Homework 3

Benjamin Noland

Note: I discussed some of these problems with Joseph Purtill and Don Walpola (both 654 students), but my write-up is my own work.

Problems from the book

Exercise 1.13

Let $x, y, z \in \mathcal{X}$.

- *Reflexivity*: Since $x = x$, $x \leftrightarrow x$ trivially.
- *Symmetry*: Suppose $x \leftrightarrow y$. If $x = y$, then $y \leftrightarrow x$ by reflexivity. If $x \to y$ and $y \to x$, then we see immediately that $y \leftrightarrow x$ as well.
- *Transitivity*: Suppose $x \leftrightarrow y$ and $y \leftrightarrow z$. If $x = y$ or $y = z$, then $x \leftrightarrow z$ trivially. If $x \to y$ and $y \to x$, and $y \to z$ and $z \to y$, then we see that $x \to z$ and $z \to x$ by transitivity of $\to$. Hence $x \leftrightarrow z$.

Therefore $\leftrightarrow$ is an equivalence relation on $\mathcal{X}$.

Exercise 1.14

Let P denote the transition matrix of the chain. Let $C_1, \ldots, C_k$ denote the essential communicating classes of the chain, and let S be the set of stationary distributions for the chain. First, note that S is closed under convex combinations. To see this, let $\alpha_1, \ldots, \alpha_n$ be non-negative real numbers with $\alpha_1 + \cdots + \alpha_n = 1$, and let $\pi_1, \ldots, \pi_n \in S$. Then

$$\sum_{x \in \mathcal{X}} \sum_{i=1}^{n} \alpha_i \pi_i(x) = \sum_{i=1}^{n} \alpha_i \sum_{x \in \mathcal{X}} \pi_i(x) = \sum_{i=1}^{n} \alpha_i = 1$$

and

$$\sum_{x \in \mathcal{X}} \sum_{i=1}^{n} \alpha_i \pi_i(x) \geq 0 \quad \text{for any } x \in \mathcal{X},$$

and finally,

$$\sum_{i=1}^{n} \alpha_i(\pi_i P) = \sum_{i=1}^{n} \alpha_i \pi_i.$$

Hence $\sum_{i=1}^{n} \alpha_i \pi_i \in S$. Let $\pi_{C_i}, \ldots, \pi_{C_k}$ be the unique stationary distributions corresponding to $C_1, \ldots, C_k$ (they exist and are unique since the restriction P_{C_i} is irreducible for any $1 \leq i \leq k$). For each $1 \leq i \leq k$, define

$$\pi'_{C_i}(x) = \begin{cases} \pi_{C_i}(x) & \text{if } x \in C_i \\ 0 & \text{if } x \notin C_i \end{cases} \quad \text{for each } x \in \mathcal{X}. \tag{1}$$

Then it follows that π'_{C_i} is stationary for P for each $1 \leq i \leq k$, and hence $\pi'_{C_i} \in S$. Let S' denote the set of convex combinations of $\pi'_{C_1}, \ldots, \pi'_{C_k}$. Then $S' \subseteq S$ by the comments above. Moreover, note that since $C_1, \ldots, C_k$ are disjoint, the definition (1) implies that $\pi'_{C_1}, \ldots, \pi'_{C_k}$ are linearly independent. Thus S' is a polytope whose vertices are $\pi'_{C_1}, \ldots, \pi'_{C_k}$.

To see that $S \subseteq S'$ as well, let $\pi \in S$. Let $1 \leq i \leq k$. Define

$$\pi''_{C_i}(x) = \begin{cases} \frac{\pi(x)}{k_i} & \text{if } x \in C_i \\ 0 & \text{if } x \notin C_i \end{cases} \quad \text{for each } x \in \mathcal{X},$$

where $k_i = \sum_{y \in C_i} \pi(y)$. Thus $\sum_{y \in \mathcal{X}} \pi''_{C_i}(x) = 1$, so that π''_{C_i} is a valid probability distribution. Moreover, since π is stationary for P it follows that π''_{C_i} is also stationary for P, and since the stationary for C_i is unique, it follows that $\pi''_{C_i} = \pi'_{C_i}$. Thus, for any $x \in \mathcal{X}$,

