Small Problem 3: Discrete-time Discrete-observation HMM

Introduction

The file "problem-3-generator.R" contains R code to generate a simple discrete-time, discrete-observation HMM. The specific instance for this small problem has 5 states and 5 observations and sequences of length 20. The transition distribution has probability $\frac{1}{3}$ of staying in the same state s, $\frac{1}{3}$ of transitioning to state $s+1$ and $\frac{1}{3}$ of transitioning to state $s+2$. The starting state is $s=1$, and the states wrap around mod 5.

The observation for state s is equal to s with probability 0.6 and is equal to 0.1 for each of the other observation values.

Name	*Description*
problem-3-true-state.csv	The true state of the MDP for each time step 1:20
problem-3-outputs.csv	The sequence of observations 1:20

Note that this problem involves no learning. The task is just to perform probabilistic inference with a given model and data.

Queries and Metrics

Query 1: The MAP state sequence.

Metric: Minimum Hamming distance between the predicted and the true MAP state sequences.

Query 2: Smoothing: For each time step $t > 1$, the marginal distribution $P(s_t|O)$, where O is the entire output sequence.

Metric: Total variation distance between the true and computed marginal posteriors at each time step.

Query 3: Filtering: For each time step $t > 1$, compute the marginal distribution $P(s_t|O_{1:t})$, where $O_{1:t}$ is the vector of outputs from time 1 up to the current time t.

Metrics: Same as for Query 2.

Please note that the true posterior distributions are provided for both queries 2 and 3. Please reference the provided solution for problem 3, e.g. `ppaml-cp4/solutions/problem3`.

Submission

The metric value should be computed for each elapsed time step (by calling the provided code or by implementing yourself). The metric value should be reported for several elapsed time steps. The number of elapsed time steps should be sufficient to establish an "informative profile".

For further details regarding submission of the metric and your code, please refer to the main CP4 problem description document, e.g. PPAML-Challenge-Problem-4.pdf.

Sample output for this problem has been provided in the “sampleoutput” folder:

`problem-3-query-1-metric-1.csv`

`problem-3-query-2-metric-2.csv`

`problem-3-query-3-metric-2.csv`

Notes

Further details on this problem can be found in the provided sample solution, e.g.

`ppaml-cp4/solutions/problem3`