ECE 194E HW #4

Geffen Cooper

June 2022

1 Problem 1

The state transition graph for the original markov chain is shown:

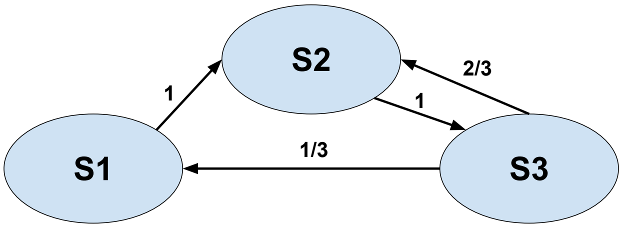

Figure 1: X_n: State Graph

Now we attempt to draw the state graph for the new markov chain which is a function of the original markov chain:

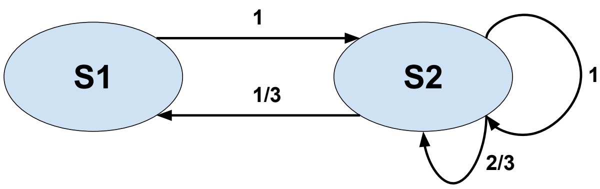

Figure 2: Y_n: State Graph

We see that there is some ambiguity when we are in state 2. Here the arrows indicate that once we enter state 2, the next state is always state 2. Then the next state can be state 2 with probability 2/3 or state 1 with probability 1/3. This is obviously not a markov chain since it relies on a prior state, i.e. we need to know if we just entered state 2 or have been in state 2 for one step in order to describe the transition probability of the next state. An example for a series of states is shown:

$$X_n : 1, 2, 3, 1, 2, 3, 2, 3$$

$$Y_n : 1, 2, 2, 1, 2, 2, 2, 2$$

In the sequence for Y_n, when we reach state 2, we cannot describe the probability of the next state with out using knowledge of the prior state. This violates the markov chain property.

2 Problem 2

First let's write the transition probability matrix. It is given that $P(1, 2) = P(2, 1) = \alpha$ and we can fill out the missing entries since the rows add up to one.

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}$$

It is also given that $Pr(y_n \neq x_n) = \epsilon$ and that $Pr(y_n = x_n) = 1-\epsilon$. Next, we can express this using the following notation:

$$Pr(Y_n = y_n | X_n = x_n) = Q(x_n, y_n)$$

so that

$$Q(1,1) = 1-\epsilon$$
$$Q(1,2) = \epsilon$$
$$Q(2,1) = \epsilon$$
$$Q(2,2) = 1-\epsilon$$

We need to find the MAP estimate of X^n given that $Y^n = y^n$ where the notation $X^n = (X_0, X_1, ..., X_n)$. The MAP estimate is then expressed as:

$$\underset{x^n}{argmax}\ P[X^n = x^n | Y^n = y^n]$$

Using Baye's rule and ignoring the denominator which is a constant we can rewrite as:

$$\underset{x^n}{argmax}\ P(y^n | x^n) P(x^n)$$

Using the $Q(x_n, y_n)$ notation from above we can write:

$$\underset{x^n}{argmax}\ Q(x_0, y_0), ..., Q(x_n, y_n) * \pi_0(x_0) P(x_0, x_1), ..., P(x_{n-1}, x_n)$$

Now taking the log we get:

$$\underset{x^n}{argmax}\ log\Big(\pi_0(x_0) Q(x_0, y_0)\Big) + \sum_{m=1}^{n} log\Big(P(x_{m-1}, x_m) Q(x_m, y_m)\Big)$$

Next we can introduce additional notation:

$$d_0(x_0) = -log\Big(\pi_0(x_0) Q(x_0, y_0)\Big)$$

$$d_m(x_{m-1}, x_m) = -log\Big(P(x_{m-1}, x_m) Q(x_m, y_m)\Big)$$

and write the MAP estimate equation as:

$$\underset{x^n}{argmin}\ d_0(x_0) + \sum_{m=1}^{n} d_m(x_{m-1}, x_m)$$

As shown in class, this can be treated as a shortest path problem where $d_0(x_0)$ is the length from the initial state to x_0 and $d_m(x_{m-1}, x_m)$ is the length from any future state x_{m-1} to x_m. We can use this approach to write the dynamic programming equations.