$$\pi(x) = \sum_{i=1}^{k} k_i \pi'_{C_i}(x),$$

and hence $\pi(x) = \sum_{i=1}^{k} k_i \pi'_{C_i}$. But also,

$$1 = \sum_{x \in \mathcal{X}} \pi(x) = \sum_{x \in \mathcal{X}} \sum_{i=1}^{k} k_i \pi'_{C_i}(x) = \sum_{i=1}^{k} k_i \sum_{x \in \mathcal{X}} \pi'_{C_i}(x) = \sum_{i=1}^{k} k_i.$$

Thus π is a convex combination of $\pi'_{C_1}, \ldots, \pi'_{C_k}$. Hence $\pi \in S'$, so that $S = S'$.

The Strong Markov Property

Exercise 1

$$I\{\tau < \infty\} = I\left\{\bigcup_{t=0}^{\infty} \{\tau = t\}\right\} = \sum_{t=0}^{\infty} I\{\tau = t\}$$

Exercise 2

By exercise 1 and the Dominated Convergence Theorem,

$$
\begin{aligned}
\mathrm{E}_x[I\{\tau < \infty\} f(X_\tau, \ldots, X_{\tau+m})] &= \mathrm{E}_x\left[\sum_{t=0}^{\infty} I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})\right] \\
&= \sum_{t=0}^{\infty} \mathrm{E}_x[I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})] \\
&= \sum_{t=0}^{\infty} \mathrm{E}_x\{\mathrm{E}_x[f(X_\tau, \ldots, X_{\tau+m}) | \tau = t] I\{\tau = t\}\} \\
&= \sum_{t=0}^{\infty} \mathrm{E}_x\{\mathrm{E}_{X_\tau}[f(X_0, \ldots, X_m)] I\{\tau = t\}\} \\
&= \mathrm{E}_x\left\{\sum_{t=0}^{\infty} \mathrm{E}_{X_\tau}[f(X_0, \ldots, X_m)] I\{\tau = t\}\right\} \\
&= \mathrm{E}_x[I\{\tau < \infty\} F(X_\tau)].
\end{aligned}
$$

Note that the Dominated Convergence Theorem justified switching sums and expectations in the above. As an example, note that if for each $n \geq 1$ we set

$$
Y_n = \sum_{t=0}^{n} I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})
$$

and

$$
Y = \sum_{t=0}^{\infty} I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})
$$

and $Z = 1$, then since $E(|Z|) = 1 < \infty$ trivially, and $\mathrm{P}\{|Y_n| \leq |Z|\} = 1$, the Dominated Convergence Theorem implies that

$$
\begin{aligned}
\mathrm{E}_x\left[\sum_{t=0}^{\infty} I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})\right] &= \mathrm{E}_x(Y) \\
&= \lim_{n \to \infty} \mathrm{E}_x(Y_n) \\
&= \lim_{n \to \infty} \mathrm{E}_x\left[\sum_{t=0}^{n} I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})\right] \\
&= \lim_{n \to \infty} \sum_{t=0}^{n} \mathrm{E}_x[I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})] \\
&= \sum_{t=0}^{\infty} \mathrm{E}_x[I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})].
\end{aligned}
$$

Exercise 3

$$
\begin{aligned}
\mathrm{E}_x[I\{\tau < \infty\} f(X_\tau, \ldots, X_{\tau+m})] &= \mathrm{E}_x[I\{\tau < \infty\} F(X_\tau)] \\
&= \mathrm{E}_x[I\{\tau < \infty\} F(z)] \\
&= F(z)\mathrm{E}_x[I\{\tau < \infty\}] \\
&= F(z)\mathrm{P}_x\{\tau < \infty\}
\end{aligned}
$$

Exercise 4

For any $(x_0, \ldots, x_m) \in \mathcal{X}^{m+1}$,

$$
f(x_0, \ldots, x_m) = \sum_{(y_0, \ldots, y_m) \in \mathcal{X}^{m+1}} I\{x_0 = y_0, \ldots, x_m = y_m\} f(y_0, \ldots, y_m).
$$

(The indicator functions in the sum are maps $\mathcal{X}^{m+1} \to \{0, 1\}$).

