Stoch. Proc. HW assignment 5. Due Friday, October 13 at start of class

1. Let us consider the Markov chain with state space $S = \{1, 2, 3, 4, 5\}$ and transition matrix

$$\underline{\underline{P}} = \begin{pmatrix} 1/3 & 0 & 1/3 & 1/3 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1/2 & 1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 2/3 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

   (a) Partition S into irreducible components and find out which state is recurrent and which one is transient.

   (b) Characterize the family of stationary distributions of this Markov chain. Check by hand that the distributions you have found are indeed stationary.

   (c) Starting from state 1, what is the probability of reaching state 2 before state 5?

   (d) Find $\lim_{n\to\infty} \underline{\underline{P}}^n$. *Hint:* Use the ideas from parts (b) and (c).

Solution:

(a) The irreducible components are: $\{1, 3\}$, $\{2\}$ and $\{4, 5\}$. The states in the component $\{1, 3\}$ are transient, while the states in the components $\{2\}$ and $\{4, 5\}$ are recurrent, since these irreducible components are *closed*, i.e., there is no way out of them. The state 2 is an *absorbing state* of the Markov chain: once the walker reaches state 2, it gets stuck there forever.

(b) First we have to find the stationary distributions of the Markov chain if we restrict it to the irreducible components.

The stationary distribution $\pi^{\{2\}}$ corresponding to the absorbing state 2 is the Dirac mass concentrated on the state 2, so the vector

$$\pi^{\{2\}} = (0, 1, 0, 0, 0)$$

satisfies $\pi^{\{2\}} \underline{\underline{P}} = \pi^{\{2\}}$, i.e., it is a stationary distribution.

The stationary distribution $\pi^{\{4,5\}}$ corresponding to the closed irreducible component $\{4, 5\}$ is of form $\pi^{\{4,5\}} = (0, 0, 0, a, b)$ where (a, b) is a left eigenvector corresponding to the eigenvalue 1 of the stochastic matrix $\begin{pmatrix} 1/3 & 2/3 \\ 1 & 0 \end{pmatrix}$. That is, $a = \frac{1}{3}a + b$ and $b = \frac{2}{3}a$, thus $a = \frac{3}{5}$, $b = \frac{2}{5}$, so the vector

$$\pi^{\{4,5\}} = (0, 0, 0, \frac{3}{5}, \frac{2}{5})$$

satisfies $\pi^{\{4,5\}} \underline{\underline{P}} = \pi^{\{4,5\}}$, i.e., it is a stationary distribution.

By the theorem stated on page 69 of the scanned lecture notes, if π is a stationary distribution of $\underline{\underline{P}}$ then π is of form

$$\pi = (1 - \alpha)\pi^{\{2\}} + \alpha \pi^{\{4,5\}} = (0, 1 - \alpha, 0, \frac{3}{5}\alpha, \frac{2}{5}\alpha), \qquad \alpha \in [0, 1].$$

To check by hand that for any $\alpha \in [0, 1]$ we indeed obtain a probability distribution, note that the coordinates of π are indeed non-negative and indeed sum up to 1. Moreover, π is indeed stationary:

$$\pi \underline{\underline{P}} = (1 - \alpha)\pi^{\{2\}} \underline{\underline{P}} + \alpha \pi^{\{4,5\}} \underline{\underline{P}} = (1 - \alpha)\pi^{\{2\}} + \alpha \pi^{\{4,5\}} = \pi.$$

(c) Let $h(x) = P_x(T_{\{2\}} < T_{\{4,5\}})$. In words, let $h(x)$ be the probability of reaching $\{2\}$ before reaching $\{4, 5\}$ if we start from x. Of course we have $h(2) = 1$ and $h(4) = h(5) = 0$, moreover

$$h(1) = \frac{1}{3}h(1) + \frac{1}{3}h(3) + \frac{1}{3} \cdot 0,$$
$$h(3) = \frac{1}{2}h(1) + \frac{1}{4} \cdot 1 + \frac{1}{4}h(3).$$

Solving this system of linear equations we obtain $h(1) = \frac{1}{4}$ and $h(3) = \frac{1}{2}$. So the answer is $\frac{1}{4}$.

(d) Let $\underline{\underline{Q}} = (q(x,y))_{x,y\in S}$, where $\underline{\underline{Q}} := \lim_{n\to\infty} \underline{\underline{P}}^n$. If $x, y \in S = \{1,\ldots,5\}$ then $q(x,y)$ is the probability that we are in state y at time n if we start from x and n is very large.

