DS 122 Homework 8

Xiang Fu
xfu@bu.edu
Boston University Faculty of Computing & Data Sciences

Contents

1 Q1 Brand Loyalty Check

1.1 Q1.1

A brand asserts that 60% of their consumers are repeat customers. This prior belief is well represented by a Beta distribution with paramters $\alpha = 15$ and $\beta = 10$. A random sample of 40 customers is selected, and 28 are found to be returning customers. If we consider a beta distribution with parameters $\alpha = 15$ and $\beta = 10$ as the prior distribution for the probability of a customer being a repeat customer, what would the posterior distribution for this probability.

1.2 Solution

Given:
- Prior distribution: Beta distribution with parameters (alpha = 15) and (beta = 10)
- Sample data: Out of 40 customers, 28 are returning customers

The Beta distribution is updated with the data as follows:
- Posterior α = Prior α + Number of successes (returning customers)
- Posterior β = Prior β + Number of failures (non-returning customers)

In this case:
- Number of successes = 28
- Number of failures = 40 - 28 = 12

So, the posterior distribution parameters are:
- Posterior (alpha) = 15 + 28 = 43
- Posterior (beta) = 10 + 12 = 22

Therefore, the posterior distribution for the probability of a customer being a repeat customer is a Beta distribution with parameters (alpha = 43) and (beta = 22).

1.3 Q1.2

Compute the MAP (Maximum A Posteriori) and the MMSE (Minimum Mean Square Error) estimators for the probability.

1.4 Solution

Given the posterior distribution parameters α = 43 and β = 22.

$$\text{MPA} = \frac{43 - 1}{43 + 22 - 2} = \frac{42}{63} \approx 0.667$$

$$\text{MMSE} = \frac{43}{43 + 22} = \frac{43}{65} \approx 0.662$$

2 Q2 Biology

A biologist is studying the growth rate of a type of bacteria, which follows a Poisson distribution. Before any observations are made, the biologist's prior belief in the growth rate is that it increases by 5 colonies per day, though it may grow a lot faster, and the biologist is certain the growth rate is greater than 0. They therefore decide to model their prior beliefs in the growth rate as gamma distribution with parameters $\alpha = 5$ and $\beta = 1$. The biologist forgets to check on their experiment for a few days, but 3 days later they observe 20 new colonies. What is the posterior distribution of the growth rate?

2.1 Solution

Given:

- Prior distribution: Gamma distribution with parameters α = 5 and β = 1
- Observed data: 20 new colonies in 3 days

The posterior distribution's parameters are updated as follows:
- Posterior α = Prior α + Total number of events (new colonies)
- Posterior β = Prior β + Total time interval

In our case:

- Total number of events = 20
- Total time interval = 3 days

So, the posterior distribution parameters will be:
- Posterior α = 5 + 20 = 25
- Posterior β = 1 + 3 = 4

3 Q3 Email Campaign Effectiveness

Imagine that a marketing company runs two email campaigns: Campaign X and Campaign Y. Recipients are randomly assigned to receive one of the two campaigns when they sign up for the mailing list, and they continue to receive the same campaign for future mailings. The probability of a recipient being assigned to Campaign X at sign-up is 0.7. If a recipient is currently receiving Campaign X, the probability that they will request a switch to Campaign Y for their next email is 0.1; if a recipient is currently receiving Campaign Y, the probability that they will request to switch to Campaign X for their next email is 0.4.

3.1 Q3.1

What is the probability that a recipient who initially receives Campaign X will continue to receive Campaign X after 2 emails?

3.2 Solution

Consider the probability of staying with Campaign X after each email.

The probability of staying with Campaign X after the first email is simply 1, as they start with Campaign X.

The probability of staying with Campaign X after the second email involves not switching to Campaign Y.

Given that the probability of switching from Campaign X to Campaign Y is 0.1, the probability of staying with Campaign X is

$$1 - 0.1 = 0.9$$

Therefore, the probability that a recipient who initially receives Campaign X will continue to receive Campaign X after 2 emails is the product of the probabilities of staying with Campaign X after each email:

$$\text{Probability of staying with Campaign X after 2 emails} = 1 \times 0.9 = 0.9$$

3.3 Q3.2

What is the probability that a recipient initially receiving Campaign Y will still receive Campaign Y after 2 emails?