As for the generalization,

$$
\begin{aligned}
&\mathrm{E}_x[I\{\tau < \infty\} f(X_\tau, \ldots, X_{\tau+m})] \\
&\quad = \sum_{(y_0, \ldots, y_m) \in \mathcal{X}^{m+1}} \mathrm{E}_x[I\{\tau < \infty\} I\{X_\tau = y_0, \ldots, X_{\tau+m} = y_m\} f(y_0, \ldots, y_m)] \\
&\quad = \sum_{(y_0, \ldots, y_m) \in \mathcal{X}^{m+1}} f(y_0, \ldots, y_m) \mathrm{E}_x[I\{\tau < \infty\} I\{X_\tau = y_0, \ldots, X_{\tau+m} = y_m\}] \\
&\quad = \sum_{(y_0, \ldots, y_m) \in \mathcal{X}^{m+1}} f(y_0, \ldots, y_m) \mathrm{E}_x\{I\{\tau < \infty\} \mathrm{E}_{X_\tau}[I\{X_\tau = y_0, \ldots, X_{\tau+m} = y_m\}]\}
\end{aligned}
$$

by way of exercise 2.

Exercise 5

(Note that we must have $C_m = \mathcal{X}^{m+1}$ for each $m \geq 0$ in order for this problem to make sense). We have the following:

$$
\begin{aligned}
&\mathrm{E}_x[g(X_0, \ldots, X_\tau) I\{\tau < \infty\} f(X_\tau, \ldots, X_{\tau+m})] \\
&\quad = \mathrm{E}_x\left[\sum_{t=0}^{\infty} g(X_0, \ldots, X_\tau) I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m})\right] \\
&\quad = \mathrm{E}_x\left\{\mathrm{E}_x\left[\sum_{t=0}^{\infty} g(X_0, \ldots, X_\tau) I\{\tau = t\} f(X_\tau, \ldots, X_{\tau+m}) | \tau = t\right] I\{\tau = t\}\right\} \\
&\quad = \mathrm{E}_x\left\{\mathrm{E}_x\left[\sum_{t=0}^{\infty} g(X_0, \ldots, X_t) f(X_\tau, \ldots, X_{\tau+m})\right] I\{\tau = t\}\right\} \\
&\quad = \mathrm{E}_x\left\{\mathrm{E}_x\left[\sum_{t=0}^{\infty} \sum_{(y_0,\ldots,y_t)\in\mathcal{X}^{t+1}} g(y_0, \ldots, y_t) I\{X_0 = y_0, \ldots, X_t = y_t\} f(X_\tau, \ldots, X_{\tau+m})\right] I\{\tau = t\}\right\} \\
&\quad = \sum_{t=0}^{\infty} \mathrm{E}_x\left\{\mathrm{E}_x\left[\sum_{(y_0,\ldots,y_t)\in\mathcal{X}^{t+1}} g(y_0, \ldots, y_t) I\{X_0 = y_0, \ldots, X_t = y_t\} f(X_\tau, \ldots, X_{\tau+m})\right] I\{\tau = t\}\right\} \\
&\quad = \sum_{t=0}^{\infty} \mathrm{E}_x\left\{\sum_{(y_0,\ldots,y_t)\in\mathcal{X}^{t+1}} g(y_0, \ldots, y_t) I\{X_0 = y_0, \ldots, X_t = y_t\} \mathrm{E}_x\left[f(X_\tau, \ldots, X_{\tau+m})\right] I\{\tau = t\}\right\} \\
&\quad = \sum_{t=0}^{\infty} \mathrm{E}_x\left\{g(X_0, \ldots, X_t) \mathrm{E}_x\left[f(X_\tau, \ldots, X_{\tau+m})\right] I\{\tau = t\}\right\} \\
&\quad = \sum_{t=0}^{\infty} \mathrm{E}_x\left\{g(X_0, \ldots, X_t) \mathrm{E}_{X_\tau}\left[f(X_0, \ldots, X_m)\right] I\{\tau = t\}\right\} \\
&\quad = \mathrm{E}_x\left\{\sum_{t=0}^{\infty} g(X_0, \ldots, X_t) \mathrm{E}_{X_\tau}\left[f(X_0, \ldots, X_m)\right] I\{\tau = t\}\right\} \\
&\quad = \mathrm{E}_x\left\{g(X_0, \ldots, X_\tau) \mathrm{E}_{X_\tau}\left[f(X_0, \ldots, X_m)\right] I\{\tau < \infty\}\right\} \\
&\quad = \mathrm{E}_x\left\{g(X_0, \ldots, X_\tau) F(X_\tau) I\{\tau < \infty\}\right\}.
\end{aligned}
$$