After a long time, we will certainly leave the transient states (i.e., 1 and 3) behind and we are going to be in one of the closed irreducible components (i.e., either $\{2\}$ or $\{4,5\}$).

Moreover, if we ended up in a particular irreducible component then we have spent so much time there that we are already in the stationary state corresponding to that component (note that all closed irreducible components of our $\underline{\underline{P}}$ are also aperiodic, so convergence to stationarity does hold if we are inside a particular irreducible component).

Putting these observations together we obtain that

$$q(x,y) = P_x(T_{\{2\}} < T_{\{4,5\}})\pi^{\{2\}}(y) + P_x(T_{\{2\}} > T_{\{4,5\}})\pi^{\{4,5\}}(y) = h(x)\pi^{\{2\}}(y) + (1-h(x))\pi^{\{4,5\}}(y).$$

Writing this in matrix form we obtain

$$\lim_{n\to\infty} \underline{\underline{P}}^n = \underline{\underline{Q}} = \begin{pmatrix} 0 & 1/4 & 0 & 9/20 & 3/10 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 3/10 & 1/5 \\ 0 & 0 & 0 & 3/5 & 2/5 \\ 0 & 0 & 0 & 3/5 & 2/5 \end{pmatrix}.$$

Remark: Note that each row of $\underline{\underline{Q}}$ is a stationary distribution of $\underline{\underline{P}}$, i.e., $\underline{\underline{Q}}\,\underline{\underline{P}} = \underline{\underline{Q}}$.

This is not surprising, since we have

$$\underline{\underline{Q}}\,\underline{\underline{P}} = \left(\lim_{n\to\infty} \underline{\underline{P}}^n\right) \underline{\underline{P}} = \lim_{n\to\infty} \underline{\underline{P}}^{n+1} = \underline{\underline{Q}}. \tag{1}$$

In other words, each row of $\underline{\underline{Q}}$ is a left eigenvector of $\underline{\underline{P}}$ corresponding to the eigenvalue 1.

Comparing with the solution of part (b), indeed for each row of $\underline{\underline{Q}}$ there exists an α for which that row is of form $(0, 1-\alpha, 0, \frac{3}{5}\alpha, \frac{2}{5}\alpha)$, namely $\alpha = P_x(T_{\{2\}} > T_{\{4,5\}})$ for the row indexed by $x \in \{1,\ldots,5\}$.

Also note that each column of $\underline{\underline{Q}}$ is a *harmonic function* on S with respect to $\underline{\underline{P}}$, where a function $H : S \to \mathbb{R}$ is harmonic if

$$H(x) = \sum_{y\in S} p(x,y)H(y), \qquad \forall\, x \in S.$$

In other words, if $(X_n)_{n\geq 0}$ is our Markov chain then $\mathbb{E}[H(X_1)\,|\,X_0 = x] = H(x)$.

For example the constant function $H(x) \equiv 1$ is a harmonic function and the function h that we have found in part (c) is also a harmonic function. Any linear combination of harmonic functions is a harmonic function. In fact, each column of $\underline{\underline{Q}}$ is a linear combination of the column vector $\mathbb{1} = (1,1,1,1,1)^T$ and the column vector $\underline{h} = (h(1),\ldots,h(5))^T$. To be more precise, the column of $\underline{\underline{Q}}$ indexed by $y \in S = \{1,\ldots,5\}$ can be written as $(\pi^{\{2\}}(y) - \pi^{\{4,5\}}(y))\underline{h} + \pi^{\{4,5\}}\mathbb{1}$.

A function H is harmonic if and only if the column vector $\underline{H} = (H(1),\ldots,H(5))^T$ satisfies $\underline{\underline{P}}\,\underline{H} = \underline{H}$.

A calculation similar to (1) shows that $\underline{\underline{P}}\,\underline{\underline{Q}} = \underline{\underline{Q}}$, so that each column of $\underline{\underline{Q}}$ is indeed a harmonic function.

In other words, each column of $\underline{\underline{Q}}$ is a right eigenvector of $\underline{\underline{P}}$ corresponding to the eigenvalue 1.

It is easy to check that if $|S| < +\infty$ and the Markov chain is irreducible then every harmonic function has to be a constant multiple of the trivial harmonic function $H(x) \equiv 1$.

More generally, the dimension of the vector space of harmonic functions is equal to the number of closed irreducible components of the state space.

2. Let $\underline{\underline{P}} = (p(x,y))_{x,y\in S}$ denote the transition matrix of an irreducible Markov chain (X_n) with finite state space S. Show that the following three statements are equivalent:

   (a) The Markov chain (X_n) is reversible.

