Exercises # 4: Markov Chains

Léo Neufcourt

October 4, 2019

Each problem is worth 10 points, points above 50 are bonus. Due at the beginning of class on October 4. Python codes must be sent by email before October 4, 12:40pm.

Exercise 1. *Consider a Markov chain with transition matrix*

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

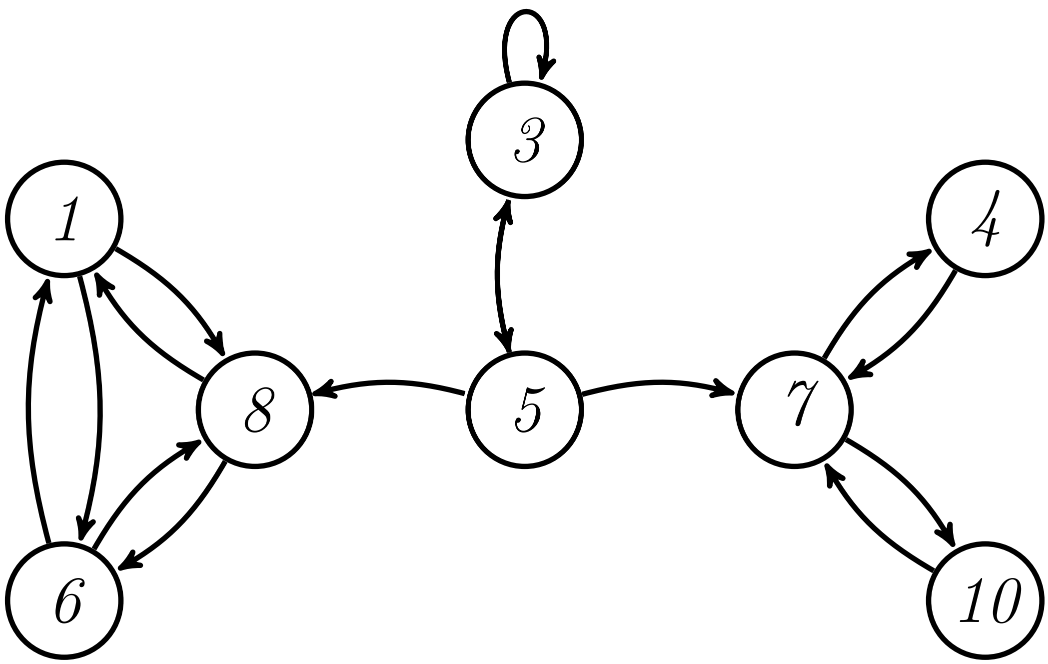
a) Represent the graph of this Markov chain and determine its communication classes, their nature (recurrent or transient) and their periodicity.

b) Determine the stationary probabilities of this Markov chain.

c) Is there a limiting distribution?

d) How much time does the process spend in average in each of the states (at the limit where the time $n \to \infty$)?

Solution 1.



a) *This is a Markov chain on a finite state space, with five communication classes: two recurrent classes,* $R_1 := \{1, 6, 8\}$*, aperiodic; and* $R_2 := \{4, 7, 10\}$*, with period two; and three transient classes,* $T_1 := \{9\}$*,* $T_2 = \{2\}$ *and* $T_3 = \{3, 5\}$ *(aperiodic).*

b) *Let* π *be a stationary distribution of the Markov chain, if it exists. First we must have* $\pi_2 = \pi_3 = \pi_5 = \pi_9 = 0$ *since these transient states. Then, for a Markov chain on a recurrent classes* R_1 *or* R_2*, there exists unique stationary probabilities, which we respectively denote* $\pi^{(1)}$ *and* $\pi^{(2)}$*. The stationary distributions for the Markov process will be all the distributions of the form* $\pi = \lambda\pi^{(1)} + (1 - \lambda)\pi^{(2)}$*, for some* $0 \leq \lambda \leq 1$*.*

Now let us find the stationary distributions $\pi^{(1)}$ *and* $\pi^{(2)}$ *on respectively* R_1 *and* R_2*. On* R_1*, symmetry yields*

$$\pi_1^{(1)} = \pi_6^{(1)} = \pi_8^{(1)} = \frac{1}{3}.$$

On R_2, solving $\pi^{(2)} P_{R_2} = \pi^{(2)}$ with $\sum_{i=1}^{10} \pi_i^{(2)} = 1$ yields

$$\pi_4^{(2)} = \pi_{10}^{(2)} = \frac{1}{2}\pi_7^{(2)} = \frac{1}{4}.$$

We can conclude that the stationary distributions are:

$$\pi_\lambda = \left(\frac{\lambda}{3}, 0, 0, \frac{1-\lambda}{4}, 0, \frac{\lambda}{3}, \frac{1-\lambda}{2}, \frac{\lambda}{3}, 0, \frac{1-\lambda}{4}\right), 0 \leq \lambda \leq 1.$$

c) In general there is no limiting distributions, since the recurrent class R_2 is periodic; however, if the process starts in the ergodic class R_1, it will converge in distribution to $\pi^{(1)}$.

