Computational Intelligence SEW SS17
Homework 6
Hidden Markov Models and Markov Models

Christian Knoll

Tutor:	Philipp Gabler, pgabler@student.tugraz.at
Points to achieve:	15 pts
Extra points:	4* pts
Info hour:	03.07.2017, 11:00-12:00, HS i12
Deadline:	05.07.2017 23:59
Hand-in mode:	Submit your **python files and a colored version of your report (as.pdf)** at `https://courses-igi.tugraz.at/`. (Please name your archive *hw6-Familyname1Familyname2Familyname3.zip*)
Hand-in instructions:	`https://www.spsc.tugraz.at/courses/computational-intelligence-cs`
Newsgroup:	tu-graz.lv.ew

General remarks

Your report must be self-contained and must therefore include all relevant plots, results, and discussions. Your submission will be graded based on:

- The correctness of your results (Is your code doing what it should be doing? Are your plots consistent with what algorithm XY should produce for the given task? Is your derivation of formula XY correct?)
- The depth and correctness of your interpretations (Keep your interpretations as short as possible, but as long as necessary to convey your ideas)
- The quality of your plots (Is everything important clearly visible in the print-out, are axes labeled, . . . ?)

1 Hidden Markov Models

General: In this homework, we want to compute the optimal state sequence of given Hidden Markov Models (HMMs) for sequences of observations. Further, we will use HMMs to classify sequential data. We will stick to our very simple weather model from the problem classes. Remember that we defined the state $q_n \in \mathcal{Q}$ as the weather on the n^{th} day, and the set of possible states as $\mathcal{Q} = \{ , , \}$. You can use the provided script `skeleton_hw6.py`.

1.1 Viterbi Algorithm and optimal State Sequence [7 Points]

We want to compute the most likely sequence of hidden states given an observation sequence. To this end, there is a provided class `HMM` with the following fields:

```
#   A ... Transition Probabilities
#   B ... Emission Probabilities
#   pi ...Start Probabilities
#   N_s... Number of States
#   N_o... Number of possible observations
```

Write a method `q_opt = HMM.viterbi_discrete(X)` for this class that implements the Viterbi algorithm for HMMs with discrete emission probabilities and computes the optimal state sequence `q_opt` for a given observation sequence `X`.

1. To evaluate your implementation, use the observation sequences $X_1 = \{ , , , , , \}$ and $X_2 = \{ , , , , , , \}$. For the HMM, try the one from the problem classes:

$$\text{HMM}_1: \quad \boldsymbol{A}_1 = \begin{bmatrix} 0.8 & 0.05 & 0.15 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}, \quad \boldsymbol{B}_1 = \begin{bmatrix} 0.1 & 0.9 \\ 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}, \quad \boldsymbol{\pi}_1 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$$

   Also try another HMM with the following parameters:

$$\text{HMM}_2: \quad \boldsymbol{A}_2 = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.05 & 0.7 & 0.25 \\ 0.05 & 0.6 & 0.35 \end{bmatrix}, \quad \boldsymbol{B}_2 = \begin{bmatrix} 0.3 & 0.7 \\ 0.95 & 0.05 \\ 0.5 & 0.5 \end{bmatrix}, \quad \boldsymbol{\pi}_2 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$$

2. What are the optimal state sequences for the two HMMs? Explain your results!

3. The computations in the iteration phase of the Viterbi algorithm include a lot of multiplications of probabilities. Which problem can arise here and how can you prevent it? If you have not already done so, modify your implementation accordingly!

1.2 Sequence Classification [4 Points]

HMMs can be used to classify sequential data, where different classes are modeled using HMMs. In our example, we want to find out which of the two weather models from above is more likely to have caused the observation sequences X_1 and X_2. Hence, given HMMs specified by parameters $\boldsymbol{\Theta}_i$, we need to compute $\mathrm{P}(X|\boldsymbol{\Theta}_i)$ for all i and choose the largest.

1. Write down the formula for $\mathrm{P}(X|\boldsymbol{\Theta}_i)$ and state why it is computationally expensive to solve it.

2. Rewrite the formula using the product rule! Using the result, can you think of a way that the Viterbi algorithm could help in calculating at least an approximate classification result?

3. Perform this approximate classification and comment on the results!

 Tutor:Philipp Gabler, pgabler@student.tugraz.at

1.3 Samples from an HMM [2 Points]

1. Write a method `Q,X = HMM.sample(N)` that draws state sequences Q and according observation sequences X of length N from HMMs with discrete observations. You can use the function `Y = sample_discrete_pmf(X, PM, N)` that you know already from homework 5. Then, sample from the HMMs specified as above!

2. In homework 5, you were asked to sample from a GMM. Can you see any connections to this task?

2 Markov Model

Now consider a Markov Model with the same properties as HMM_1, except that the states can be directly observed. We use the model from the problem class specified by the following transition- and prior-probabilities

$$\boldsymbol{A} = \begin{bmatrix} 0.8 & 0.05 & 0.15 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}, \quad \boldsymbol{\pi}_1 = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}$$

2.1 Graphical Representation [2 Points]

1. Draw the graphical representation of this Markov model.

2.2 Weather Prediction [4* Points]

The probabilities for all states on day n are collected in the vector $P_n =$ [$P(q_n =$ $)$; $P(q_n =$ $)$; $P(q_n =$ $)$].

1. Compute the state-probabilities for day 2 and 3, i.e., P_2 and P_3.

2. Can you think of a way to compute the state-probabilities in the infinite future; if so explain your aproach and compute all values of P_∞.

3. Assume the prior-probabilities are given by $\boldsymbol{\pi}_2 = \begin{bmatrix} 0.9 & 0.05 & 0.05 \end{bmatrix}$. How would this influence P_∞. Explain your answer!