   (b) There exists a function $\varphi : S \to (0,+\infty)$ satisfying $\varphi(x)p(x,y) = \varphi(y)p(y,x)$ for any $x, y \in S$.

   (c) For any $k \in \mathbb{N}$ and any $x_1, x_2, \ldots, x_k \in S$ we have

$$p(x_1,x_2)p(x_2,x_3)\ldots p(x_{k-1},x_k)p(x_k,x_1) = p(x_1,x_k)p(x_k,x_{k-1})\ldots p(x_3,x_2)p(x_2,x_1). \tag{2}$$

Solution: The equivalent characterization (c) of reversibility is called *Kolmogorov's cycle condition.* The condition (2) means that if $(x_1, x_2, \ldots, x_k, x_1)$ is cycle then the walker jumps along the edges of this cycle in the clockwise direction and in the anti-clockwise direction with the same probability. This condition is not surprising in view of the fact that the time-reversed process should have the same law as the original process. Intuitively, a Markov chain is reversible if there is no „turbulence" in the flow of mass as time evolves, and Kolmogorov's cycle condition manages to capture this property without any explicit reference to the stationary distribution.

Before we prove the equivalence of (a), (b) and (c), let us equivalently rewrite (c): For any $k \in \mathbb{N}$ and any $x_1, x_2, \ldots, x_k \in S$ if either the left-hand side or the right-hand side of (2) is positive then we have

$$\frac{p(x_1,x_2)}{p(x_2,x_1)}\frac{p(x_2,x_3)}{p(x_3,x_2)}\cdots\frac{p(x_{k-1},x_k)}{p(x_k,x_{k-1})}\frac{p(x_k,x_1)}{p(x_1,x_k)} = 1. \tag{3}$$

First we prove that (a) implies (c). We have seen in class (see page 41 of the scanned lecture notes) that reversibility is equivalent to the detailed balance condition, so (a) implies that $\pi(x)p(x,y) = \pi(y)p(y,x)$ for any $x, y \in S$, where π is the stationary distribution. Rearranging this we obtain that if either $p(x,y)$ or $p(y,x)$ is positive then $\frac{p(x,y)}{p(y,x)} = \frac{\pi(y)}{\pi(x)}$. Thus (3) holds:

$$\frac{p(x_1,x_2)}{p(x_2,x_1)}\frac{p(x_2,x_3)}{p(x_3,x_2)}\cdots\frac{p(x_{k-1},x_k)}{p(x_k,x_{k-1})}\frac{p(x_k,x_1)}{p(x_1,x_k)} = \frac{\pi(x_2)}{\pi(x_1)}\frac{\pi(x_3)}{\pi(x_2)}\cdots\frac{\pi(x_k)}{\pi(x_{k-1})}\frac{\pi(x_1)}{\pi(x_k)} = \frac{\prod_{i=1}^k \pi(x_i)}{\prod_{i=1}^k \pi(x_i)} = 1.$$

Now we show that (c) implies (b). Let us fix some $x^* \in S$ and declare $\varphi(x^*) = 1$. By irreducibility, for any $x \in S$ there exists an $n \in \mathbb{N}$ and a sequence of vertices $x_1, \ldots, x_n$ starting at $x_1 = x^*$ and ending at $x_n = x$ for which $p(x_1,x_2)\ldots p(x_{n-1},x_n) > 0$. Given this path from x^* to x, let us define

$$\varphi(x) = \varphi(x^*)\frac{p(x_1,x_2)}{p(x_2,x_1)}\frac{p(x_2,x_3)}{p(x_3,x_2)}\cdots\frac{p(x_{n-1},x_n)}{p(x_n,x_{n-1})}. \tag{4}$$

First let us argue that this definition is unambiguous, i.e., that it does not depend on the path that we chose to travel from x^* to x. Indeed, if $\widetilde{x}_1, \ldots, \widetilde{x}_m$ is another path with $\widetilde{x}_1 = x^*$ and $\widetilde{x}_m = x$ then we will show that they give the same $\varphi(x)$ value by creating a cycle by concatenating the first path and the reversed second path: $(x_1, x_2 \ldots, x_{n-1}, x_n, \widetilde{x}_{m-1}, \widetilde{x}_{m-2}, \ldots, \widetilde{x}_1)$. Now if we apply the cycle condition (3) to this cycle, we obtain

$$\frac{p(x_1,x_2)}{p(x_2,x_1)}\frac{p(x_2,x_3)}{p(x_3,x_2)}\cdots\frac{p(x_{n-1},x_n)}{p(x_n,x_{n-1})}\frac{p(\widetilde{x}_m,\widetilde{x}_{m-1})}{p(\widetilde{x}_{m-1},\widetilde{x}_m)}\frac{p(\widetilde{x}_{m-1},\widetilde{x}_{m-2})}{p(\widetilde{x}_{m-2},\widetilde{x}_{m-1})}\cdots\frac{p(\widetilde{x}_2,\widetilde{x}_1)}{p(\widetilde{x}_1,\widetilde{x}_2)} = 1. \tag{5}$$