d) There is no unique stationary distribution here that we can use it to determine immediately the average time spent in each of the states. First, the process will spend a finite time in transient states, hence will have a limiting probability 0 to be in the transient states. The time spent in recurrent states depends where the process starts initially: if the process starts in the class R_1, it will always stay in R_1 and spend in the long run average proportions $\pi^{(1)}$ of its time in the respective states of R_1; if the process starts in the class R_2, it will always stay in R_2 and spend in the long run average proportions $\pi^{(2)}$ of its time in the respective states of R_2; now if the process starts in the transient class, if will enter one of the recurrent classes after a finite number of steps, and stay there forever: this will be R_1 with probability $\frac{p_{5,8}}{p_{5,8}+p_{5,7}} = \frac{2}{3}$ and R_2 with probability $\frac{1}{3}$.

Exercise 2. *A kangaroo jumps between five points on a circle, At every step he jumps from its location to one of the two neigboring points on the circle with probability* 0.5*. Show that the locations of the kangaroo at each step compose a Markov chain and provide its state space, graph and transition matrix; determine the communication classes as well as the period and nature (recurrent or transient) of all states. How much time does the kangaroo spend in average in each of the states (at the limit where the time $n \to \infty$)?*

Solution 2. *Clearly at every time the future moves of the kangaroo depend only in its current locations, hence its locations form a Markov chain on the* *state space $S = \{1,2,3,4,5\}$. The transitions matrix is:*

$$P = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

The chain is irreducible; it is also aperiodic because it is possible to go from state 1 to state 1 in two steps and in five steps with the paths $1 \to 2 \to 1$ and $1 \to 2 \to 3 \to 4 \to 5 \to 1$, so any period d should divide 2 and 5 and there is no such number $d \geq 2$.

Since it has a finite state space the chain is positive recurrent and ergodic, and it follows from the symmetries of the matrix that the unique stationary distribution is the uniform distribution $\pi = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$.

Thus the kangaroo will spend in average $\frac{1}{5}$ of the time at each locations. Note that we know additionally that this ergodic Markov chain converges to its stationary distribution, i.e. the probability that after any long time we find the kangaroo in each of the locations is $\frac{1}{5}$.

Exercise 3. *You commute between home and office and you have four umbrellas. Every time, if it rains you take your umbrella; if it doesn't rain you leave the umbrella behind (at home or in the office). It may happen that all umbrellas are in one place, you are at the other, it starts raining and you must leave: in that case, you get wet.*

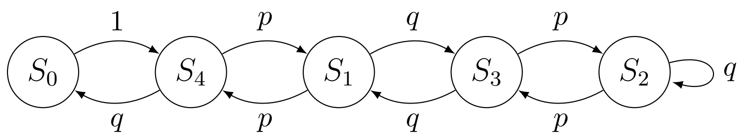
1. *Show that the number of umbrellas that are in the same location as you are at the times when you need to commute is a Markov chain, draw its graph and give its transition probabilities.*

2. *What are the stationary distributions?*

3. *If the probability of rain is p, what is the probability that you get wet?*

4. *(Bonus) Current estimates show that $p = 0.6$ in East Lansing. How many umbrellas should you have so that, following the strategy described above, the probability that you get wet is less than 10%?*

Solution 3.



1. *Let us consider a Markov chain X taking values in the set $S = \{0, 1, 2, 3, 4\}$ and representing the number of umbrellas in the place where you are currently at (home or office). For instance, if at some time n there are $X_n = 1$ umbrella(s) at your location and it rains, then you take the umbrella, commute to the other location where there are already 3 umbrellas, so that including the one you are bringing bring you will have next $X_{n+1} = 4$ umbrellas. Thus $p_{1,4} = p$ which is the probability of rain. If $X_n = 1$ but it does not rain then you do not take the umbrella and you go to the other place where you will find 3 umbrellas. Thus, $p_{1,3} = 1 - p =: q$. Continuing in the same manner we obtain the transition matrix*

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & q & p \\ 0 & 0 & q & p & 0 \\ 0 & q & p & 0 & 0 \\ q & p & 0 & 0 & 0 \end{bmatrix}$$

2. *The stationary distributions are the row matrices π solutions of $\pi P = \pi$ with $\sum_i \pi_i = 1$. This leads first to $\pi_0 = q\pi_4$. The remaining equations, once π_0 has been replaced by $q\pi_4$, are symmetric in π_1, π_2, π_3 and π_4, so that $\pi_1 = \pi_2 = \pi_3 = \pi_4 = \frac{1}{4+q}$ and $\pi_0 = \frac{q}{4+q}$.*

3. *You will get wet every time you happen to be in state 0 and it rains. The long run-probability to be in state 0 is π_0, and independently the probability that it rains is p. Hence $\mathbb{P}(wet) = \pi_0 p = \frac{pq}{q+4}$.*

4. *With $p = 0.6$ and $q = 0.4$, we have $\mathbb{P}(wet) \approx 5.45\%$ If you want the chance to be less than 1% then clearly you need more umbrellas. Suppose that you have k umbrellas. Seting up the Markov chain as above, it is clear that the unique stationary distribution is $\pi_1 = \pi_2 = ... = \pi_k = \frac{1}{q}\pi_0$ whence $\pi_0 = \frac{q}{q+N}$ and $P(wet) = \frac{pq}{q+N}$. In order to have $P(wet) \leq 1\%$ you just need*

to take

$$N \geq 100pq - q = 23.6.$$

In order to reduce the chance of getting wet from 6% to less than 1% you will need 24 umbrellas instead of 4. It is certainly cheaper to get wet.

