1 Problem sheet 1: Discrete-time Markov chains

1.1 Prerequisites: Lecture 2

Exercise 1- 1: Show by induction that for a discrete Markov chain $(X_n)_{n\in\mathbb{N}_0}$, we have

$$P(X_{n+m} = x_{n+m}|X_n = x_n, \dots, X_0 = x_0) = P(X_{n+m} = x_{n+m}|X_n = x_n),$$

for $m \in \mathbb{N}$ and for all $x_{n+m}, x_n, \dots, x_0 \in E$.

Solution: For $m = 1$, this is the definition of the Markov property. Assume now the identity holds for m. We apply the law of total probability with additional conditioning, which leads to

$$\begin{aligned}
&P(X_{n+m+1} = x_{n+m+1}|X_n = x_n, \dots, X_0 = x_0) \\
&= \sum_{x_{n+m}\in E} P(X_{n+m+1} = x_{n+m+1}|X_{n+m} = x_{n+m}, X_n = x_n, \dots, X_0 = x_0) \\
&\qquad \cdot P(X_{n+m} = x_{n+m}|X_n = x_n, \dots, X_0 = x_0).
\end{aligned}$$

Now we apply the Markov property to the first term

$$\begin{aligned}
&P(X_{n+m+1} = x_{n+m+1}|X_{n+m} = x_{n+m}, X_n = x_n, \dots, X_0 = x_0) \\
&= P(X_{n+m+1} = x_{n+m+1}|X_{n+m} = x_{n+m}, X_n = x_n),
\end{aligned}$$

where the green term $X_n = x_n$ could have been deleted as well, but we are keeping it for now, and the induction hypothesis to the second term

$$P(X_{n+m} = x_{n+m}|X_n = x_n, \dots, X_0 = x_0) = P(X_{n+m} = x_{n+m}|X_n = x_n).$$

Hence we get

$$\begin{aligned}
&P(X_{n+m+1} = x_{n+m+1}|X_n = x_n, \dots, X_0 = x_0) \\
&= \sum_{x_{n+m}\in E} P(X_{n+m+1} = x_{n+m+1}|X_{n+m} = x_{n+m}, X_n = x_n)P(X_{n+m} = x_{n+m}|X_n = x_n) \\
&= P(X_{n+m+1} = x_{n+m+1}|X_n = x_n) = p_{x_n x_{n+m+1}}(m+1),
\end{aligned}$$

where we applied the law of total probability with additional conditioning (”backwards”) in the penultimate identity. Now we see why we kept the green term $X_n = x_n$: Then we have exactly the functional form of the law of total probability with additional conditioning.

Exercise 1- 2: Show that a $K \times K$-dimensional *stochastic matrix* has at least one eigenvalue equal to 1. Hence show that if $\mathbf{P}$ is a stochastic matrix, then so is $\mathbf{P}^n$ for all $n \in \mathbb{N}$.

Solution: A stochastic matrix $\mathbf{P}$ must satisfy

$$\sum_{j=1}^{K} p_{ij} = 1, \qquad \text{for all } i \in \{1, \dots, K\}.$$

We can write the above condition in matrix notation:

$$\mathbf{P}\mathbf{1} = 1 \cdot \mathbf{1}, \tag{1.1}$$

where $\mathbf{1}$ is the vector of 1's. So we see that 1 is an eigenvalue.

We use induction to show that $\mathbf{P}^n$ is a stochastic matrix. We already know that this is true for $n = 1$. Then for $n + 1$ we have

$$\mathbf{P}^{n+1}\mathbf{1} = \mathbf{P}(\mathbf{P}^n\mathbf{1}) = \mathbf{P}\mathbf{1} = \mathbf{1},$$

where we used that according to the induction hypothesis $\mathbf{P}^n$ is a stochastic matrix (in particular $\mathbf{P}^n\mathbf{1}$). In addition, since all elements of $\mathbf{P}$ are nonnegative, this is also true for the elements of $\mathbf{P}^{n+1}$ (since both $\mathbf{P}$ and by induction hypothesis $\mathbf{P}^n$ are non-negative, hence their product has non-negative elements).