3.4 Solution

Consider the probability of staying with Campaign Y after each email.

The probability of staying with Campaign Y after the first email is 1, as they start with Campaign Y.

The probability of staying with Campaign Y after the second email involves not switching to Campaign X. Given that the probability of switching from Campaign Y to Campaign X is 0.4, the probability of staying with Campaign Y is

$$1 - 0.4 = 0.6$$

Therefore, the probability that a recipient initially receiving Campaign Y will still receive Campaign Y after 2 emails is the product of the probabilities of staying with Campaign Y after each email:

$$\text{Probability of staying with Campaign Y after 2 emails} = 1 \times 0.6 = 0.6$$

4 Q4 Transition matrices

4.1 Q4.1

For the diagram illustrating a Markov chain below, write out the transition matrix P. Then, if your current state is 80% sunny, 10% rainy, 10% cloudy, what will your state be in two days?

4.2 Solution

The rows of the matrix represent the current state, while the columns represent the next state.

Set up the transition matrix:

$$\begin{Bmatrix} \text{P(sunny} \mid \text{sunny)} & \text{P(rainy} \mid \text{sunny)} & \text{P(cloudy} \mid \text{sunny)} \\ \text{P(sunny} \mid \text{rainy)} & \text{P(rainy} \mid \text{rainy)} & \text{P(cloudy} \mid \text{rainy)} \\ \text{P(sunny} \mid \text{cloudy)} & \text{P(rainy} \mid \text{cloudy)} & \text{P(cloudy} \mid \text{cloudy)} \end{Bmatrix}$$

Given the transition probabilities from the diagram:

- From sunny to sunny: 0.7
- From sunny to rainy: 0.2
- From sunny to cloudy: 0.1
- From rainy to sunny: 0.2
- From rainy to rainy: 0.6
- From rainy to cloudy: 0.2
- From cloudy to sunny: 0.1
- From cloudy to rainy: 0.3
- From cloudy to cloudy: 0.6

So, the transition matrix P is:

$$\begin{Bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.6 & 0.1 \\ 0.1 & 0.3 & 0.6 \end{Bmatrix}$$

Given the current state vector So = [0.8, 0.1, 0.1], to find the state in two days, which is equivalent to applying the transition matrix twice, we calculate $S_2 = S_0 \cdot P \cdot P$.

$$S_1 = S_0 \cdot P$$

$$S_1 = \{0.8 \;\; 0.1 \;\; 0.1\} \cdot \begin{Bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.6 & 0.1 \\ 0.1 & 0.3 & 0.6 \end{Bmatrix}$$

$$S_1 = [0.8 \times 0.7 + 0.1 \times 0.2 + 0.1 \times 0.1, 0.8 \times 0.2 + 0.1 \times 0.6 + 0.1 \times 0.3, 0.8 \times 0.1 + 0.1 \times 0.2 + 0.1 \times 0.6]$$

$$S_1 = \{0.59 \;\; 0.25 \;\; 0.16\}$$

$$S_2 = S_1 \cdot P$$

$$S_2 = \{0.59 \;\; 0.25 \;\; 0.16\} \cdot \begin{Bmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.6 & 0.1 \\ 0.1 & 0.3 & 0.6 \end{Bmatrix}$$

$$S_2 = [0.59 \times 0.7 + 0.25 \times 0.2 + 0.16 \times 0.1, 0.59 \times 0.2 + 0.25 \times 0.6 + 0.16 \times 0.3, 0.59 \times 0.1 + 0.25 \times 0.2 + 0.16 \times 0.6]$$

$$S_2 = [0.479, 0.316, 0.205]$$

So, after two days, the state is approximately 47.9% sunny, 31.6% rainy, and 20.5% cloudy.

4.3 Q4.2

On any given day Gary is either cheerful, so-so, or glum. If he is cheerful today, then he will be so-so the next day 40% of the time and glum the next day 10% of the time. If he is feeling so-so today, then he will be cheerful or glum tomorrow each with a 30% probability. If he is glum today, then he will be cheerfull the next day 10% of the time and so-so 50% of the time. Write out the transition probability matrix for gary's mood.

