Exercises

Viviana Arrigoni
Internet of Things 2024-2025
Computer Science Department, Sapienza University of Rome

March 2025

Exercise 1 Consider the discrete-time homogeneous Markov chain described by the following transition matrix:

$$\mathbf{P} = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \end{array} \begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.3 & 0.3 \end{pmatrix} \end{array} \tag{1}$$

Let X_n be the state of the system at time n.

1. What does $\mathbb{P}[X_2 = s_1|X_0 = s_1]$ represent? What is its value? Show the reasoning.

2. What is $\mathbb{P}[X_5 = s_2|X_2 = s_3]$?

3. Find the stationary probability distribution π.

Solution:

1. $\mathbb{P}[X_2 = s_1|X_0 = s_1]$ is the probability of transitioning from state s_1 to state s_1 in two time steps. We have seen that $\mathbb{P}[X_n = s'|X_0 = s] = p^n_{s,s'}$, that is, the (s, s') entry of the matrix $P^n = P \cdot \ldots \cdot P$ (n times). Hence, $\mathbb{P}[X_2 = s_1|X_0 = s_1]$ can be computed as the dot product of the first row (that is the row related to s_1) of matrix $\mathbf{P}$ and the first column (that is the column related to s_1) of matrix $\mathbf{P}$. Hence:
   $\mathbb{P}[X_2 = s_1|X_0 = s_1] = 0.1 \cdot 0.1 + 0.5 \cdot 0.2 + 0.4 \cdot 0.4 = 0.27.$

2. $\mathbb{P}[X_5 = s_2|X_2 = s_3]$ is the probability of transitioning from state s_3 at time 2 to state s_2 at time 5. Since the Markov chain is homogeneous, this is just the probability of transitioning from state s_3 to state s_2 in 5-2=3 time steps. Such a probability is the entry (3,2) of matrix $\mathbf{P}^3$. To compute this entry we need to multiply $\mathbf{P}$ by itself to get $\mathbf{P}^2$, and then multiply

the 3rd row of $\mathbf{P}^2$ by the second column of $\mathbf{P}$:

$$\begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.3 & 0.3 \end{pmatrix} \cdot \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.3 & 0.3 \end{pmatrix} =$$
$$\begin{pmatrix} 0.2700 & 0.4200 & 0.3100 \\ 0.2400 & 0.4400 & 0.3200 \\ 0.2200 & 0.4400 & 0.3400 \end{pmatrix}$$

$p^3_{s_3,s_2} = 0.22 \cdot 0.5 + 0.44 \cdot 0.5 + 0.34 \cdot 0.3 = 0.432.$

3. To find the stationary probability distribution $\pi = (\pi_1, \pi_2, \pi_3)$, we need to solve the linear system $\{P^T \pi^T = \pi^T; \sum \pi_i = 1\}$, that is:

$$\begin{cases} 0.1\pi_1 + 0.2\pi_2 + 0.4\pi_3 = \pi_1 \\ 0.5\pi_1 + 0.5\pi_2 + 0.3\pi_3 = \pi_2 \\ 0.4\pi_1 + 0.3\pi_2 + 0.3\pi_3 = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

that is equivalent to:

$$\begin{cases} -0.9\pi_1 + 0.2\pi_2 + 0.4\pi_3 = 0 \\ 0.5\pi_1 - 0.5\pi_2 + 0.3\pi_3 = 0 \\ 0.4\pi_1 + 0.3\pi_2 - 0.7\pi_3 = 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

We can use any method to solve this system, for instance, the Gaussian elimination.

$$\begin{pmatrix} -0.9 & 0.2 & 0.4 & 0 \\ 0.5 & -0.5 & 0.3 & 0 \\ 0.4 & 0.3 & -0.7 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \xrightarrow{\text{swap rows}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -0.9 & 0.2 & 0.4 & 0 \\ 0.5 & -0.5 & 0.3 & 0 \\ 0.4 & 0.3 & -0.7 & 0 \end{pmatrix} \xrightarrow[+0.9\cdot\text{1st row}]{\text{2nd row = 2nd row +}}$$
$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1.1 & 1.3 & 0.9 \\ 0.5 & -0.5 & 0.3 & 0 \\ 0.4 & 0.3 & -0.7 & 0 \end{pmatrix} \xrightarrow[-0.5\cdot\text{1st row}]{\text{3rd row = 3rd row -}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1.1 & 1.3 & 0.9 \\ 0 & -1 & -0.2 & -0.5 \\ 0.4 & 0.3 & -0.7 & 0 \end{pmatrix} \xrightarrow[-0.4\cdot\text{1st row}]{\text{4th row = 4th row -}}$$
$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1.1 & 1.3 & 0.9 \\ 0 & -1 & -0.2 & -0.5 \\ 0 & -0.1 & -1.1 & -0.4 \end{pmatrix} \xrightarrow[\text{multiply by -1}]{\text{swap rows and}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0.2 & 0.5 \\ 0 & 1.1 & 1.3 & 0.9 \\ 0 & -0.1 & -1.1 & -0.4 \end{pmatrix} \xrightarrow[-1.1\cdot\text{2nd row}]{\text{3rd row = 3rd row -}}$$
$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0.2 & 0.5 \\ 0 & 0 & 1.08 & 0.35 \\ 0 & -0.1 & -1.1 & -0.4 \end{pmatrix} \xrightarrow[+0.1\cdot\text{2nd row}]{\text{4th row = 4th row +}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0.2 & 0.5 \\ 0 & 0 & 1.08 & 0.35 \\ 0 & 0 & -1.08 & -0.35 \end{pmatrix}$$

