 E Peltola

3 Long-term behavior of Markov chains

In this exercise you learn to recognize whether a Markov chain is reducible or periodic, and whether the chain admits a limiting distribution, by inspecting the transition matrix and the transition diagram of the chain. You also learn to compute the invariant distributions of a given transition matrix. It is recommended to bring a laptop or a calculator to the exercise session to make it easier to calculate the numerical results of the exercises.

Classroom problems

3.1 *Periodicity of an irreducible chain.* Justify why the following results are true for the transition matrix P of a Markov chain with a finite state space S. (Recall that $P^t(x,y)$ denotes the entry on row x and column y of the t-th matrix power of P.)

(a) If $P(x,x) > 0$, then also $P^t(x,x) > 0$ for all $t \geq 1$.

Solution. The inequality $P(x,x) > 0$ means that there is a link from state x to itself, whereas $P^t(x,x) > 0$ means that it is possible to get from state x to itself in t steps. The latter is true when the first holds, since we may simply take t steps $x \to x$. Formally we may write this idea as:

$$\begin{aligned} P^t(x,x) &= \mathbb{P}(X_t = x | X_0 = x) \\ &= \sum_{a_1,a_2,\ldots,a_{t-1} \in S} \mathbb{P}(X_t = x, X_{t-1} = a_{t-1}, \ldots, X_1 = a_1 \mid X_0 = x) \\ &\geq \mathbb{P}(X_t = x, X_{t-1} = x, ..., X_1 = x | X_0 = x) \\ &= P(x,x)^t > 0, \text{ because } P(x,x) > 0. \end{aligned}$$

(b) If $P(x,x) > 0$, then the period of state x is 1.

Solution. If $P(x,x) > 0$, then the set of possible return times is $\mathcal{T}_x = \{1,2,3,\ldots\}$, with the greatest common divisor 1. Thus, the period of state x is 1.

(c) If $P(x,x) > 0$ and $x \leftrightsquigarrow y$ (both states are reachable from each other by directed paths in the transition diagram), then there exists an integer $t_0 \geq 1$ such that $P^t(y,y) > 0$ for all $t \geq t_0$.

Solution. If $x \leftrightsquigarrow y$, then there are integers s_1 ja s_2 s.t. $P^{s_1}(y,x) > 0$ and $P^{s_2}(x,y) > 0$. Denote $s = s_1 + s_2$. Now $P^{s+k}(y,y) \geq P^{s_1}(y,x)P^k(x,x)P^{s_2}(x,y) > 0$ for all $k \geq 1$.

(d) An irreducible chain is aperiodic if $P(x,x) > 0$ holds for some state x.

Solution. The chain is irreducible if for all states z, y it holds that $z \leftrightsquigarrow y$. Let $z = x$. By the previous part for every y there is $s \geq 1$ such that $P^t(y,y) > 0$ for all $t = s, s+1, \ldots$. So, the possible return times $\mathcal{T}_y$ of state y contains $\{s, s+1, s+2, \ldots\}$. The only positive integer that divides both s and $s+1$ is 1. Thus the gcd of the numbers $\mathcal{T}_y$ is 1, i.e., the period of y is 1.

Additional information. More generally, all states of an irreducible Markov chain have the same period. The proof is essentially the same as part (d).

Homework problems

3.2 Determine the long-term behavior of the following Markov chains.

(a) The bike of a bicycle commuter on a given work day is either `unbroken` or `broken`. If the bike is `unbroken` on a given work day, then it's also `unbroken` the following day with probability 95% and otherwise `broken`. If the bike is `broken`, then it's `unbroken` the next work day with probability 33% and otherwise `broken`. In both cases, the state of the bike is independent of any earlier states. In the long term, what is the proportion of work days that the bike is `broken`?

Solution.