4.4 Solution

Given the three states: cheerful, so-so, and glum. we can have its transition matrix as:

$$\begin{Bmatrix} \text{P(cheerful} \mid \text{cheerful)} & \text{P(so-so} \mid \text{cheeful)} & \text{P(glum} \mid \text{cheerful)} \\ \text{P(cheeful} \mid \text{so-so)} & \text{P(so-so} \mid \text{so-so)} & \text{P(glum} \mid \text{so-so)} \\ \text{P(cheerful} \mid \text{glum)} & \text{P(so-so} \mid \text{glum)} & \text{P(glum} \mid \text{glum)} \end{Bmatrix}$$

And also given in the question:

- If Gary is cheerful today (C), he can be so-so (S) with a probability of 40% and glum (G) with a probability of 10%. Therefore, the probability that he remains cheerful is the remaining percentage which is 50%.
- If he is so-so today, he has a 30% chance of being cheerful and a 30% chance of being glum tomorrow. Therefore, the probability that he remains so-so is the remaining percentage which is 40%.
- If he is glum today, he has a 10% chance of being cheerful and a 50% chance of being so-so tomorrow. Therefore, the probability that he remains glum is the remaining percentage which is 40%.

So the transition probability matrix P for Gary's mood is:

$$P = \begin{Bmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.5 & 0.4 \end{Bmatrix}$$

5 Q5 Identifying Primitive Markov Chains

For each of the matrices provided below, determine whether it is a primitive Markov chain. Provide a thorough explanation for your conclusion based on the properties of primitive Markov chain. In your explanations, consider the definition of a primitive Markov chain, which requires that the matrix represents a Markov chain, and that the matrix is irreducible and aperiodic. Remember to address these aspects in your analysis.

5.1 Q5.1 Matrix A

$$\begin{Bmatrix} 0 & 1 \\ 1 & 0 \end{Bmatrix}$$

5.2 Solution

Irreducibility

The matrix is irreducible because we can move from each state to the other in one step. Specifically, we can go from state 1 to state 2 and vice versa.

Aperiodicity

This matrix represents a two-state Markov chain where the only transitions are from one state to the other. This means that the period for each state is 2 because we can only return to the state we started from in an even number of steps (2, 4, 6, …). Therefore, the GCD is not 1, but 2.

So, this Markov chain is not aperiodic, it is not a primitive Markov chain.

5.3 Q5.2 Matrix B

$$\begin{Bmatrix} 0.7 & 0.5 \\ 0.3 & 0.5 \end{Bmatrix}$$

5.4 Solution

Irreducibility

The matrix appears to be irreducible because there are non-zero probabilities for transitioning from any state to any other state. This means it is possible to get from one state to the other in one step, and thus, over time, it is possible to reach any state from any state.

Aperiodicity

We can look at the cycle lengths possible for each state, and also for the chain to be aperiodic, the greatest common divisor (GCD) of the lengths of the cycles for each state must be 1. In this matrix, since there are non-zero probabilities for staying in the same state (0.7 for the first state and 0.5 for the second), it means that the chain does not require a fixed number of steps to return to the same state. It can return in one step, two steps, three steps, and so on. This flexibility in the number of steps to return to the same state indicates that the GCD is 1.

Since the matrix is irreducible and the GCD of the number of steps to return to a state is 1, making the chain aperiodic, we can conclude that this Markov chain is indeed primitive.

5.5 Q5.3 Matrix C

$$\begin{Bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{Bmatrix}$$

5.6 Solution

Irreducibility

The matrix is irreducible because it is possible to move from any state to any other state. Specifically, there is a probability of 0.5 of moving from state 1 to state 2, and a probability of 0.2 of moving from state 2 to state 1. This non-zero probability for transitions between states ensures that the Markov chain is irreducible.