Note that since the state space $\mathcal{X}$ is finite, $\mathcal{X}^{t+1}$ is finite for any $t \geq 0$, and so g is bounded in the expression above. Thus the Dominated Convergence Theorem justifies the switching of infinite sums and expectations.

Exercise 6

The sequence (Y_t) satisfies the Markov property since for any states $y_0, y_1, \dots, y_t$ we have

$$\begin{aligned} \mathrm{P}\{Y_t = y_t | Y_{t-1} = y_{t-1}, \dots, Y_0 = y_0\} &= \mathrm{P}\{X_{t+\tau} = y_t | X_{t-1+\tau} = y_{t-1}, \dots, X_\tau = y_0\} \\ &= \mathrm{P}\{X_{t+\tau} = y_t | X_{t-1+\tau} = y_{t-1}\} \\ &= \mathrm{P}\{Y_t = y_t | Y_{t-1} = y_{t-1}\} \end{aligned}$$

because the original chain (X_t) satisfies the Markov property. Moreover, for any states x and y,

$$\mathrm{P}\{Y_{t+1} = y | Y_t = x\} = \mathrm{P}\{X_{t+\tau+1} = y | X_{t+\tau} = x\} = P(x, y).$$

Finally, for any state y, we immediately have $\mathrm{P}\{Y_0 = y\} = \mathrm{P}_x\{X_\tau = y\}$.

Exercise 7

I will proceed by induction on r. When $r = 1$ the result holds trivially. Now let $r > 1$ and assume that $\mathrm{E}_x(\tau_x^r) = r\mathrm{E}_x(\tau_x^1)$. We have:

$$\tau_x^{r+1} = \inf\{t > \tau_x^r : X_t = x\} = \inf\{t - \tau_x^r > 0 : X_t = x\} = \inf\{t' > 0 : X_{t'+\tau_x^r} = x\} = \tau_x^1 + \tau_x^r,$$

and thus $\mathrm{E}_x(\tau_x^{r+1}) = \mathrm{E}_x(\tau_x^r) + \mathrm{E}_x(\tau_x^1) = (r+1)\mathrm{E}_x(\tau_x^1)$ by the inductive hypothesis. So the result follows by induction.

Since the chain is irreducible (we need this assumption to use Lemma 1.13), $\mathrm{E}_x(\tau_x^+) < \infty$ by Lemma 1.13, and thus for any $r > 0$,

$$0 < \mathrm{E}_x(\tau_x^r) = r\mathrm{E}_x(\tau_x^1) = r\mathrm{E}_x(\tau_x^+) < \infty.$$

Thus, by the Markov inequality, for any $t > 0$,

$$\mathrm{P}_x\{\tau_x^r < \infty\}^c = \mathrm{P}_x\left\{\bigcap_{s=0}^{\infty}\{\tau_x^r \geq s\}\right\} \leq \mathrm{P}_x\{\tau_x^r \geq t\} \leq \frac{\mathrm{E}_x(\tau_x^r)}{t},$$

so that

$$\mathrm{P}_x\{\tau_x^r < \infty\}^c \leq \lim_{t\to\infty} \frac{\mathrm{E}_x(\tau_x^r)}{t} = 0,$$

and hence $\mathrm{P}_x\{\tau_x^r < \infty\} = 1$ upon taking the complement.