Exercise 4. *Let S_n be a simple (symmetric and one-dimensional) random walk on $\mathbb{Z}$ and let $N_i := \inf\{n \geq 1 : S_n = 0\}$. What is the expectation of the number of times that S, starting at 0, will visit a given state $i \in \mathbb{Z}, i > 0$, before it will come back to 0 for the first time?*

Hint: *The probability that, starting at 0, the process will hit i exactly k times before coming back to 0 for the first time is the probability that the process will first, starting at 0, hit i without hitting 0; then, starting at i, return to i without hitting 0, $k-1$ times; and finally, starting at i, hit 0 without hitting i.*

Solution 4. *Let us define $T_{x,y}$ to be the first time a simple random walk reaches state y, starting in state x. Following the hint we have*

$$\begin{aligned}\mathbb{P}(N_i = k) &= P(T_{0,i} < T_{0,0})P(T_{i,i} < T_{i,0})^{k-1}P(T_{i,0} < T_{i,i}) \\ &= P(T_{0,i} < T_{0,0})P(T_{0,0} < T_{0,i})^{k-1}P(T_{0,i} < T_{0,0})\end{aligned}$$

after using that $P(T_{i,i} < T_{i,0}) = P(T_{0,0} < T_{0,i})$ and $P(T_{i,0} < T_{i,i}) = P(T_{0,i} < T_{0,0})$ thanks to symmetries. Now for $i > 0$ we have $S_1 = -1$ implies $T_{0,0} < T_{0,i}$, thus $P(T_{0,i} < T_{0,0}) = P(T_{0,i} < T_{0,0}, S_1 = 1) = \frac{1}{2}P(T_{0,i} < T_{0,0}, S_1 = 1)$. The last term is given by the next problem (symmetric case): we are looking at whether a gambler starting with \$1 will reach a fortune of \$$i$ or become ruined. Hence, $P(T_{0,0} < T_{0,i}) = \frac{1}{i}$ so that

$$\mathbb{P}(T_{0,0} = k) = \frac{1}{2k}.$$

Finally

$$E[N_i] = \sum_{k=1}^{\infty} k \frac{1}{(2i)^2}(1 - \frac{1}{2i})^{k-1} = 1,$$

i.e. the process will make in average one visit to i before returning to 0.

Exercise 5. *Consider a gambler who starts with an initial fortune of \$$x$ and then places independently successive bets, at each of which he wins or loses*

\$1 *with probabilities* p *and* $q := 1-p$*, respectively. The gambler stops playing when (if and only if) his fortune reaches* \$0 *or* \$$y$*,* $y > x$*. Let* S_n *denote the total fortune after the* n^{th} *bet,* $T := (x,y) := \inf\{n \geq 0 : S_n = 0 \text{ or } S_n = y\}$ *the time when the game stops, and let* $\phi(x) := \phi(x,y) := P(S_T = y)$*.*

1. *Show that* S *is a Markov chain and give its states, transition probabilities, and communication classes as well as classes nature (recurrent or transient) and periodicity.*

2. *Find a recursion on* $\phi(x)$*.*

3. *If* $p = q = \frac{1}{2}$*, show that* $\phi(x) = \frac{x}{y}$*.*

4. *What is* $\phi(x)$ *when* $p \neq q$*?*

5. *Does the Markov chain* S *have any stationary distribution? any limiting distribution?*

6. *Consider the alternative strategy where the gambler's bet all his wealth* x *at the first bet. Which strategy is best? Does this depend on* x*,* y*, or* p*?*

Solution 5.

1-4. *See textbook for this classical problem.*

5. *To compare the two strategies let us consider the case* $x = y$*. When* $p = \frac{1}{2}$ *(fair game) both strategies have the same expected outcome – and the same risk; if* $p > \frac{1}{2}$ *the gambler should follow the first strategy consisting in incremental* \$1 *bets, and earn* \$$x$ *with probability 1 in the long run; if* $p < \frac{1}{2}$ *the gambler should use the second strategy consisting in a single* \$$x$ *bets and earn* \$$x$ *with probability* p*.*

Exercise 6. *Consider the simple random walk on the signed integers (as defined in class). Is there any stationary distribution? Is there any limiting distribution?*

Solution 6. *If there were a stationary distribution, then the stationary probabilities shall be the same for all states due to the symmetry of the transition probabilities (both in the symmetric case and non-symmetric case). As there is an infinite number of states this is simply impossible. Since a limiting distribution must be a stationary distribution there can be no limiting distribution, either.*