Exercise 1- 3: For each matrix, decide whether it is stochastic. If it is, draw the corresponding transition diagram (assuming the state space is given by $E = \{1, 2, 3\}$).

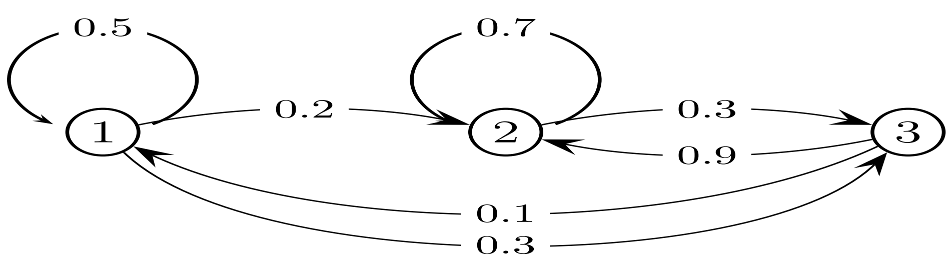
$$\text{(a)} \begin{pmatrix} 0 & 0 & 1 \\ 0.5 & -0.5 & -1 \\ 0.25 & 0.25 & 0.5 \end{pmatrix} \qquad \text{(b)} \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0 & 0.7 & 0.3 \\ 0.1 & 0.9 & 0 \end{pmatrix} \qquad \text{(c)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.9 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}$$

Solution:

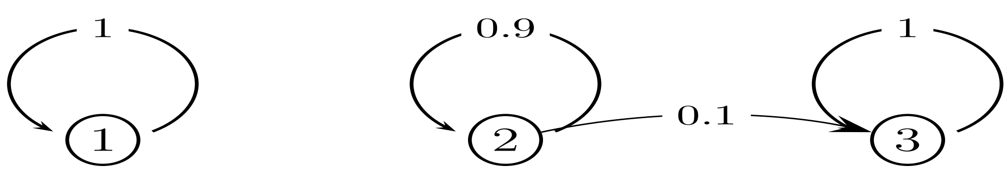
(a) The matrix is not stochastic since it has negative entries.

(b) and (c) Both matrices are stochastic since they have non–negative entries, and the row sums are 1.

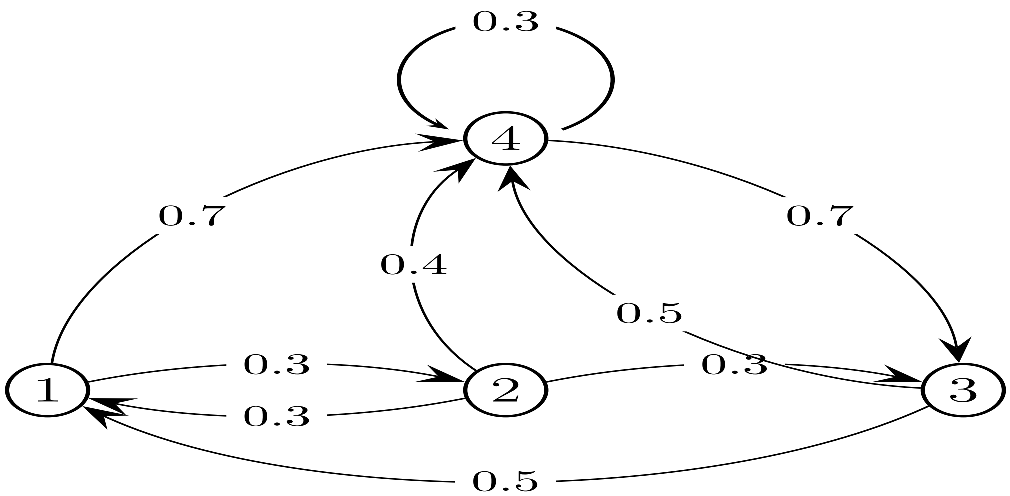
The transition diagram for (b) is given by



The transition diagram for (c) is given by



Exercise 1- 4: Derive the transition matrix from the following transition diagram.



