Small Problem 3: Discrete-time Discrete-observation HMM

Queries:

Query 1: The MAP state sequence.

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

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:

Metric 1 (for query 1 only):
Minimum Hamming distance between the predicted and the given MAP state sequences. Notice that there are 18 true MAP state sequences provided in this solution.

Metric 2 (for queries 2 and 3):

a) Total variation distance between the true and computed marginal posteriors at each time step.
b) The mean and variance of the per-step total variation distance computed across all time steps.

Ground Truth:

The answers to the queries are in the attached spreadsheet (`problem-3-solution.xlsx`). The Excel format file is represented by three CSV files, one for each query, i.e. map, smoothing, and filtering.

TODO:

Compute Metric 1 for Query 1.
Compute Metrics 2a and 2b for Queries 2 and 3. The provided Matlab evaluation code uses the following command line:

```
matlab -nospash -nojvm -nodisplay -nodesktop -r
"TVDScoreHMM(<number-of-states>, <number-of-timesteps>,
<ground-truth-path>,<input-path>,<output-path>)"
```

The output will be a CSV file written to `<output-path>`.

Note that the Matlab program for Metric 1 is compatible with GNU Octave.

Submit the metric and your code as described in the main CP4 problem description document, e.g. PPAML Challenge Problem 4-v7.pdf.

Ground Truth Details:

This discrete-time Discrete-observation HMM problem was solved using Kevin Murphy's Bayes Net Toolbox (BNT) for MATLAB (Murphy, 2001).