Aperiodicity

For aperiodicity, we need to consider the probabilities of returning to the same state. State 1 has a self-transition probability of 0.5, and state 2 has a self-transition probability of 0.8. Since both states have non-zero probabilities of returning to themselves in one step, the chain does not require a fixed number of steps to return to the same state. It can return in one step, two steps, three steps, etc. This implies that the Markov chain is aperiodic because the number of steps to return to a given state does not have a common divisor greater than 1.

Given that the matrix is both irreducible and aperiodic, we can conclude that the Markov chain represented by this matrix is primitive.

6 Q6 Steady State and Detailed Balance

6.1 Q6.1

Given the transition matrix and state distribution below, determine if the state is a steady state. Show your work.

$$P = \begin{Bmatrix} 0.2 & 0.4 & 0.1 \\ 0.5 & 0.1 & 0.9 \\ 0.3 & 0.5 & 0 \end{Bmatrix}$$

State distribution:

$x = (0.2, 0.5, 0.3)$

6.2 Solution

Check if the state distribution remains unchanged after one transition.

Compute the product $x \cdot \text{mat}$ and compare it with x. If they are equal, then x is a steady state.

After performing the matrix multiplication $x \cdot \text{mat}$, we get the result:

$$x = (0.38, 0.28, 0.47)$$

Comparing this with the original state distribution $x = (0.2, 0.5, 0.3)$, we see that they are not equal. Therefore, the given state distribution x is not a steady state for the given transition matrix.

6.3 Q6.2

Given the transition matrix and a steady state distribution below, determine if the steady state statisfies the detailed balance condition. Show your work.

$$P = \begin{Bmatrix} 0.2 & 0.1 & 0.4 \\ 0.3 & 0.2 & 0.3 \\ 0.5 & 0.7 & 0.3 \end{Bmatrix}$$

State distribution:

$\pi = \left(\frac{35}{132}, \frac{3}{11}, \frac{61}{132}\right)$

6.4 Solution

Check for every pair of states i and j, the following condition holds:

$$\pi_i P_{\text{ij}} = \pi_j P_{\text{jj}}$$

We need to check the detailed balance condition for each pair of states. There are three states, so we will check the condition for the pairs (1,2), (1,3), and (2,3).

After checking the detailed balance condition for each pair of states, we find that none of the conditions are satisfied:

- For the pair (1,2): $\pi_1 P_{12} \neq \pi_2 P_{21}$
- For the pair (1,3): $\pi_1 P_{13} \neq \pi_3 P_{31}$
- For the pair (2,3): $\pi_2 P_{23} \neq \pi_3 P_{32}$

Since none of these conditions are met, the steady state distribution $\pi = \left(\frac{35}{132}, \frac{3}{11}, \frac{61}{132}\right)$ does not satisfy the detailed balance condition for the given transition matrix P.

7 Q7 Metropolis-Hastings

Consider sampling a posterior distribution using MCMC.

Assume the target posterior is the uniform distribution on $[-2, 2]$ and 0 elsewhere, i.e.

$p_T(x) = \frac{1}{4}$ for $-2 \leq x \leq 2$

$p_T(x) = 0$ for $x < -2, x > 2$

Assume the candidate distribution is uniform on $[-4, 4]$, i.e. $p_C(x) = \frac{1}{8}$ for $-4 \leq x \leq 4$.

Recall that the Metropolis-Hastings rule is to accept a proposed transition from j to i with probability:

$a_{\text{ij}} = \min\left(1, \frac{\pi H_{\text{ji}}}{\pi_j H_{\text{ij}}}\right)$

7.1 Q7.1

Consider a transition from $x = 3$ to $x = \frac{1}{2}$. What is the probability this transition will be accepted.

7.2 Solution

Given:

- The target distribution $p_T(x)$ is uniform on $[-2, 2]$ and zero elsewhere.
- The candidate distribution $p_C(x)$ is uniform on $[-4, 4]$.

For the transition from $x = 3$ to $x = \frac{1}{2}$:

- $\pi_j = p_T(3) = 0$ (since 3 is outside the range $[-2, 2]$).
- $\pi_i = p_T\left(\frac{1}{2}\right) = \frac{1}{4}$ (since $\frac{1}{2}$ is within the range $[-2, 2]$).
- $H_{\text{ij}} = p_C(3) = \frac{1}{8}$ (since 3 is within the range $[-4, 4]$).
- $H_{\text{ji}} = p_C\left(\frac{1}{2}\right) = \frac{1}{8}$ (since 1/2 is within the range $[-4, 4]$).