Denote the state `unbroken` by 1 and the state `broken` by 2. The transition matrix for the state space {1,2} is now

$$P = \begin{bmatrix} 0.95 & 0.05 \\ 0.33 & 0.67 \end{bmatrix}.$$

The corresponding Markov chain is irreducible. The percentage of days when the bike is `broken` is $\pi(\texttt{broken})$, where π is the invariant distribution. We solve the invariant distribution from the balance equations:

$$\pi = \pi P \qquad \text{and} \qquad \sum_i \pi(x_i) = 1,$$

so that

$$\begin{cases} 0.95\pi_1 + 0.33\pi_2 = \pi_1 \ \Rightarrow \ \pi_2 = 5/33\pi_1 \\ 0.05\pi_1 + 0.67\pi_2 = \pi_2 \\ \pi_1 + \pi_2 = 1. \end{cases}$$

The first and the second equations are equivalent, which we may verify by substituting the former to the latter. Substituting the first row to the third yields $\pi_1 = 33/38$, so $\pi_2 = 5/38$. The invariant distribution, and hence the limiting distribution, is

$$\pi = [33/38, 5/38] \approx [0.8684, 0.1316].$$

In the long run the bike is broken on about 13.2% of the days.

Additional information. Numerically: Since the Markov chain is irreducible and aperiodic, the state distributions converge to the limiting distribution π, irrespectively of the initial state. The limiting distribution can be deduced by computing large powers of the transition matrix:

$$P^{100} \approx \begin{bmatrix} 0.8684 & 0.1316 \\ 0.8684 & 0.1316 \end{bmatrix}.$$

Up to numerical accuracy, $P^{101} = P^{100}$, so the state distribution has ended up in the equilibrium $[0.8684, 0.1316]$ irrespectively of the initial state. We see that in the long run the bike is broken on approximately 13.2 % of the days.

Additional information. In the long run the relative times spent in the states are given by the invariant distribution. Formally: Let $(X_t)_{t\in\mathbb{N}}$ be an irreducible Markov chain on a finite state space, and let $N_T(y)$ be the visit count at y by time T, i.e.,

$$N_T(y) = \sum_{t=0}^{T} \mathbb{I}\{X_t = y\}.$$

For all states y and all distributions μ_0 on the initial states,

$$\lim_{T\to\infty} \mathbb{E}\left(\frac{N_T(y)}{T+1}\right) = \lim_{T\to\infty} \frac{1}{T+1}(\mu_0 P^t)_y = \pi(y),$$

where π is the unique invariant distribution of the chain. Thus, the *expected* relative times spent in each state are given by the invariant distribution. The word "expected" can be removed, however, because of the stronger result:

$$\lim_{T\to\infty} \frac{N_T(y)}{T+1} = \pi(y)$$

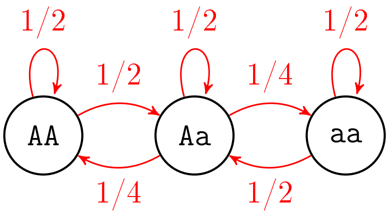
with probability 1, where π is the unique invariant distribution of the chain.

(b) Consider the Markov chain of Problem 2.3 with state space $S = \{\mathtt{AA}, \mathtt{Aa}, \mathtt{aa}\}$ and transition matrix

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{bmatrix}.$$

Calculate the proportion of each genotype in the chain of descendants in the long term.

Solution. The transition diagram is



where we see that the chain is irreducible. Therefore, in the long run the relative frequencies of the genotypes are given by the unique invariant distributions π. This may be justified as in (a).

The balance equations are

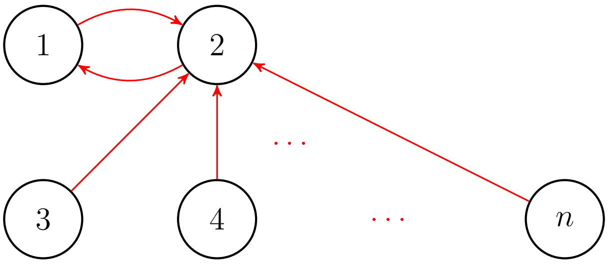
$$\begin{cases} 1/2\pi_1 + 1/4\pi_2 = \pi_1 & \Leftrightarrow \quad \pi_2 = 2\pi_1 \\ 1/2\pi_1 + 1/2\pi_2 + 1/2\pi_3 = \pi_2 & \\ 1/4\pi_2 + 1/2\pi_3 = \pi_3 & \Leftrightarrow \quad \pi_2 = 2\pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 & \end{cases}$$

and from the right-hand side we immediately see that $\pi = [1/4, 1/2, 1/4]$.