Rearranging (5) we obtain

$$\frac{p(x_1,x_2)}{p(x_2,x_1)}\frac{p(x_2,x_3)}{p(x_3,x_2)}\cdots\frac{p(x_{n-1},x_n)}{p(x_n,x_{n-1})} = \frac{p(\widetilde{x}_1,\widetilde{x}_2)}{p(\widetilde{x}_2,\widetilde{x}_1)}\frac{p(\widetilde{x}_2,\widetilde{x}_3)}{p(\widetilde{x}_3,\widetilde{x}_2)}\cdots\frac{p(\widetilde{x}_{m-1},\widetilde{x}_m)}{p(\widetilde{x}_m,\widetilde{x}_{m-1})},$$

thus the definition of $\varphi(x)$ by (4) is indeed unambiguous.

Now by (4) we see that $\varphi(x_n) = \varphi(x_{n-1})\frac{p(x_{n-1},x_n)}{p(x_n,x_{n-1})}$ holds for any path $(x_1, \ldots, x_n)$, thus (b) holds.

Finally we prove that (b) implies (a). This is simple: given $\varphi : S \to (0,+\infty)$, let us define $\pi(x) := \frac{\varphi(x)}{Z}$, where $Z = \sum_{y\in S}\varphi(y)$. With this definition we have $\sum_{y\in S}\pi(y) = 1$, thus π is a distribution and it inherits the property $\pi(x)p(x,y) \equiv \pi(y)p(y,x)$ from φ. Therefore the detailed balance condition holds, and thus the Markov chain is indeed reversible.

3. Recall the notion of a *renewal process* from page 73 of the lecture notes. Let $\tau_1, \tau_2, \dots$ denote i.i.d. random variables which have the same distribution as the random variable τ. We assume that τ takes positive integer values and also assume that $\mathbb{P}(\tau \le M) = 1$ for some $M \in \mathbb{N}$. Now τ_k is the length of the k'th renewal interval. Denote by $p_x = \mathbb{P}[\tau = x]$ for any $x \in \{1, \dots, M\}$. Let us define $T_n = \sum_{i=1}^n \tau_i$, thus T_n is the time of the n'th renewal. Denote by α_t the time that has elapsed since the last renewal at time t and denote by β_t the time until the next renewal at time t, where $t = 0, 1, 2, \dots$ (if $t = T_n$ is a renewal time then we let $\alpha_t = 0$ and $\beta_t = \tau_{n+1}$). Denote by $\gamma_t = \alpha_t + \beta_t$ the total length of the renewal interval that contains t.

(a) Briefly argue that $Y_t := (\alpha_t, \beta_t)$, $t = 0, 1, 2, \dots$ is an irreducible Markov chain on the state space

$$S = \{\, (x, y) \,:\, x \in \{0, \dots, M-1\},\ y \in \{1, \dots, M\},\ p_{x+y} > 0 \,\}$$

and describe the transition rules of this Markov chain.

(b) What is the period of this Markov chain?

(c) Find the stationary distribution π of this Markov chain.

Hint: Denote by $\alpha_\infty, \beta_\infty, \gamma_\infty$ the time that has elapsed since the last renewal, time until the next renewal and the total length of the current renewal interval in the stationary state. Show that

$$\pi(\alpha_\infty = x, \beta_\infty = y) = \frac{p_{x+y}}{\mathbb{E}(\tau)}.$$

(d) Deduce from the result of the previous sub-exercise that $\mathbb{P}(\gamma_\infty = z) = \frac{z p_z}{\mathbb{E}(\tau)}$ for any $z \in \{1, \dots, M\}$.