Calculate the acceptance probability:

$$a_{\text{ij}} = \min\left(1, \frac{\pi H_{\text{ji}}}{\pi_j H_{\text{ij}}}\right) = \min\left(1, \frac{\frac{1}{4} \cdot \frac{1}{8}}{0 \cdot \frac{1}{8}}\right)$$

From the equation, we can see that given the $j = 0$, the fraction becomes undefined. However, in the context of the Metropolis-Hastings algorithm, when the current state j has a target probability of zero, the transition to any other state with a non-zero target probability is always accepted. Therefore, the acceptance probability a_{ij} is 1.

7.3 Q7.2

Consider a transition from $x = \frac{1}{2}$ to $x = -1$. What is the probability this transition will be accepted.

7.4 Solution

For the transition from $x = 3$ to $x = \frac{1}{2}$:

- $\pi_j = p_T\left(\frac{1}{2}\right) = \frac{1}{4}$ (since $\frac{1}{2}$ is within the range $[-2, 2]$).
- $\pi_i = p_T(-1) = \frac{1}{4}$ (since -1 is within the range $[-2, 2]$).
- $H_{\text{ij}} = p_C\left(\frac{1}{2}\right) = \frac{1}{8}$ (since $\frac{1}{2}$ is within the range $[-4, 4]$).
- $H_{\text{ji}} = p_C\left(\frac{1}{2}\right) = \frac{1}{8}$ (since -1 is within the range $[-4, 4]$).

Calculate the acceptance probability:

$$a_{\text{ij}} = \min\left(1, \frac{\pi H_{\text{ji}}}{\pi_j H_{\text{ij}}}\right) = \min\left(1, \frac{\frac{1}{4} \cdot \frac{1}{8}}{\frac{1}{4} \cdot \frac{1}{8}}\right)$$

So, we find that the acceptance probability $a$ij is 1, since the ratio inside the min function equals 1. This means the transition from x = 2 to x = - 1 will be accepted with a probability of 1.

7.5 Q7.3

Consider a transition from $x = -1$ to $x = -3$. What is the probability this transition will be accepted.

7.6 Solution

For the transition from $x = -1$ to $x = -3$:

- $\pi_j = p_T(-1) = \frac{1}{4}$ (since $\frac{1}{2}$ is within the range $[-2, 2]$).
- $\pi_i = p_T(-3) = 0$ (since -3 is outside the range $[-2, 2]$).
- $H_{\text{ij}} = p_C(-1) = \frac{1}{8}$ (since 1 is within the range $[-4, 4]$).
- $H_{\text{ji}} = p_C(-3) = \frac{1}{8}$ (since -3 is within the range $[-4, 4]$).

Calculate the acceptance probability:

$$a_{\text{ij}} = \min\left(1, \frac{\pi H_{\text{ji}}}{\pi_j H_{\text{ij}}}\right) = \min\left(1, \frac{0 \cdot \frac{1}{8}}{\frac{1}{4} \cdot \frac{1}{8}}\right)$$

We find that the acceptance probability a_{ij} is 0, since the numerator of the fraction is 0. This means the transition from $x = -1$ to $x = -3$ will not be accepted.

8 Q8 Computational

These problems have a coding solution. Code your solutions in Python in the space provided and upload the notebook in the given space (Please directly upload your notebooks and not python files without outputs).

Use of LLM's such as ChatGPT are not allowed for this computational task, please try not to use such applications as we do check for the same during grading.

- Notebook Link: https://piazza.com/class_profile/get_resource/llqwp5rdfue104/loq97qrxqqwgf
- Batting.csv: https://piazza.com/class_profile/get_resource/llqwp5rdfue104/loq97t5zavhn8
- Pitching.csv: https://piazza.com/class_profile/get_resource/llqwp5rdfue104/loq97smafcwld