Assignment4 – Markov Models

Given: May 22   Due: May 27

Problem 4.1 (The Value of Information)

Chef Giordana runs a kitchen that provides food for a large organisation. A salad is sold for €6 and costs €4 to prepare. So a sold salad is profit of €2, and an unsold salad a loss of €4. Actual demand will be 40 (with probability 0.5) or 60 (also with probability 0.5) each day.

Each day, Giordana must decide in advance between two options: prepare 40 or 60 salads.

1. In the absence of additional information, compute the expected utility of each decision option, and choose the best option.
2. She is considering a new ordering system, where she knows the demand perfectly in advance. So she can always choose the better of the two options. State the formula for computing the value of this perfect information, explain the components in Giordana's case, and compute the value.

Solution: Giordana's payoff table looks as follows:

Demand	Probability	40 salads	60 salads
40	0.5	€80	€0
60	0.5	€80	€120

1. Thus, the expected utility for making 40 salads is 80 and the expected utility for making 60 salads is 60. Based on these expected values without additional information, Giordana would choose to make 40 salads per day with an *expected utility* of €80 per day.
2. The value of information is equal to the expected value of best action given the information minus expected value of best action without information. Consider a random variable X with possible values $x_1, \dots, x_n$. Let A_k be the optimal action to take if $X = x_k$, and let B be the optimal action to take in the absence of information on X. The general formula for the value of perfect information on variable X given evidence E is

$$VPI_E(X) = \sum_k P(X = x_k \mid E) \cdot \mathrm{EU}(A_k|E, X = x_k) - \mathrm{EU}(B|E)$$

In Giordana's case, X is the demand, i.e., $x_1 = 40$ and $x_2 = 60$, and E is all background information already known. B is to prepare 40 salads, and $\mathrm{EU}(B|E) = 80$. A_1 and A_2 are to prepare 40 and 60 salads, respectively, and

$$\sum_k P(X = x_k \mid E) \cdot \mathrm{EU}(A_k|E, X = x_k) = 0.5 \cdot 80 + 0.5 \cdot 120 = 100$$

So $VPI_E(X) = 100 - 80 = 20$. So Giordana should pay at most €20 for the perfect information.

Problem 4.2 (Expected Utility)

1. State the formal definition of *expected utility* of an action in the current state of an agent? Explain the meaning of every variable in the defining equation.

Solution: The *expected utility* $(\cdot|\cdot)$ is defined as

$$\mathrm{EU}(a|e) = \sum_{s'} P(R(a) = s' \mid a, e) \cdot U(s')$$

where

1. a is the action for which we want to find out the expected utility, given the evidence e.
2. $U(s')$ is the utility of a state s'.
3. $R(a)$ is the result of the action a.

---

2. How do we use expected utility to make decisions?

Solution: The principle of maximum expected utility says that a rational agent should choose the action that *maximizes* the agent's expected utility.

Problem 4.3 (Decision Theory)

You are offered the following game: You pay x dollars to play. A fair coin is then tossed repeatedly until it comes up heads for the first time. Your payout is 2^n, where n is the number of tosses that occurred.

1. Assume your utility function is exactly the monetary value. How much should you, as a rational agent, be willing to pay to play? Use the formal definition of "expected utility" from the *lecture*.

Solution: We have

$$\begin{aligned}\mathrm{EU}(\mathsf{Play}) &= \sum_{s'} P(\mathsf{Payout} = s' \mid \mathsf{Play}) \cdot U(s') = \sum_{k\in\mathbb{N}^+} P(\mathsf{Payout} = k) \cdot 2^k \\ &= \sum_{k\in\mathbb{N}^+} \frac{1}{2^k} \cdot 2^k = \sum_{k\in\mathbb{N}^+} 1 \to \infty\end{aligned}$$

We should be willing to pay any amount for a chance to play.

---

2. Assume now, that your utility function for having k dollars is $U(k) = m\log_n k$ for some $m, n \in \mathbb{N}^+$. How does this change the result?