Solution:

$$\mathbf{P} = \begin{pmatrix} 0 & 0.3 & 0 & 0.7 \\ 0.3 & 0 & 0.3 & 0.4 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0.7 & 0.3 \end{pmatrix}$$

Exercise 1- 5: Suppose that the movements of a particle are recorded at discrete times on a discrete lattice $\{0, 1, \ldots, a\}$, such that if it hits either 0 or a it never leaves. Given that the particle is at the points $\{1, \ldots, a-1\}$ it may jump up the lattice with probability p and down with probability $1-p$. What are the transition probabilities for this Markov chain?

Solution: Suppose that $x_{n-1} \in \{1, \ldots, a-1\}$ then

$$\mathrm{P}(X_n = x_{n-1} + 1 | X_{n-1} = x_{n-1}) = p$$

and

$$\mathrm{P}(X_n = x_{n-1} - 1 | X_{n-1} = x_{n-1}) = 1 - p.$$

If $x_{n-1} \in \{0, a\}$ then

$$\mathrm{P}(X_n = x_{n-1} | X_{n-1} = x_{n-1}) = 1.$$

All the other transition probabilities are equal to 0. We note that the Markov chain described here is also referred to as a random walk with absorption at 0 and a.

Exercise 1- 6: Let X_n be the maximum reading obtained in the first n rolls of a fair die (for $n \in \mathbb{N}$). Show that $\{X_n\}_{n \in \mathbb{N}}$ is a Markov chain, and give the transition probabilities.

Solution: Let D_n be the score of the die at time n. Then D_n is a uniform random variable on $\{1, \ldots, 6\}$. Then

$$X_n = \max\{D_1, \ldots, D_n\} = \max\{X_{n-1}, D_n\}.$$

Clearly X_n depends on $(X_1, \ldots, X_{n-1})$ only through X_{n-1} and D_n is independent of $X_{n-2}, \ldots, X_1$, so (X_n) is a Markov chain, i.e.

$$\mathrm{P}(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_1 = x_1) = \mathrm{P}(X_n = x_n | X_{n-1} = x_{n-1}).$$

It also follows that

$$\mathrm{P}(X_n = x_{n-1} | X_{n-1} = x_{n-1}) = \frac{x_{n-1}}{6}.$$

Also, if $x_{n-1} \leq 5$ and $y \in \{x_{n-1}+1, \ldots, 6\}$, we have

$$\mathrm{P}(X_n = y | X_{n-1} = x_{n-1}) = \frac{1}{6}.$$

Altogether, we get the following transition matrix

$$\mathbf{P} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 2/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 3/6 & 1/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 4/6 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 6/6 \end{pmatrix}.$$

Exercise 1- 7: Consider the transition matrix of a Markov chain with state space $E = \{1, 2, 3\}$ given by

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \end{pmatrix}.$$

Derive a general formula $p_{12}(n)$ for $n \in \{0, 1, 2, \ldots\}$. *Hint:* Compute the eigenvalues of $\mathbf{P}$ and diagonalise $\mathbf{P}$ to compute $\mathbf{P}^n$!

Solution: First we compute the eigenvalues of $\mathbf{P}$: Solve the characteristic equation

$$\begin{aligned} 0 = \det(\mathbf{P} - \lambda \mathbf{I}) &= \det \begin{pmatrix} -\lambda & 1 & 0 \\ 0 & 2/3 - \lambda & 1/3 \\ 1/3 & 0 & 2/3 - \lambda \end{pmatrix} \\ &= -\lambda \left(\frac{2}{3} - \lambda\right)^2 + \left(\frac{1}{3}\right)^2 = -\lambda^3 + \frac{4}{3}\lambda^2 - \frac{4}{9}\lambda + \frac{1}{9} \\ &= -(\lambda - 1)\left(\lambda^2 - \frac{1}{3}\lambda + \frac{1}{9}\right). \end{aligned}$$

Hence, $\mathbf{P}$ has three distinct eigenvalues given by

$$\lambda_1 = 1, \quad \lambda_2 = \frac{1 + i\sqrt{3}}{6}, \quad \lambda_3 = \frac{1 - i\sqrt{3}}{6},$$

and is therefore diagonalisable.