Let $f, g : \cup_{m=0}^{\infty} \to \{0, 1\}$. Then

$$\begin{aligned} &\mathrm{E}_x[g(X_0, \dots, X_{\tau_x^1}) f(X_{\tau_x^1}, \dots, X_{\tau_x^r})] \\ &\quad= \mathrm{P}_x\{g(X_0, \dots, X_{\tau_x^1}) = 1, f(X_{\tau_x^1}, \dots, X_{\tau_x^r}) = 1\} \\ &\quad= \mathrm{P}_x\{f(X_{\tau_x^1}, \dots, X_{\tau_x^r}) = 1 | g(X_0, \dots, X_{\tau_x^1}) = 1\}\mathrm{P}_x\{g(X_0, \dots, X_{\tau_x^1}) = 1\} \\ &\quad= \mathrm{P}_x\{f(X_{\tau_x^1}, \dots, X_{\tau_x^r}) = 1\}\mathrm{P}_x\{g(X_0, \dots, X_{\tau_x^1}) = 1\} \\ &\quad= \mathrm{E}_x[g(X_0, \dots, X_{\tau_x^1})]\mathrm{E}_x[f(X_{\tau_x^1}, \dots, X_{\tau_x^r})] \end{aligned}$$

by way of the Markov property. In addition,

$$\begin{aligned}
\mathrm{E}_x[f(X_{\tau_x^1},\ldots,X_{\tau_x^r})] &= \mathrm{P}_x\{f(X_{\tau_x^1},\ldots,X_{\tau_x^r})=1\} \\
&= \sum_{t=0}^{\infty} \mathrm{P}_x\{f(X_{\tau_x^1},\ldots,X_{\tau_x^r})=1|\tau_x^1=t\}\mathrm{P}_x\{\tau_x^1=t\} \\
&= \sum_{t=0}^{\infty} \mathrm{P}_x\{f(X_0,\ldots,X_{\tau_x^{r-1}})=1\}\mathrm{P}_x\{\tau_x^1=t\} \\
&= \mathrm{P}_x\{f(X_0,\ldots,X_{\tau_x^{r-1}})=1\}\mathrm{P}_x\{\tau_x^1<\infty\} \\
&= \mathrm{P}_x\{f(X_0,\ldots,X_{\tau_x^{r-1}})=1\} = \mathrm{E}_x[f(X_0,\ldots,X_{\tau_x^{r-1}})]
\end{aligned}$$

due to the fact that $\tau_x^r = \tau_x^{r-1} + \tau_x^1$ for any $r > 1$, and $\mathrm{P}\{\tau_x^1 < \infty\} = 1$. Therefore,

$$\begin{aligned}
\mathrm{E}_x[g(X_0,\ldots,X_{\tau_x^1})f(X_{\tau_x^1},\ldots,X_{\tau_x^r})] &= \mathrm{E}_x[g(X_0,\ldots,X_{\tau_x^1})]\mathrm{E}_x[f(X_{\tau_x^1},\ldots,X_{\tau_x^r})] \\
&= \mathrm{E}_x[g(X_0,\ldots,X_{\tau_x^1})]\mathrm{E}_x[f(X_0,\ldots,X_{\tau_x^{r-1}})].
\end{aligned}$$

In particular,

$$\begin{aligned}
\mathrm{E}_x[g(X_0,\ldots,X_{\tau_x^1})f(X_{\tau_x^1},\ldots,X_{\tau_x^2})] &= \mathrm{P}_x\{g(X_0,\ldots,X_{\tau_x^1})=1, f(X_{\tau_x^1},\ldots,X_{\tau_x^2})=1\} \\
&= \mathrm{P}_x\{g(X_0,\ldots,X_{\tau_x^1})=1\}\mathrm{P}_x\{f(X_{\tau_x^1},\ldots,X_{\tau_x^2})=1\},
\end{aligned}$$

so that the events $\{g(X_0,\ldots,X_{\tau_x^1})=1\}$ and $\{f(X_{\tau_x^1},\ldots,X_{\tau_x^2})=1\}$ are independent.