Solution:

(a) The transition rules are as follows:

- If $\beta_t > 1$ then $\alpha_{t+1} = \alpha_t + 1$ and $\beta_{t+1} = \beta_t - 1$.
- If $\beta_t = 1$ then $\alpha_{t+1} = 0$ and β_{t+1} is a random variable which is independent from everything that happened earlier and β_{t+1} has the same distribution as τ.

$$\mathbb{P}\Big((\alpha_{t+1}, \beta_{t+1}) = (x', y') \mid (\alpha_t, \beta_t) = (x, y)\Big) = \begin{cases} \mathbb{1}[\, x' = x + 1,\ y' = y - 1\,] & \text{if } y > 1, \\ p_{y'} \mathbb{1}[\, x' = 0\,] & \text{if } y = 1. \end{cases}$$

Or, equivalently, if $p\Big((x, y), (x', y')\Big) := \mathbb{P}\Big((\alpha_{t+1}, \beta_{t+1}) = (x', y') \mid (\alpha_t, \beta_t) = (x, y)\Big)$, then

$$p\Big((x, y), (x', y')\Big) = \mathbb{1}[\, x' = x + 1,\ y' = y - 1\,] + p_{y'} \mathbb{1}[\, x' = 0,\ y = 1\,], \quad (x, y) \in S,\ (x', y') \in S. \quad (6)$$

(b) The period of (Y_t) is the greatest common divisor of the numbers contained in $\{\, x \,:\, p_x > 0\,\}$, since this number divides T_n for any $n \in \mathbb{N}$.

(c) Let us define $\mu(x, y) = p_{x+y}$ for any $x, y \in S$. We will show that μ is a stationary measure of our Markov chain, i.e., for any $(x', y') \in S$, we have

$$\sum_{(x,y) \in S} \mu(x, y) p\Big((x, y), (x', y')\Big) = \mu(x', y').$$

We will argue as follows. If state (x, y) gives mass $\mu(x, y) p\Big((x, y), (x', y')\Big)$ to state (x', y'), then we need to show that the total mass received by state (x', y') is exactly $\mu(x', y')$. Indeed:

- If $x' > 0$ then state (x', y') receives the total mass of state $(x' - 1, y' + 1)$, and this mass happens to be equal to $p_{x'+y'}$, i.e., $\mu(x', y')$.
- If $x' = 0$ then $(0, y')$ receives mass from states of form $(x, 1)$. Each such state sends mass $p_{x+1} p_{y'}$ to state $(0, y')$, so the total mass received by state $(0, y')$ is $\sum_{x=0}^{M-1} p_{x+1} p_{y'} = p_{y'}$, which happens to be equal to $\mu(0, y')$.

Of course, we can do the the same calculation more formally:

$$\sum_{(x,y)\in S} \mu(x,y) p\Big((x,y),(x',y')\Big) \overset{(6)}{=}$$

$$\sum_{(x,y)\in S} p_{x+y} \cdot (\mathbb{1}[\, x' = x+1,\, y' = y-1\,] + p_{y'}\mathbb{1}[\, x' = 0,\, y = 1\,]) =$$

$$\sum_{x=0}^{M-1}\sum_{y=1}^{M} p_{x+y}\mathbb{1}[\, x' = x+1,\, y' = y-1\,] + \sum_{x=0}^{M-1}\sum_{y=1}^{M} p_{x+y}p_{y'}\mathbb{1}[\, x' = 0,\, y = 1\,] =$$

$$\sum_{x=0}^{M-1}\sum_{y=1}^{M} p_{x+y}\mathbb{1}[\, x = x'-1,\, y = y'+1\,] + \sum_{x=0}^{M-1} p_{x+1}p_{y'}\mathbb{1}[\, x' = 0\,] =$$

$$p_{(x'-1)+(y'+1)}\mathbb{1}[x' > 0,\, y' < M] + p_{y'}\mathbb{1}[\, x' = 0\,]\left(\sum_{x=0}^{M-1} p_{x+1}\right) =$$

$$p_{x'+y'}\mathbb{1}[x' > 0,\, y' < M] + p_{y'}\mathbb{1}[\, x' = 0\,] = p_{x'+y'}\mathbb{1}[x' > 0\,] + p_{x'+y'}\mathbb{1}[\, x' = 0\,] = p_{x'+y'} = \mu(x',y').$$

It is easy to check that that $\sum_{(x,y)\in S} \mu(x,y) = \sum_{z=1}^{M} z p_z = \mathbb{E}(\tau)$ (see the calculation in part (d)). Given this assumption, we indeed obtain that $\mu/\mathbb{E}(\tau)$ is the stationary distribution.

(d)

$$\mathbb{P}(\gamma_\infty = z) = \sum_{x=0}^{M-1}\sum_{y=1}^{M} \frac{p_{x+y}}{\mathbb{E}(\tau)}\mathbb{1}[x+y=z] = \frac{p_z}{\mathbb{E}(\tau)}\sum_{x=0}^{M-1}\sum_{y=1}^{M}\mathbb{1}[x+y=z] = \frac{p_z}{\mathbb{E}(\tau)} z$$