I.e. we know that there exists an invertible matrix $\mathbf{U}$ such that

$$\mathbf{U}^{-1}\mathbf{P}\mathbf{U} = \mathrm{diag}\left(1, \frac{1+i\sqrt{3}}{6}, \frac{1-i\sqrt{3}}{6}\right),$$

i.e.

$$\mathbf{P} = \mathbf{U}\mathrm{diag}\left(1, \frac{1+i\sqrt{3}}{6}, \frac{1-i\sqrt{3}}{6}\right)\mathbf{U}^{-1}.$$

Then we have for $n \in \{0, 1, 2, \dots\}$

$$\mathbf{P}^n = \mathbf{U}\mathrm{diag}\left(1, \left(\frac{1+i\sqrt{3}}{6}\right)^n, \left(\frac{1-i\sqrt{3}}{6}\right)^n\right)\mathbf{U}^{-1}.$$

That means that each element of $\mathbf{P}^n$ is of the form

$$a + b\left(\frac{1+i\sqrt{3}}{6}\right)^n + c\left(\frac{1-i\sqrt{3}}{6}\right)^n,$$

for some (possibly complex) coefficients a, b, c.

The coefficients can be computed using the initial values $n = 0, 1, 2$. Note that $\mathbf{P}^0$ is the identity matrix, $\mathbf{P}^1 = \mathbf{P}$ is given and $\mathbf{P}^2$ can be easily computed, here we have

$$\mathbf{P}^2 = \begin{bmatrix} 0 & 2/3 & 1/3 \\ 1/9 & 4/9 & 4/9 \\ 2/9 & 1/3 & 4/9 \end{bmatrix}$$

Now, we need to solve

$$p_{12}(n) = a + b\left(\frac{1+i\sqrt{3}}{6}\right)^n + c\left(\frac{1-i\sqrt{3}}{6}\right)^n,$$

for $n = 0, 1, 2$ with respect to the coefficients a, b, c.

I.e. solve the following system of equations:

$$\begin{aligned}
p_{12}(0) = 0 &= a + b\left(\frac{1+i\sqrt{3}}{6}\right)^0 + c\left(\frac{1-i\sqrt{3}}{6}\right)^0 = a + b + c, \\
p_{12}(1) = 1 &= a + b\left(\frac{1+i\sqrt{3}}{6}\right) + c\left(\frac{1-i\sqrt{3}}{6}\right), \\
p_{12}(2) = \frac{2}{3} &= a + b\left(\frac{1+i\sqrt{3}}{6}\right)^2 + c\left(\frac{1-i\sqrt{3}}{6}\right)^2.
\end{aligned}$$

Now we simplify the computations by getting rid off the imaginary part. Note that the n-step transition probabilities are real non-negative numbers!

Recall that, for a complex number $z \in \mathbb{C}$, with Cartesian form $z = x + iy$ for $x, y \in \mathbb{R}$, we can write it in polar form as $z = r(\cos(\phi) + i\sin(\phi)) = re^{i\phi}$, where $r = |z| = \sqrt{x^2 + y^2}$, $\phi = \arg(z) = \mathrm{atan2}(y, x)$.

Observe that

$$\frac{1 \pm i\sqrt{3}}{6} = \frac{1}{3}\frac{1 \pm i\sqrt{3}}{2} = \frac{1}{3}e^{\pm i\pi/3} = \frac{1}{3}\left(\cos\frac{\pi}{3} \pm i\sin\frac{\pi}{3}\right).$$

Hence

$$\left(\frac{1 \pm i\sqrt{3}}{6}\right)^n = \left(\frac{1}{3}\right)^n e^{\pm i\pi n/3} = \left(\frac{1}{3}\right)^n \left(\cos\frac{\pi n}{3} \pm i\sin\frac{\pi n}{3}\right),$$

and

$$p_{12}(n) = \alpha + \left(\frac{1}{3}\right)^n \left(\beta\cos\frac{\pi n}{3} + \gamma\sin\frac{\pi n}{3}\right),$$

where $\alpha = a$, $\beta = b + c$ and $\gamma = i(b - c)$ must be real numbers.

