe one thing in both the parts that the KL Divergence Loss value is not changing over all the iterations. Whatever KLD value was calculated in the first iteration remains the same till the last iteration. One reason could be that the number of observation sequences given (10000) is not sufficient for the Baum Welch algorithm to update the transition probability matrix and the Observation Probability Matrix. Another reason could be that I have used the number of states as 225, ie the all the cells in the grid are valid states for me in all time steps. Thus, there may be a case of underflow since the transition matrix and the observation matrix are sparse matrices of sizes (225,225) and (16,225) respectively.

For more clarity on the implementation related details, please refer to the ipython notebook.