Solution:

$$\mathrm{EU}(\mathsf{Play}) = \sum_{k\in\mathbb{N}^+} \frac{1}{2^k} \cdot m\log_n(2^k) = m\log_n(2) \sum_{k\in\mathbb{N}^+} \frac{k}{2^k} = 2m\log_n(2)$$

Of course, this is wrong insofar as the utility, being logarithmic, is in particular not linear, i.e. the actual utility depends on our original capital K, but this just makes everything more complicated. The point is that using a logarithmic utility for money yields a finite result.

---

3. What is wrong with the result from the first exercise? Which implicit assumption leads to the apparently nonsensical result? How could it be fixed?

Solution: This model assumes that the payout is potentially infinite (which is unrealistic), as well as that we have an unlimited amount of money at our disposal. One way to fix this is to calculate our overall utility as the difference of $\mathrm{EU}(\mathsf{Play})$ and the utility $U(k)$ of paying the cost k. Then we can set an upper limit of how much we can afford to pay by setting $U(k) = -\infty$ when k is greater than the amount in our bank account.

Hint: The series $\sum_{k=1}^{\infty} \frac{k}{2^k}$ is convergent with limit 2.

Problem 4.4 (Markov Mood Detection)

On any given day d, your roommate Moody is either happy or sad, so $M_d \in \{h, s\}$. Usually when he is sad, he stays sad for a while, and $P(M_{d+1} = s \mid M_d = s) = 0.7$. But aside from that he is a cheery guy, and $P(M_{d+1} = h \mid M_d = h) = 0.85$.

He either listens to jazz or metal music, so $L_d \in \{j, m\}$. On a happy day he usually listens to Jazz, and $P(L_d = j \mid M_d = h) = 0.7$. On a sad day, he slightly prefers metal, and $P(L_d = m \mid M_d = s) = 0.6$.

1. Model this situation as a Markov process. Explain what the state and evidence variables are. What order does the process have? Is the transition model stationary? Is the sensor model stationary? Does the transition model have the Markov property? Does the sensor model have the sensor Markov property? Explain all answers in one short sentence each.

Solution: The state variables are the M_d. The evidence variables are the L_d.

The process is first-order: $P(M_d \mid M_{0:d-1}) = P(M_d \mid M_{d-1})$ and thus has the Markov property (and thus is a Markov process). The transition model is stationary: $P(M_d \mid M_{d-1})$ does not depend on d. The sensor Markov property holds: $P(L_d \mid M_{0:d-1}, L_{1:d-1}) = P(L_d \mid M_d)$. The sensor model is stationary: $P(L_d \mid M_d)$ does not depend on d.

---

2. State the formula for the full joint probability distribution. You know that he was happy at day d_0. What is the probability that he is happy and plays Jazz for the next two days?

Solution: The distribution is

$$P(M_{0:n}, L_{1:n}) = P(M_0) \cdot \prod_{i=1}^{n} P(M_i \mid M_{i-1}) \cdot P(L_i \mid M_i)$$

In our case, $n = 2$, $P(M_0 = h) = 1$, $P(M_i = h \mid M_{i-1} = h) = 0.85$ and $P(L_i = j \mid M_i = h) = 0.7$. So $P(M_0 = h, M_1 = h, M_2 = h, L_1 = j, L_2 = j) = 0.85^2 * 0.7^2$.

Problem 4.5 (Moody HMM)

Consider the Markov process from before about the roommate Moody (which in particular gives the concrete probabilities needed below). We have already modeled it as an HMM with state variables M_d and evidence variables L_d.

Because the transition model is first-order and stationary, we can collect the conditional probabilities for the state transitions into a matrix $T_{ij} = P(M_d = x_j \mid M_{d-1} = x_i)$ where x_i, x_j are two states (i.e., two possible values of the state variable). We use k for the number of states, and T is an $k \times k$ matrix.

Because the sensor model is stationary and has the sensor Markov property, we can collect the conditional probabilities for the observations into a matrix $S_{ij} = P(L_d = y_j \mid M_d = x_i)$ where x_i is a state and the y_j are the possible observations. If there are l possible observations, this is an $k \times l$ matrix. For a fixed observation e, the diagonal $k \times k$ matrices O_e from the *lecture notes* are obtained from the columns of this matrix.