The new system of equation we need to solve is given by

$$\alpha + \beta = 0, \alpha + \frac{1}{3}\left(\frac{1}{2}\beta + \frac{\sqrt{3}}{2}\gamma\right) = 1, \alpha + \frac{1}{3^2}\left(-\frac{1}{2}\beta + \frac{\sqrt{3}}{2}\gamma\right) = \frac{2}{3},$$

which has solution

$$\alpha = \frac{3}{7}, \quad \beta = -\frac{3}{7}, \quad \gamma = \frac{9}{7}\sqrt{3}.$$

Altogether, we have

$$p_{12}(n) = \frac{3}{7} + \left(\frac{1}{3}\right)^n \left(-\frac{3}{7}\cos\frac{\pi n}{3} + \frac{9}{7}\sqrt{3}\sin\frac{\pi n}{3}\right).$$

1.2 Prerequisites: Lecture 3

Exercise 1- 8: Show that a discrete random process $(X_n)_{0\le n\le N}$ $(N < \infty)$ is a discrete-time, time-homogeneous Markov chain on the state space E if and only if

$$\mathrm{P}(X_0 = x_0, X_1 = x_1, \ldots, X_N = x_N) = \mathrm{P}(X_0 = x_0)p_{x_0x_1}p_{x_1x_2}\cdots p_{x_{N-1}x_N},$$

for all $x_0, x_1, \ldots, x_N \in E$.

Solution: Suppose that $(X_n)_{0\le n\le N}$ is a discrete-time, time-homogeneous Markov chain. Then

$$\begin{aligned}
&\mathrm{P}(X_0 = x_0, X_1 = x_1, \ldots, X_N = x_N)\\
&= \mathrm{P}(X_N = x_N | X_0 = x_0, X_1 = x_1, \ldots, X_{N-1} = x_{N-1})\mathrm{P}(X_0 = x_0, X_1 = x_1, \ldots, X_{N-1} = x_{N-1})\\
&= \mathrm{P}(X_N = x_N | X_{N-1} = x_{N-1})\mathrm{P}(X_0 = x_0, X_1 = x_1, \ldots, X_{N-1} = x_{N-1}),
\end{aligned}$$

where we used the Markov property. By iterating the argument, we get

$$\begin{aligned}
&\mathrm{P}(X_0 = x_0, X_1 = x_1, \ldots, X_N = x_N)\\
&= \mathrm{P}(X_N = x_N | X_{N-1} = x_{N-1})\mathrm{P}(X_0 = x_0, X_1 = x_1, \ldots, X_{N-1} = x_{N-1})\\
&= \mathrm{P}(X_N = x_N | X_{N-1} = x_{N-1})\cdots\mathrm{P}(X_1 = x_1 | X_0 = x_0)\mathrm{P}(X_0 = x_0).
\end{aligned}$$

Using the notation from the lecture notes this can be written as

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1, \ldots, X_N = x_N) &= P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{N-1} x_N} \\ &= \nu^{(0)}_{x_0} p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{N-1} x_N}. \end{aligned}$$

Now suppose that

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1, \ldots, X_N = x_N) &= P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{N-1} x_N} \\ &= \nu^{(0)}_{x_0} p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{N-1} x_N} \qquad (1.2) \end{aligned}$$

holds. Sum both sides of the equation over $x_N \in E$ using the law of total probability and the fact that the row sums of the transition matrix are equal to 1. Then

$$\sum_{x_N \in E} P(X_0 = x_0, X_1 = x_1, \ldots, X_N = x_N) = \sum_{x_N \in E} P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{N-1} x_N},$$

hence

$$P(X_0 = x_0, X_1 = x_1, \ldots, X_{N-1} = x_{N-1}) = P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{N-2} x_{N-1}}.$$

So, equation (1.2) holds for $N - 1$.

Now we iterate to conclude that

$$P(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{n-1} x_n}$$

holds for all $n = 0, \ldots, N$. Then

$$\begin{aligned} &P(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) \\ &= P(X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) / P(X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) \\ &= \frac{P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{n-1} x_n}}{P(X_0 = x_0) p_{x_0 x_1} p_{x_1 x_2} \cdots p_{x_{n-2} x_{n-1}}} \\ &= p_{x_{n-1} x_n}, \end{aligned}$$

which is the (time-homogeneous) Markov property.

