STAT516: Stochastic Modeling of Scientific Data I	Fall 2021

Homework 2

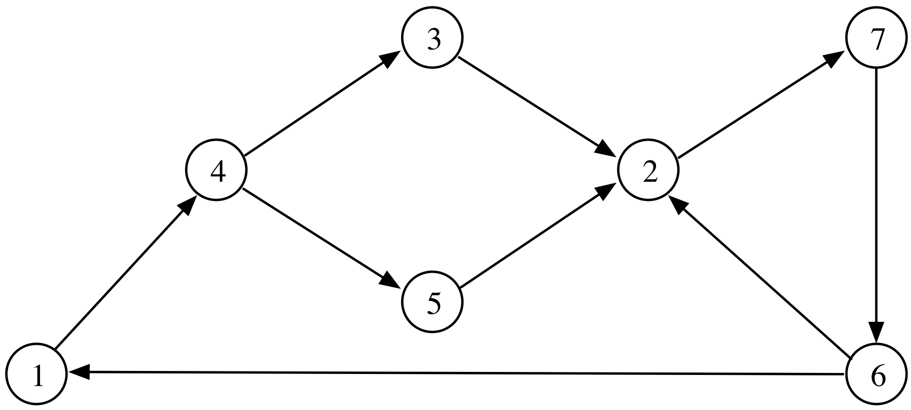
Due **October 19th** by 11:59pm

Instructor: Vincent Roulet *Grader: Alex Jiang*

Upload your answers to the following question on Gradescope in a pdf.
Make a new page for each exercise that will facilitate the grading.
For the coding questions, you can use the programming language of your choice, e.g. Python or R.

State classification

Exercise 1. Show that the transition graph below is irreducible. Find its period and its cyclic classes.



Solution. **It is irreducible**. In this diagram, we see loops 1-4-3-2-7-6-1 and 1-4-5-2-7-6-1. Thus $\{1, 4, 3, 2, 7, 6, 1\}$ and $\{4, 5\}$ are communicative. Moreover, let state 2 be the reference point, we notice that $\left\{n : p_{22}^{(n)} > 0\right\} = \{n \mid n = 3k, k \in \mathbb{N}^+\}$, and thus $d_2 = 3$. Since this is an irreducible chain, all states(and thus the MC chain) has a period of 3. Let state 1 be the reference point, notice that

$$p_{14}^{(1)} > 0, p_{15}^{(2)} > 0, p_{13}^{(2)} > 0, p_{12}^{(3+0)} > 0, p_{17}^{(1+3)} > 0, p_{16}^{(2+3)} > 0$$

According to the lattice theorem, the cylic classes are:

$$C_0 = \{1, 2\}, C_1 = \{4, 7\}, C_2 = \{3, 5, 6\}$$

■

Exercise 2. Prove that recurrence is a communication class property: $i \leftrightarrow j$ and i is recurrent $\Rightarrow j$ is recurrent.

Hint: Use the characterization of recurrence as the fact that $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$ and the Chapman Kolmogorov equation.

Solution. If i communicates with j, there exists $k, l \in \mathbf{Z}$ such that $p_{ij}^{(k)} := \alpha > 0$ and $p_{ji}^{(l)} := \beta > 0$. Since i is a recurrent state, $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$.

For $\forall n \geq k + l$, we have

$$\begin{aligned}
p_{jj}^{(n)} &= P(X_n = j | X_0 = j) \\
&\geq P(X_n = j, X_{n-k} = i, X_l = i | X_0 = j) \\
&= P(X_n = j | X_{n-k} = i) P(X_{n-k} = i | X_l = i) P(X_l = i | X_0 = j) \quad \text{(By Markov property)} \\
&= p_{ji}^{(l)} p_{ij}^{(k)} p_{ii}^{(n-k-l)} = \alpha\beta p_{ii}^{(n-k-l)}
\end{aligned}$$

Thus

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} \geq \sum_{n=k+l+1}^{\infty} p_{jj}^{(n)} \geq \sum_{n=k+l}^{\infty} \alpha\beta p_{ii}^{(n+1-k-l)} = \alpha\beta \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$$

Thus $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ and we conclude from proposition 3.10 that state j is a recurrent state. ∎

Recurrence

Exercise 3. Consider $(X_n)_{n\in\mathbb{N}}$ to be a random walk on the set of all integers $\mathbb{Z}$ i.e.

$$p_{ij} = \begin{cases} p & \text{if } j = i+1 \\ 1-p & \text{if } j = i-1. \end{cases}$$

We want to study the recurrence of this chain, which amounts to study the recurrence of the state e.g. 0 since the chain is irreducible.

