CMPUT 466/551 — Programming Exercise #7

Instructor: R Greiner
Due Date: 5pm, Wednesday 30/Nov/2016

This exercise is intended to further your understanding of Hidden Markov Models.

Relevant reading: Rabiner: "A Tutorial on Hidden Markov Models ..."
+ [HTF: parts of Ch 17]
Total points: 20
Percentage of total course mark: Ugrad: 8% Grad: 6%
Note: We will impose a 20% deduction for any question if you do not follow the input/output format specified in that question.

You are playing a dice game in the *Hogwart's Casino*, against Malfoy. Unfortunately, Malfoy cheats... when he can, he jinxes the single dice to change from being fair (with distribution f) to rigged (r), based on the distributions

v	$P(\, v \mid D = f\,)$	$P(\, v \mid D = r\,)$
1	1/6	0.80
2	1/6	0.04
3	1/6	0.04
4	1/6	0.04
5	1/6	0.04
6	1/6	0.04

(1)

where in general the variable D (later D_t) refers to the current state of the die, which is a hidden variable. As he needs to cast the jinx *undetected*, which is difficult, he only does this occasionally — only 20% of the time. Fortunately for fair play, Dobby can counterjinx it back (from r to f). However, he also does not want to be caught, and so can only change the die back 10% of the time. So we have

$$\begin{aligned} P(\, D_{t+1} = r \mid D_t = f\,) &= 0.20 \\ P(\, D_{t+1} = f \mid D_t = f\,) &= 0.80 \\ P(\, D_{t+1} = f \mid D_t = r\,) &= 0.10 \\ P(\, D_{t+1} = r \mid D_t = r\,) &= 0.90 \end{aligned} \qquad (2)$$

Now imagine you observe the sequence of die rolls:

$$\vec{O} \quad = \quad [4, 1, 2, 3, 1, 3, 1, 1, 5, 6]$$

Let $\vec{O}_{i:j}$ be the subsequence between roll#i and roll#j (inclusive), so $\vec{O}_{3:5} = [2, 3, 1]$ and $\vec{O}_{1:10} = \vec{O}$. You may assume the $\{f, r\}$ are the only two possible states of the die, and that the initial "jinx-state" of the die is $50/50 - P(\, D_0 = r\,) = 0.5$.

In `PE7-code.zip`, we have provided three skeleton MATLAB functions: `forward.m`, `backward.m`, `viterbi.m` and one top-level script `PE7.m` that calls these functions. Also, we will let k refer to the number of states and m refer to the number of observed values – here $k = |\{r,\ f\}| = 2$ and $m = |\{1, 2, 3, 4, 5, 6\}| = 6$.

a [1]: Note that the initial state distribution(called $P(\,D_0\,)$ above) is [0.5 0.5]. Write the appropriate code in `PE7.m` to predict state distribution before evidence – *i.e.*, predict $P(\,D_1\,)$.

b [5]: The `forward.m` function takes a sequence of observations $\vec{O}_{1:T} = [O_1,\ \ldots,\ O_T]$, as well as a description of the HMM: `phi` = the state distribution before evidence (called $P(\,D_1\,)$ above); `A` = HMM transition matrix (of size $k \times k$); and `B` = HMM emission matrix (of size $m \times k$).

You should complete the function `forward.m`, so it returns the `alpha` matrix (of size $T \times k$) and the probabilities of the observation sequence, where $\alpha_t(i) = P(\ \vec{O}_{1:t},\ D_t = d_i\,)$ is probability of the partial observation sequence $\vec{O}_{1:t}$ and being in state d_i (where $d_i \in \{f, r\}$) at time `t`. (See Equations 18 and 21 of [Rabiner 1989].)

Write appropriate code in `PE7.m` to compute $P(\,D_t = r \,|\, \vec{O}_{1:t}\,)$ for $t = 1..10$ and report the results.

[Hint: You may need the variables returned by `forward.m` *.]*

c [5]: The `backward.m` function takes an observation sequence $\vec{O}_{1:T}$ as well as a description of the HMM. You should complete the `backward.m` function so it returns the `beta` matrix (of size $T \times k$), where $\beta_t(i) = P(\,\vec{O}_{t+1:T} \,|\, D_t = d_i\,)$ is probability of observation sequence from `t+1` to `T`, $\vec{O}_{t+1:T} = [O_{t+1},\ \ldots,\ O_T]$, given the state at time `t` is d_i. (See Equation 23 of [Rabiner 1989].)

Fill-in the appropriate part in `PE7.m` that computes $P(\,D_t = r \,|\, \vec{O}_{1:10}\,)$ for $t = 1..10$. Report the results. (See Equation 26 of [Rabiner 1989].)

[Hint: As a sanity check, notice we also computed $P(\,D_{10} = r \,|\, \vec{O}_{1:10}\,)$ in (b)...]

d [5]: The `viterbi.m` function takes an observation sequence $\vec{O}_{1:T}$, as well as a description of the HMM. Now, you should complete the `viterbi.m` function so it returns the most likely interpretation of observed sequence, $\vec{q}^{\,*} = \mathrm{argmax}_{\vec{d}}\, P(\,\vec{D}_{1:10} = \vec{d} \,|\, \vec{O}_{1:10}\,)$ and the probability of this interpretation. (See Equations 32a-35 of [Rabiner 1989].)

Report this most likely interpretation $\vec{d}^{\,*}$. Write the code in `PE7.m` that prints the results, and also report the results.

e [4]: If Lucius is watching, Malfoy is more likely to try to show-off and Dobby is less able to change the die back, meaning the transition probabilities will be:

$$\begin{aligned} P_{+L}(\,D_{t+1} = r \,|\, D_t = f\,) &= 0.25 \\ P_{+L}(\,D_{t+1} = f \,|\, D_t = f\,) &= 0.75 \\ P_{+L}(\,D_{t+1} = r \,|\, D_t = r\,) &= 0.95 \\ P_{+L}(\,D_{t+1} = f \,|\, D_t = r\,) &= 0.05 \end{aligned} \tag{3}$$

which is different from the $P(\,D_{t+1} \,|\, D_t\,)$ values presented in Equation 2. (You should view those values as a shorthand for $P_{-L}(\,D_{t+1} \,|\, D_t\,)$.) Note that Equation 1 is unaffected.

Do you think Lucius is watching? That is, is $P_{+L}(\,\vec{O}\,) > P_{-L}(\,\vec{O}\,)$? (Btw, you should NOT be surprised if these numbers are very small.)

Modify the appropriate portion in `PE7.m` to see the results. You should also report the probability values.

You should turn in your well documented `forward.m`, `backward.m`, `viterbi.m` and `PE7.m` files in a single folder. You may not change the arguments of the functions. The folder should also include the report, as `.pdf` or `.txt` file. Please do not submit `.doc` or `.docx` files.

Submit a single `zip` file with all of your deliverables. Do not submit `7z` or `rar` files.