Exercise 1- 9: Using the notation from the lecture notes, show by induction that the marginal distribution of a Markov chain satisfies

$$\boldsymbol{\nu}^{(n)} = \boldsymbol{\nu}^{(0)} \mathbf{P}_n,$$

for $n \in \mathbb{N}$.

Solution: For $n = 1$, we have (for $j \in E$)

$$\begin{aligned} \nu^{(1)}_j &= P(X_1 = j) = \\ &\stackrel{\text{LTP}}{=} \sum_{i \in E} P(X_1 = j | X_0 = i) P(X_0 = i) \\ &= \sum_{i \in E} p_{ij} \nu^{(0)}_i = \sum_{i \in E} \nu^{(0)}_i p_{ij}. \end{aligned}$$

Hence $\boldsymbol{\nu}^{(1)} = \boldsymbol{\nu}^{(0)}\mathbf{P}$.

Now we consider the case $n+1$:

$$\begin{aligned}\nu_j^{(n+1)} &= \mathrm{P}(X_{n+1} = j)\\ &\overset{\text{LTP}}{=} \sum_{i\in E} \mathrm{P}(X_{n+1} = j | X_n = i)\mathrm{P}(X_n = i)\\ &= \sum_{i\in E} p_{ij}\nu_i^{(n)}.\end{aligned}$$

Now we apply the induction hypothesis and get

$$\begin{aligned}\nu_j^{(n+1)} &= \sum_{i\in E} p_{ij}\nu_i^{(n)} = \sum_{i\in E} p_{ij} \sum_{k\in E} \nu_k^{(0)} p_{ki}(n)\\ &= \sum_{k\in E} \nu_k^{(0)} \underbrace{\sum_{i\in E} p_{ki}(n)p_{ij}}_{=p_{kj}(n+1)},\end{aligned}$$

where we used the Chapman-Kolmogorov equations. Hence $\boldsymbol{\nu}^{(n+1)} = \boldsymbol{\nu}^{(0)}\mathbf{P}_{n+1}$.

1.3 Prerequisites: Lecture 4

Exercise 1- 10: Suppose that $p, q \in (0,1)$ and $p+q=1$. The transition matrix of a Markov chain with state space $E = \{0, 1, \ldots,\}$ is:

$$\mathbf{P} = \begin{pmatrix} q & p & 0 & 0 & 0 & \ldots \\ q & 0 & p & 0 & 0 & \ldots \\ q & 0 & 0 & p & 0 & \ldots \\ q & 0 & 0 & 0 & p & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

(a) Find the first return probabilities for state 0, i.e. find $f_{00}(n)$ for all $n \in \mathbb{N}$.

(b) Use your results from (a) to find f_{00}. Is state 0 recurrent?

Solution: The first part of the transition diagram is given by:

$$\begin{tikzcd} 0 \arrow[loop above, "q"] \arrow[r, bend left, "p"] & 1 \arrow[l, bend left, "q"] \arrow[r, bend left, "p"] & 2 \arrow[ll, bend left, "q"] & \text{etc.} \end{tikzcd}$$

(a) We denote the Markov chain by $X = (X_n)_{n\in\mathbb{N}_0}$. Recall that, for $n \in \mathbb{N}$, $f_{00}(n)$ denotes the probability that the first return to 0 is after n steps. Here we find that

$$f_{00}(1) = p_{00} = q,$$

$$f_{00}(2) = pq,$$
$$f_{00}(3) = p^2 q,$$

etc. We read off that, for $n \in \mathbb{N}$,

$$f_{00}(n) = \mathrm{P}(X_n = 0, X_{n-1} \neq 0, \ldots, X_1 \neq 0 | X_0 = 0) = p^{n-1} q.$$

(b) Consider

$$f_{00} = \sum_{n=1}^{\infty} f_{00}(n) = q \sum_{n=1}^{\infty} p^{n-1} = q \sum_{n=0}^{\infty} p^n = \frac{q}{1-p} = 1,$$

where we worked with the geometric series expansion for $|p| = p < 1$. That is, state 0 is recurrent. (In fact it is positive recurrent.)