1. What is $p_{00}^{(2n+1)}$ for any n?

   Solution. If the random walk is initiated at $X_0 = 0$, this means we cannot return to it after an odd number of steps. So $p_{00}^{(2n+1)} = 0$ for any n. ∎

2. Now, note that $X_{2n} = \xi_1 + \ldots + \xi_{2n}$ where ξ_i are i.i.d. r.v. s.t. $\mathbb{P}(\xi_i = 1) = 1 - \mathbb{P}(\xi_i = -1) = p$. Deduce an expression for $p_{00}^{(2n)}$ as the number of ways to move n times forward ($\xi_i = 1$) and n times backward ($\xi_i = -1$).

   Solution. We know that if $X_{2n} = 0$ given $X_0 = 0$, there need to be n 1s and n -1-s in the sequence $\{\xi_i\}_{i=1}^{2n}$. Thus we have

   $$p_{00}^{(2n)} = P\left(X_{2n} = 0 \mid X_0 = 0\right) = \left(\begin{array}{c} 2n \\ n \end{array}\right) p^n q^n = \frac{(2n)!}{n!n!} p^n q^n$$

   ∎

3. Use the ratio test explained below to conclude whether the state 0 is transient/recurrent depending on the value of p (use the characterization of the recurrence in terms of the n-steps transition probabilities). The Stirling formula $n! \sim n^{n+\frac{1}{2}} e^{-n}\sqrt{2\pi}$ can be useful.

   Solution. By the Stirling formula,

   $$p_{00}^{(2n)} = \frac{(2n)!}{n!n!} p^n q^n \sim \frac{(2n)^{2n+\frac{1}{2}} e^{-2n}\sqrt{2\pi}}{n^{2n+1} e^{-2n} 2\pi} (pq)^n = \frac{(4pq)^n}{\sqrt{\pi n}}$$

   From (a) we know $p_{00}^{(2n+1)} = 0$, thus if state 0 is recurrent, $\sum_{n=1}^{\infty} p_{00}^{(n)} = \sum_{n=1}^{\infty} p_{00}^{(2n)} = \infty$. From the ratio test, we know that $\lim_n |\frac{p_{00}^{2n+2}}{p_{00}^{2n}}| = (4p(1-p))^2$, thus the series $\sum_{n=1}^{\infty} p_{00}^{(n)} = \sum_{n=1}^{\infty} p_{00}^{(2n)}$ is convergent when $p \neq \frac{1}{2}$ and thus the state is transient. Also, by lemma 1 in lecture notes 4, we see that only when $p = q = \frac{1}{2}$, we have that $\frac{(4pq)^n}{\sqrt{\pi n}} = \frac{1}{\sqrt{\pi n}} \sim \frac{c}{\sqrt{n}}$. So it is divergent if and only if $p = q = \frac{1}{2}$ and in this case the state is recurrent. ∎

4. Finally, can the state 0 be positive recurrent? (This will be clarified in lecture 5, so you can wait until Thursday to treat this question)

   Solution. From lecture 5 we know that it cannot be positive recurrent since the measure space is not normalizable and we cannot find an invariant measure that can be normalized. ∎

Ratio test: Let $\sum_{n=1}^{\infty} a_n$ be a series which satisfies $\lim_{n\to\infty} |\frac{a_{n+1}}{a_n}| = c$. If $c > 1$ the series diverges, if $c < 1$ the series converges.)

Note: For two sequences $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ we say that they are equivalent as $n \to +\infty$, denoted $u_n \sim v_n$, if $\lim_{n\to+\infty} u_n/v_n = 1$.

Transience

Exercise 4. Let $(X_i)_{i\in\mathbb{N}}$ be a homogeneous Markov chain with state space S and n-step transition probability matrix $\mathbf{P}^n = (p_{ij}^{(n)})_{i,j\in S}$. Prove that if $i \in S$ is a transient state, then $\sum_{n=1}^{\infty} p_{ji}^{(n)} < \infty$ which implies that $\lim_{n\to\infty} p_{ji}^{(n)} = 0$ for all $j \in S$.

Hint: Relate $\sum_{n=1}^{\infty} p_{ji}^{(n)}$ to $\mathbb{E}_j(N_i)$ and express the later using $\mathbb{P}_j(N_i > k)$ which is given in the proof of Proposition 4.5 in the lecture note 4. Conclude using that i is transient.