Clarify the modeling as an HMM. Concretely:

1. What is k? Give the transition matrix T.

Solution: $k = 2$ and we use the ordering $[h, s]$ for the states. Then $T = \begin{pmatrix} 0.85 & 0.15 \\ 0.3 & 0.7 \end{pmatrix}$. For example, $T_{12} = T_{hs} = P(M_d = s \mid M_{d-1} = h) = 0.15$.

---

2. What is l? Give the sensor matrix S.

Solution: $l = 2$ and we use the ordering $[j, m]$. Then $S = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}$. For example, $S_{12} = S_{hm} = P(L_d = m \mid M_d = h) = 0.3$.

Note that you need to choose and state orderings for the states and observations so that it is clear which state/observation corresponds to which row/column of T and S.

Now consider a fixed sequence $L_1 = e_1, L_2 = e_2$ of observations that we have made for two days. Concretely, you heard Moody play metal on day $d = 1$ and jazz on day $d = 2$.

3. Give the diagonal sensor matrices O_1 and O_2 corresponding to the observation at $d = 1$ and $d = 2$.

Solution: We have $e_1 = m$ and $e_2 = j$.
We have $O_1 = O_m = \begin{pmatrix} O_{hm} & 0 \\ 0 & O_{sm} \end{pmatrix} = \begin{pmatrix} 0.3 & 0 \\ 0 & 0.6 \end{pmatrix}$ and $O_2 = O_j = \begin{pmatrix} O_{hj} & 0 \\ 0 & O_{sj} \end{pmatrix} = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.4 \end{pmatrix}$.

---

4. You are not sure what kind of mood your flatmate was in on day $d = 0$, but it was either good or bad with equal probability. The HMM algorithm for filtering and smoothing uses compact matrix/vector equation to compute f and b. Use those equation to determine the probability distribution of Moody's mood on day $d = 1$.
Note that T^T in the filtering equation in the *lecture notes* denotes the transpose of T.
Grading will focus on writing out the matrices with the correct probabilities in them and on formally stating the computations that need to applied to those matrices. You should also actually do those computations, but that is secondary.

Solution: We need to apply smoothing at $k = 1$. The general equation for smoothing is $P(M_1 \mid L_{1:2} = e_{1:2}) = \alpha \cdot (f_{1:1} \cdot b_{2:2})$. And the HMM matrix equation for f give us $f_{1:1} = \alpha \cdot (O_1 T^t f_{1:0})$ We have the prior probabilities $P(M_0) = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$, and we use that for the starting value $f_{1:0}$ of the forward iteration. The HMM matrix equation for b gives us $b_{2:2} = T O_2 b_{3:2}$. As the starting value $b_{3:2}$ of the backward iteration, we use the vector containing all 1s, i.e., $b_{3:2} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Note that, for simplicity, this example is chosen such that only one iteration step each is needed for f and b — for the general case we would need to apply the matrix equations multiple times to iterate towards the needed value.

We omit the concrete calculation.

Problem 4.6 (Prediction, Filtering, Smoothing)

Consider an HMM consisting of a stationary Markov process and a stationary sensor model. We restrict attention to HMMs with a single state-variable X_t and a single evidence variable E_t. (The sensor model does not necessarily have the Markov property.)

Explain the results of the *prediction*, *filtering*, *smoothing* algorithms. For each one, state the motivation, the expression for the conditional probability that is to be computed, and explain the components of the formula. You do not have to explain how the algorithms work.

Solution: In all three cases, we have observed the process from time 1 to time t, i.e., we have observed the events $E_i = e_i$ for $i = 1, \ldots, t$. Let $E_{1:t} = e_{1:t}$ abbreviate the conjunction of these events.

We use those observations to compute different conditional probabilities about the state X_p of the HMM at time p.

Prediction Here we predict the future of the HMM, i.e., $p = t + k > t$. We need the conditional probability

$$P(X_{t+k} \mid E_{1:t} = e_{1:t})$$

Filtering Here we estimate the current state of the HMM, i.e., $p = t$. We need the conditional probability

$$P(X_t \mid E_{1:t} = e_{1:t})$$

Smoothing Here we estimate the past states of the HMM, i.e., $p = k < t$. We need the conditional probability

$$P(X_k \mid E_{1:t} = e_{1:t})$$