We can observe that the last two rows represent the same equation, so one is redundant and can be omitted. By rewriting the matrix in a system format, we get that:

$$\begin{cases} \pi_1 + \pi_2 + \pi_3 = 1 \\ \pi_2 + 0.2\pi_3 = 0.5 \\ 1.08\pi_3 = 0.35 \end{cases}$$

$$\begin{cases} \pi_1 + \pi_2 + \pi_3 = 1 \\ \pi_2 + 0.2\pi_3 = 0.5 \\ \pi_3 = \frac{0.35}{1.08} \sim 0.324 \end{cases}$$

$$\begin{cases} \pi_1 + \pi_2 + \pi_3 = 1 \\ \pi_2 = 0.5 - 0.2\pi_3 \sim 0.5 - 0.2 \cdot 0.324 \sim 0.435 \\ \pi_3 \sim 0.324 \end{cases}$$

$$\begin{cases} \pi_1 = 1 - \pi_2 - \pi_3 \sim 1 - 0.435 - 0.324 \sim 0.24 \\ \pi_2 \sim 0.435 \\ \pi_3 \sim 0.324 \end{cases}$$

This means that around the 24% of the time, the system will be in state s_1, around the 43% of the time it will be in state s_2, and around the 32% of the time it will be in state s_3.

Exercise 2 Consider the Markov Chain described by the following transition matrix:

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Are there absorbing states? If so, what are they? Is it periodic? Why?

Solution:
There are no absorbing states because, from any state, it is possible to reach any other state in a finite number of steps. Let's see if s_1 is periodic. Starting from s_1, it is not possible to reach state s_1 in one time step, hence $1 \notin \{n : p^n_{1,1} > 0\}$. From state s_1, with probability 1, the next state is state s_2. From this state, we can only move to step s_3 in one step, hence $2 \notin \{n : p^n_{1,1} > 0\}$. From state s_3, the system gets back to state s_1, meaning that it is possible to move from s_1 to s_1 in three time steps, hence $3 \in \{n : p^n_{1,1} > 0\}$. We can see that the system is deterministic, and that from s_1, it goes back to s_1 only after n time steps, where n is a multiple of 3. Hence $d(s_1) = 3$. The same result holds for states s_2 and s_3, which also have period 3.

Exercise 3 Find the stationary probability distributions of the discrete-time homogeneous Markov chains described by the following transition matrices:

$$\mathbf{P}_1 = \begin{array}{c c} & \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.4 & 0.6 & 0 \\ 0 & 1 & 0 \end{pmatrix} \end{array} \quad \mathbf{P}_2 = \begin{array}{c c} & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \end{array} & \begin{pmatrix} 0.9 & 0.1 \\ 0.8 & 0.2 \end{pmatrix} \end{array} \tag{2}$$

Exercise 4 Consider the Markov Chain described by the following transition matrix:

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0.4 & 0 & 0.6 \\ 0 & 0 & 1 \end{pmatrix}$$

Are there absorbing states? If so, what are they?

Exercise 5 Consider the following binary encoding defined over the set of words $\mathcal{M} = \{m_1, m_2, m_3\}$ as $f(m_1) = 0$, $f(m_2) = 10$, $f(m_3) = 11$.

1. Is it prefix-free?
2. Consider the following two distributions over $\mathcal{M}$:
   $P_1 : P_1(m_1) = \frac{1}{5}, P_1(m_2) = \frac{2}{5}, P_1(m_3) = \frac{2}{5}$ and

   $P_2 : P_2(m_1) = \frac{3}{6}, P_2(m_2) = \frac{2}{6}, P_2(m_3) = \frac{1}{6}$.

   Which one does the definition of f fit better (meaning that it has the lower average length)? Why?

Exercise 6 Consider the random variables X, Y defined over the set $\{A, B, C\}$. The variable X follows the distribution P_X such that $P_X(X = A) = P_X(X = B) = P_X(X = C) = \frac{1}{3}$. The random variable Y follows the distribution $P_Y(Y = A) = \frac{1}{4}$, $P_Y(Y = B) = \frac{1}{2}$, $P_Y(Y = C) = \frac{1}{4}$.