Solution. By definition,

$$\begin{aligned}
\mathbb{E}_j[N_i] &= \mathbb{E}\left(N_i \mid X_0 = j\right) = \mathbb{E}\left(\sum_{n=1}^{\infty} I(X_n = i) \mid X_0 = j\right) \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left(I(X_n = i) \mid X_0 = j\right) = \sum_{n=1}^{\infty} \mathbb{P}\left(X_n = i \mid X_0 = j\right) \\
&= \sum_{n=1}^{\infty} p_{ji}^{(n)}
\end{aligned}$$

Also, from appendix A in lecture notes 4, we have

$$\mathbb{E}_j\left[N_i\right] = \sum_{t=0}^{\infty} \mathbb{P}_j\left[N_i > t\right] = \sum_{t=0}^{\infty} f_{ji} f_{ii}^t = \frac{f_{ji}}{1 - f_{ii}} < \infty$$

where $f_{ji} = \mathbb{P}_j\left(T_i < \infty\right) = \mathbb{P}_j\left(N_i > 0\right)$ and $f_{ii} = \mathbb{P}_i\left(T_i < \infty\right)$. Since i is transient, $f_{ii} < 1$ and the series is convergent.

Thus $\sum_{n=1}^{\infty} p_{ji}^{(n)} < \infty$ and $\lim_{n\to\infty} p_{ji}^{(n)} = 0$ for all $j \in S$.

■

Simulation

Exercise 5. The goal of this exercise is to visualize a bit a problem with absorbing states, namely, the Gambler's ruin problem described in the previous homework. We fix the total amount of money at play in the game to be e.g. $c = 10$. Fix e.g. $X_0 = 3$ and consider a t.p.m. s.t. $p_{00} = 1, p_{cc} = 1$ and for $i \notin \{0, c\}$, $p_{ii+1} = p$, $p_{ii-1} = 1 - p$ for e.g. $p = 0.5$.

1. Plot a random realization of the Markov chain, i.e., the y-axis will be the state (between 0 and 10) and the x-axis will be the time. (simply stop when the chain reaches one of the absorbing states).

2. Use your simulation routine to estimate the probability of reaching the largest state $c = 10$ starting at state 3, denoted $h(3, p)$, for probabilities $p_{i,i+1} = p \in \{0.1, 0.2, \ldots, 0.9\}$. Turn in a graph with estimated $h(3, p)$ plotted against p.

Useful lemma

Lemma 1. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. If $u_n \sim c/\sqrt{n}$ as $n \to +\infty$ for some constant $c > 0$ then $\sum_{n=0}^{+\infty} u_n = +\infty$.

Proof. By definition of $u_n \sim c/\sqrt{n}$ as $n \to +\infty$, there exists $n_0 \in \mathbb{N} \setminus \{0\}$ s.t. for any $n \geq n_0$,

$$\left| \frac{u_n}{c/\sqrt{n}} - 1 \right| \leq 1/2.$$

Hence for any $n \geq n_0$,

$$u_n \geq \frac{c}{2\sqrt{n}}$$

Now since $1/\sqrt{x}$ is decreasing, we have that for any $n \geq 1$ and any $x \in [n, n+1]$, $1/\sqrt{x} \leq 1/\sqrt{n}$. Therefore for any $n \geq 1$,

$$\frac{1}{\sqrt{n}} \geq \int_n^{n+1} \frac{1}{\sqrt{x}} dx$$

So we get that for any $N \geq n_0$,

$$\sum_{n=n_0}^{N} \frac{1}{\sqrt{n}} = \int_{n_0}^{N+1} \frac{1}{\sqrt{x}} dx = \left[2\sqrt{x}\right]_{n_0}^{N+1} = 2\sqrt{N+1} - 2\sqrt{n_0}$$

Therefore for any $N \geq n_0$,

$$\sum_{n=0}^{N} u_n = \sum_{n=0}^{n_0-1} u_n + \sum_{n=n_0}^{N} u_n \geq \sum_{n=0}^{n_0-1} u_n + c\sqrt{N+1} - c\sqrt{n_0}$$

Hence for any $A \gg 1$, denoting $N_A \geq n_0$ such that $\sum_{n=0}^{n_0-1} u_n + c\sqrt{N_A+1} - c\sqrt{n_0} = A$, we have that for any $N \geq N_A$,

$$\sum_{n=0}^{N} u_n \geq \sum_{n=0}^{n_0-1} u_n + c\sqrt{N+1} - c\sqrt{n_0} \geq A.$$

This shows that $\lim_{N\to+\infty} \sum_{n=0}^{N} u_n = +\infty$.

$\square$