Exercise 1- 11: (From the exam paper in 2019-2020) Suppose the influenza virus exists in K different strains, where $K \geq 2$. Each year, the virus either stays the same with probability $1-a$, for $a \in (0,1)$, or mutates to any of the other strains with equal probability. Suppose you can model the virus mutation by a discrete-time homogeneous Markov chain.

1. We denote the state space by $E = \{1, \ldots, K\}$. State the corresponding 1-step transition probabilities of the Markov chain.
2. You decide to group the states: You consider the modified state space $\widetilde{E} = \{I, O\}$ where I stands for the initial state and O for the collection of the other $K-1$ states.
   (a) State the corresponding 1-step transition probabilities of the Markov chain on $\widetilde{E}$.
   (b) Show that, for $n \in \mathbb{N}$,

$$p_{II}(n+1) = p_{II}(n) \left\{ 1 - a - \frac{a}{K-1} \right\} + \frac{a}{K-1},$$

   and state all results from the lectures which you apply in your proof.

Solution:

1. We have $p_{ii} = 1-a$ for all $i \in E$ and we have that $p_{ij} = b$ for all $i \neq j$ for some $b \in (0,1)$, hence $1 = \sum_{j=1}^{K} p_{ij} = 1 - a + (K-1)b \Leftrightarrow b = \frac{a}{K-1}$, which implies that $p_{ij} = \frac{a}{K-1}$ for all $i, j \in E, i \neq j$.
2. (a) Here the transition matrix corresponding to the Markov chain on $\widetilde{E} = \{I, O\}$ is given by

$$\mathbf{P} = \begin{pmatrix} 1-a & a \\ \frac{a}{K-1} & 1 - \frac{a}{K-1} \end{pmatrix}.$$

   (b) Let $n \in \mathbb{N}$. From the Chapman-Kolmogorov equations, we have that $\mathbf{P}^{n+1} = \mathbf{P}^n \mathbf{P}$, hence

$$\begin{aligned} p_{II}(n+1) &= p_{II}(n) \cdot p_{II} + p_{IO}(n) \cdot p_{OI}. \\ &= p_{II}(n) \cdot (1-a) + p_{IO}(n) \cdot \frac{a}{K-1}. \end{aligned}$$

Since $\mathbf{P}^n$ is a stochastic matrix, we have that $p_{II}(n) + p_{IO}(n) = 1 \Leftrightarrow p_{IO}(n) = 1 - p_{II}(n)$.
Hence

$$\begin{aligned} p_{II}(n+1) &= p_{II}(n) \cdot p_{II} + p_{IO}(n) \cdot p_{OI} \\ &= p_{II}(n) \cdot p_{II} + (1 - p_{II}(n)) \cdot p_{OI} \\ &= p_{II}(n)(p_{II} - p_{OI}) + p_{OI} \\ &= p_{II}(n) \left(1 - a - \frac{a}{K-1}\right) + \frac{a}{K-1}. \end{aligned}$$

1.4 Prerequisites: Lecture 6

Exercise 1- 12: Let $\alpha, \beta \in (0,1)$. We are given a Markov chain on a state space $E = \{0,1\}$, with transition matrix

$$\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}.$$

(a) Derive the generating function associated with the sequence $(f_{00}(n))_{n\in\mathbb{N}_0}$. *H*int: Find $f_{00}(n)$ for all $n \in N_0$ for this Markov chain. Then define the corresponding generating function as $G(s) = \sum_{n=0}^{\infty} f_{00}(n)s^n$ for $|s| < 1$.

(b) Find, using generating functions, the expected return time to state 0. I.e. compute $\mu_0 = \frac{d}{ds}G(s)\big|_{s=1} = \sum_{n=1} nf_{00}(n)$.

(c) What is the "holding time" (i.e. the time spent in a particular state) distribution for state 0 that you have derived in this question?