1. Compute the entropy of P_X and P_Y. Which one is larger and why?
2. Provide a prefix-free encoding for X and a prefix-free encoding for Y and compute their average length.

Exercise 7 Consider the random variable X defined over the set $\{A, B, C, D\}$. The variable X follows the distribution P_X such that $P_X(X = A) = \frac{1}{10}$, $P_X(X = B) = \frac{2}{10}$, $P_X(X = C) = \frac{3}{10}$, $P_X(X = D) = \frac{4}{10}$.

1. Compute the entropy of X.
2. Provide a prefix-free encoding for X and compute its average length.

Exercise 8 Consider the following discrete memoryless binary channel with input alphabet $\mathcal{X} = \{0, 1\}$ and output alphabet $\mathcal{Y} = \{a, b\}$ such that $\varphi(a) = 0$, $\varphi(b) = 1$:

$$W(Y|X) = \begin{array}{c c} & \begin{array}{cc} \mathbb{P}[Y = a|X] & \mathbb{P}[Y = b|X] \end{array} \\ \begin{array}{c} \mathbb{P}[Y|X = 0] \\ \mathbb{P}[Y|X = 1] \end{array} & \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \end{array} \tag{3}$$

1. Compute the following probabilities:
   $\mathbb{P}[\underline{Y} = aa | \underline{X} = 00]$,

   $\mathbb{P}[\underline{Y} = ab | \underline{X} = 00]$,

   $\mathbb{P}[\underline{Y} = ab | \underline{X} = 10]$,

   $\mathbb{P}[\underline{Y} = bbb | \underline{X} = 110]$.

2. Consider the two following options for the probability distribution of X:
   $P_1(X = 0) = 0.55, P_1(X = 1) = 0.45$,
   $P_2(X = 0) = 0.4, P_2(X = 1) = 0.6$.

   - Which one guarantees a higher channel capacity?
     (*hint*: use the law of total probability:
     $\mathbb{P}[Y = a] = \mathbb{P}[Y = a|X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = a|X = 1]\mathbb{P}[X = 1]$,

     $\mathbb{P}[Y = b] = \mathbb{P}[Y = b|X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = b|X = 1]\mathbb{P}[X = 1]$).

   - What is your intuition for your result?

   Partial solution to point 2:
   To answer this question, we need to compute the mutual information of X and Y, and see for what choice of the probability distribution of X (either P_1 or P_2) the mutual information is maximal.
   Let $X \sim P_1$. The entropy of Y is:

$$H(Y) = \mathbb{P}[Y = a] \log \frac{1}{\mathbb{P}[Y = a]} + \mathbb{P}[Y = b] \log \frac{1}{\mathbb{P}[Y = b]}$$

   Therefore, we need to compute $\mathbb{P}[Y = a]$ and $\mathbb{P}[Y = b]$

$$\begin{aligned}\mathbb{P}[Y = a] &= \mathbb{P}[Y = a|X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = a|X = 1]\mathbb{P}[X = 1] = \\ &= 0.9 \cdot 0.55 + 0.2 \cdot 0.45 = 0.585\end{aligned}$$

$$\begin{aligned}\mathbb{P}[Y = b] &= \mathbb{P}[Y = b|X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = b|X = 1]\mathbb{P}[X = 1] = \\ &= 0.1 \cdot 0.55 + 0.8 \cdot 0.45 = 0.415\end{aligned}$$

   Hence $H(Y) = 0.9791$. Let's now compute $H(Y|X)$, that is:

$$\begin{aligned}H(Y|X) = \mathbb{P}[X = 0]\mathbb{P}[Y = a|X = 0] \log \frac{1}{\mathbb{P}[Y = a|X = 0]} &+ \\ + \mathbb{P}[X = 1]\mathbb{P}[Y = a|X = 1] \log \frac{1}{\mathbb{P}[Y = a|X = 1]} &+ \\ + \mathbb{P}[X = 0]\mathbb{P}[Y = b|X = 0] \log \frac{1}{\mathbb{P}[Y = b|X = 0]} &+ \\ + \mathbb{P}[X = 1]\mathbb{P}[Y = b|X = 1] \log \frac{1}{\mathbb{P}[Y = b|X = 1]}&\end{aligned}$$

which is:

$$
\begin{aligned}
H(Y|X) = 0.55 \cdot 0.9 \log \frac{1}{0.9} + \\
+ 0.45 \cdot 0.2 \log \frac{1}{0.2} + \\
+ 0.55 \cdot 0.1 \log \frac{1}{0.1} + \\
+ 0.45 \cdot 0.8 \log \frac{1}{0.8} = 0.5828.
\end{aligned}
$$

Therefore, $I(X, Y) = 0.3962$ when $X \sim P_1$. To complete the exercise, compute the mutual information of X and Y for when $X \sim P_2$ and explain the reason for your result.