3.3 *PageRanks of nodes with high and low indegrees.* Consider a directed graph defined on node set $V = \{1, 2, \dots, n\}$ that has the links $1 \to 2$, $2 \to 1$, and $x \to 2$ for $x = 3, 4, \dots, n$. Let $(X_0, X_1, \dots)$ be a Markov chain that follows the PageRank algorithm for this graph, as discussed in the lecture notes (Example 1.4).

(a) Draw the transition diagram of the graph and determine for which values of the damping factor c the Markov chain is irreducible.

Solution. The graph V is



The PageRank transition probabilities are

$$P(x,y) = c\frac{1}{n} + (1-c)\frac{G(x,y)}{\sum_{y' \in V} G(x,y')},$$

where G is the adjacency matrix of V and $c \in [0,1]$ is a free parameter.

The graph drawn above is also the transition diagram of the PageRank Markov chain, when $c = 0$. The weights of the links are all 1. When $c > 0$, there are links between all pairs of nodes, to both directions. In this case, the weights of the links in the above picture are $c/n + (1-c)$, and all other weights are c/n. The chain is irreducible for $c > 0$.

(b) Compute the PageRanks for the nodes in the graph.

Solution. The balance equations are

$$\begin{aligned}
\pi(1) &= \pi(1)cn^{-1} + \pi(2)\left(cn^{-1} + (1-c)\right) + \left(\sum_{x=3}^{n} \pi(x)\right) cn^{-1} \\
\pi(2) &= \pi(1)\left(cn^{-1} + (1-c)\right) + \pi(2)cn^{-1} + \left(\sum_{x=3}^{n} \pi(x)\right) \left(cn^{-1} + (1-c)\right) \\
\pi(3) &= \left(\sum_{x=1}^{n} \pi(x)\right) cn^{-1} \\
&\vdots \\
\pi(n) &= \left(\sum_{x=1}^{n} \pi(x)\right) cn^{-1}
\end{aligned}$$

$$\sum_{x=1}^{n} \pi(x) = 1.$$

With the normalization condition $\sum_x \pi(x) = 1$, these equations simplify to

$$\begin{aligned}
\pi(1) &= cn^{-1} + \pi(2)(1-c), \\
\pi(2) &= cn^{-1} + (1-\pi(2))(1-c), \\
\pi(3) &= cn^{-1}, \\
&\vdots \\
\pi(n) &= cn^{-1}.
\end{aligned}$$

From here we can solve

$$\begin{aligned}
\pi(1) &= \left(1 + \frac{1-c}{1+(1-c)}\right) cn^{-1} + \frac{(1-c)^2}{1+(1-c)}, \\
\pi(2) &= \frac{cn^{-1} + (1-c)}{1+(1-c)}, \\
\pi(3) &= cn^{-1}, \\
&\vdots \\
\pi(n) &= cn^{-1}.
\end{aligned}$$

(c) How do the PageRanks behave when $c = 0$ and $c = 1$?

Solution. For $c = 0$ we get $\pi(1) = \pi(2) = \frac{1}{2}$, and $\pi(x) = 0$ for $x \geq 3$. For $c = 1$ we get $\pi(x) = 1/n$ for all x.

(d) How do the PageRanks behave when $n \to \infty$?

Solution. In the limit $n \to \infty$ we get

$$\begin{aligned}
\pi(1) &= \frac{(1-c)^2}{1+(1-c)}, \\
\pi(2) &= \frac{(1-c)}{1+(1-c)}, \\
\pi(j) &= 0, \quad \text{for } j \geq 3.
\end{aligned}$$

Note that for the limits it holds that

$$\pi(1) + \pi(2) = 1 - c,$$

so that $[\pi(1), \pi(2)]$ is not a probability distribution on the state space $\{1, 2\}$. Instead, a probability mass c has "escaped to infinity" in the limit process. It is seen that the pointwise limits of probability distributions do not necessarily give a probability distribution.