Solution:

(a) The generating function associated with the sequence $(f_{00}(n))_{n\in\mathbb{N}_0}$ is given by

$$G(s) = \sum_{n=0}^{\infty} f_{00}(n)s^n.$$

Recall that $f_{00}(0) := 0$ and $f_{00}(1) = p_{00} = (1-\alpha)$. For any subsequent $n \in \mathbb{N}$, we must first visit 1 and stay there for $n-2$ step and return to 0. For example, if $n = 4$, $f_{00}(4) = \alpha(1-\beta)^2\beta$, in other words

$$f_{00}(n) = \alpha(1-\beta)^{n-2}\beta, \quad \text{for all } n \geq 2.$$

Thus we have

$$\begin{aligned} G(s) &= (1-\alpha)s + \alpha\beta \sum_{n=2}^{\infty} (1-\beta)^{n-2}s^n \\ &= (1-\alpha)s + \alpha\beta s^2 \sum_{n=0}^{\infty} (1-\beta)^n s^n \\ &= (1-\alpha)s + \frac{\alpha\beta s^2}{1-(1-\beta)s}, \end{aligned}$$

for $|s| < 1$.

(b) In order to compute μ_0, we compute the derivative of G and set $s = 1$

$$\frac{dG}{ds} = (1-\alpha) + \frac{2\alpha\beta s(1-(1-\beta)s) + \alpha\beta s^2(1-\beta)}{(1-(1-\beta)s)^2}$$

it then follows that

$$\left.\frac{dG}{ds}\right|_{s=1} = \sum_{n=1}^{\infty} n f_{00}(n) = \mu_0 = \frac{(\alpha+\beta)}{\beta}.$$

We note that, in Lecture 7, we have recalled Abel's theorem to stress that we can evaluate the (derivative) of the generating function in 1.

(c) The important point you have derived here is that the holding time in a state is **geometrically** distributed.

Suppose that $X_0 = i$. Define $Y_i := \inf\{n \in \mathbb{N} : X_n \neq i\} \in \mathbb{N}$ and let $Z_i := Y_i - 1 \in \mathbb{N}_0$ be the holding time at state i. Then

$$\begin{aligned}
\mathrm{P}(Z_0 = 0|X_0 = 0) &= \mathrm{P}(Y_0 = 1|X_0 = 0) = \alpha = f_{01}(1),\\
\mathrm{P}(Z_0 = 1|X_0 = 0) &= (1-\alpha)\alpha = f_{01}(2),\\
&\vdots\\
\mathrm{P}(Z_0 = n|X_0 = 0) &= (1-\alpha)^n\alpha = f_{01}(n).
\end{aligned}$$

I.e. the holding time in state 0 follows a geometric distribution with parameter α.

Similarly, the holding time in state 1 follows a geometric distribution with parameter β.

1.5 Prerequisites: Lecture 6

Exercise 1- 13: The following question is adapted from Grimmett & Stirzaker (2001*b*,*a*) Problem 6.15.6: Let $i, j \in E$ and suppose that $i \leftrightarrow j$.

Show that there is positive probability of reaching j from i without revisiting i in the meantime. Deduce that, if the chain is irreducible and recurrent, then $f_{ij} = 1, \forall i, j \in E$.

Solution: Suppose that $i \neq j$ and set $m = \min\{n \in \mathbb{N} : p_{ij}(n) > 0\}$. Such an m exists since $i \leftrightarrow j$. If $X_0 = i$ and $X_m = j$ then there can be no intermediate visit to i (with probability one), since such a visit would contradict the minimality of m. [To see that, suppose that $X_k = i$ for $k < m$. Then $p_{ij}(m-k) > 0$ and $m - k < m$, which contradicts that m is minimal.]

Now suppose that the chain is irreducible and recurrent. Suppose that $X_0 = i$. Since i is recurrent, we know that $f_{ii} = 1$ which is equivalent to

$$\mathrm{P}(X_n \neq i, \forall n \in \mathbb{N}|X_0 = i) = 1 - f_{ii} = 0.$$

One way of not re-visiting state i would be to visit state j in m steps and then never return to state i which happens with probability $1 - f_{ji}$. So altogether, we get

$$0 = \mathrm{P}(X_n \neq i, \forall n \in \mathbb{N}|X_0 = i) \geq p_{ij}(m)(1 - f_{ji}),$$

since $p_{ij}(m) > 0$, we require that $1 - f_{ji} = 0$ which concludes the proof (since the result holds for all $i, j \in E$).