$m = 0, ..., n$ represents index of the states in the sequence and $x \in \{1, ..., N\}$ represents the possible states for a given index in the sequence. Now we define $V_m(x)$ to be the length of the shortest path from a state $x_m = x$ to the last state in the sequence. $V_n(x) = 0$ because the n^{th} state is the last state. We want to find $V_0(x)$ to get the shortest path (the most likely sequence of x^n given the observations y^n.)

We can define a general expression for $V_m(x)$ recursively as follows:

$$V_m(x) = \min_{x'}\Big(d_{m+1}(x, x') + V_{m+1}(x')\Big)$$

where the first term is the length from a state in index m (state x) to a state in index $m+1$ (state x') and the second term is the shortest length from the next destination state x'. We find the minimum of this sum across all the next destination states x' from a fixed source state x. This recursive definition is solved backwards because we start at the last state and then move back by one state until we reach the starting state. Overall we get the shortest path as:

$$V_0 = \min_{x}\Big(d_0(x) + V_0(x)\Big)$$

which is the result once we have recursively gone from the last state to the first state and calculated all the intermediate $V_m(x)$.

3 Problem 3

(a) At each step of the game we have two options: predict if the next card is an ace, or observe the next card. To maximize the expected gain we need to write expressions for the gain in these cases.

If we want to stop at any point, then the expected gain is the number of aces divided by the number of cards times the gain: $\frac{x}{m}(\$1) + \frac{m-x}{m}(\$0)$. If we observe the next card, then the expected gain can change in two ways depending on whether the next card is an ace. We express this recursively as

$$\frac{x}{m}V(m-1, x-1) + \frac{m-x}{m}V(m-1, x)$$

The first term is when the next card is an ace which happens with a probability of $\frac{x}{m}$ and the second term is when the next card is not an ace which happens with a probability of $1 - \frac{x}{m}$. The recursive terms are the expected gains after the next card has been shown.
Overall we want to maximize across these cases.

$$V(m, x) = \max_a \left(\frac{x}{m}, \frac{x}{m}V(m-1, x) + \frac{m-x}{m}V(m-1, x-1)\right)$$

where a is the action we take which is to stop or just observe the next card. Each term in the max() corresponds to one of two possible actions. The states are inherently defined by the number of aces and total cards left in the deck (as well as if the next card is an ace). The reward is 1 when the next card is an ace and the action was to stop.

(b) We can start by looking at the case of observing one card then saying stop on the next card. There are two cases which we take a weighted sum of to get the expected gain:

$$\frac{x-1}{m-1}\left(\frac{x}{m}\right) + \frac{x}{m-1}\left(\frac{m-x}{m}\right)$$

$$= \frac{x^2 - x + mx - x^2}{(m-1)(m)} = \frac{x(m-1)}{(m-1)m} = \frac{x}{m}$$

The left term is the case when the observed card is an ace and the gain we get if the stopped card is an ace. The right term is the case when the observed card is not an ace and the gain we get if the stopped card is an ace.

We see that even when we observe one card, the expected gain is still $\frac{x}{m}$. We now need to show that regardless of how many observations we make, the expected gain will stay the same so that the maximum expected gain is $V(m, x) = \frac{x}{m}$.

If we try to find the expression for expected gain when observing two cards we get four terms and this branching doubles for each additional card we observe. I could not figure out how to derive a general expression for this and show that it still equals $\frac{x}{m}$ beyond the base case shown above. However, the conclusion about the optimal strategy is that it does not matter when the player decides to say stop, the maximum expected gain will be the same. For example, observing 10 cards and then saying stop has the same expected gain as saying stop on the first card. Therefore, the player may as well say stop on the first card. This is similar to russian roulette, it doesn't matter what order a player is selected, they have the same probability of getting hit.