Exercise 8

Using the fact that $\tau_x^2 = \tau_x^1 + \tau_x^1$, we have for any $y \in \{0,1\}$,

$$\mathrm{P}_x\{h(X_{\tau_x^1},\ldots,X_{\tau_x^2})=y\} = \mathrm{P}_x\{h(X_0,\ldots,X_{\tau_x^1})=y\},$$

and so $h(X_{\tau_x^1},\ldots,X_{\tau_x^2})$ and $h(X_0,\ldots,X_{\tau_x^1})$ are identically distributed. Moreover, it follows from exercise 7 that $h(X_{\tau_x^1},\ldots,X_{\tau_x^2})$ and $h(X_0,\ldots,X_{\tau_x^1})$ are independent. This is due to the fact that the events $\{h(X_{\tau_x^1},\ldots,X_{\tau_x^2})=x\}$ and $\{h(X_0,\ldots,X_{\tau_x^1})=y\}$ for any choice of $x,y \in \{0,1\}$ by the final result of exercise 7.

On Proposition 1.14(i)

Exercise 1

Generalization: Fix a state $z \in \mathcal{X}$, and define a function $\tilde{\pi}$ by

$$\tilde{\pi} = \sum_{t=0}^{\infty} \mathrm{P}_z\{X_t = y, \tau > t\}.$$

If $\mathrm{P}_z\{\tau < \infty, X_\tau = z\} = 1$, then $\tilde{\pi}$ is stationary for the transition matrix P.

Proof: Since $\mathrm{P}_z\{\tau < \infty, X_\tau = z\} = 1$, we have $\mathrm{P}_z\{\tau = \tau_z^r\} = 1$ for some $r > 0$. Hence, by exercise 7 of the Strong Markov Property problem, $\mathrm{E}_z(\tau) < \infty$, and hence

$$\infty > \mathrm{E}_z(\tau) = E_z(\tau_z^r) = \sum_{t=0}^{\infty} \mathrm{P}_z\{\tau > t\} \geq \sum_{t=0}^{\infty} \mathrm{P}_z\{X_t = y, \tau > t\} = \tilde{\pi}(y) \geq 0,$$

so that $\tilde{\pi}(y)$ is finite for every $y \in \mathcal{X}$. To see that $\tilde{\pi}$ is stationary, let $y \in \mathcal{X}$. Then

$$\sum_{x\in\mathcal{X}} \tilde{\pi}(x)P(x,y) = \sum_{x\in\mathcal{X}} \sum_{t=0}^{\infty} \mathrm{P}_z\{X_t = y, \tau > t\}P(x,y). \tag{1}$$

Note that for any $x \in \mathcal{X}$,

$$\begin{aligned}\mathrm{P}_z\{X_t = x, X_{t+1} = y, \tau \geq t+1\} &= \mathrm{P}_z\{X_t = x, \tau \geq t+1\}\mathrm{P}\{X_{t+1} = y | X_t = x, \tau \geq t+1\} \\ &= \mathrm{P}_z\{X_t = x, \tau \geq t+1\}P(x,y),\end{aligned}$$

since the Markov property implies that

$$P(x,y) = \mathrm{P}\{X_{t+1} = y | X_t = x\} = \mathrm{P}\{X_{t+1} = y | X_t = x, \tau \geq t+1\}$$

because the event $\{\tau \geq t+1\}$ depends solely upon $X_0, \ldots, X_t$. Therefore (1) becomes

$$\begin{aligned}\sum_{x\in\mathcal{X}} \tilde{\pi}(x)P(x,y) &= \sum_{x\in\mathcal{X}} \sum_{t=0}^{\infty} \mathrm{P}_z\{X_t = y, \tau > t\}P(x,y) \\ &= \sum_{x\in\mathcal{X}} \sum_{t=0}^{\infty} \mathrm{P}_z\{X_t = x, X_{t+1} = y, \tau \geq t+1\} \\ &= \sum_{t=0}^{\infty} \sum_{x\in\mathcal{X}} \mathrm{P}_z\{X_t = x, X_{t+1} = y, \tau \geq t+1\} \\ &= \sum_{t=0}^{\infty} \mathrm{P}_z\{X_{t+1} = y, \tau \geq t+1\} \\ &= \sum_{t=1}^{\infty} \mathrm{P}_z\{X_t = y, \tau \geq t\} \\ &= \tilde{\pi}(y) - \mathrm{P}_z\{X_0 = y, \tau > 0\} + \sum_{t=1}^{\infty} \mathrm{P}_z\{X_t = y, \tau = t\} \\ &= \tilde{\pi}(y) - \mathrm{P}_z\{X_0 = y, \tau > 0\} + \mathrm{P}_z\{X_\tau = y\} - \mathrm{P}_z\{X_\tau = y, \tau = 0\}\end{aligned}$$

If $y = z$, then $\mathrm{P}_z\{X_\tau = y\} = 1$, and since the events $\{X_0 = y, \tau > 0\}$ and $\{X_\tau = y, \tau = 0\}$ are disjoint, it follows that $\mathrm{P}_z\{X_0 = y, \tau > 0\} = 1$ and $\mathrm{P}_z\{X_\tau = y, \tau = 0\} = 0$, or $\mathrm{P}_z\{X_0 = y, \tau > 0\} = 0$ and $\mathrm{P}_z\{X_\tau = y, \tau = 0\} = 1$. Thus in this case

$$\sum_{x\in\mathcal{X}} \tilde{\pi}(x)P(x,y) = \tilde{\pi}(y).$$

In addition, if $y \neq z$, then $\mathrm{P}_z\{X_0 = y, \tau > 0\} = \mathrm{P}_z\{X_\tau = y\} = \mathrm{P}_z\{X_\tau = y, \tau = 0\} = 0$, and we get the same result. Therefore $\tilde{\pi}$ is stationary for P.

Exercise 2

For any $t \geq 0$, define a map $p_t : \mathcal{X}^{t+1} \to \{0, 1\}$ such that for any $(x_0, \ldots, x_t) \in \mathcal{X}^{t+1}$,

- $p_t(x_0, \ldots, x_t) = 1$ if $x_t = z$, $\mathcal{X} \subseteq \{x_1, \ldots, x_t\}$, and for every $0 \leq t' < t$, $x_{t'} \neq z$ or $\mathcal{X} \not\subseteq \{x_1, \ldots, x_{t'}\}$;
- $p_t(x_0, \ldots, x_t) = 0$ otherwise.

Then the event $\{\tau = t\} = \{p_t(x_0, \ldots, x_t) = 1\}$, so that τ is a stopping time.

Note that since the chain is irreducible, it is possible to order the states $x_0, \ldots, x_n \in \mathcal{X}$ of the chain to yield a path $z = x_0 \to x_1 \to x_2 \to \cdots \to x_n \to x_0 = z$ that passes through every state in the chain and returns to the starting state z. Thus we can write

$$0 < \mathrm{E}_z(\tau) = \mathrm{E}_{x_0}(\tau_{x_1}^+) + \mathrm{E}_{x_1}(\tau_{x_2}^+) + \cdots + \mathrm{E}_{x_n}(\tau_{x_0}^+) < \infty$$

by way of Lemma 1.13. Therefore, the Markov inequality implies that for any $t > 0$,

$$\mathrm{P}_z\{\tau < \infty\}^c = \mathrm{P}_z\left\{\bigcap_{s=0}^{\infty}\{\tau \geq s\}\right\} \leq \mathrm{P}_z\{\tau \geq t\} \leq \frac{\mathrm{E}_z(\tau)}{t},$$

so that

$$\mathrm{P}_z\{\tau < \infty\}^c \leq \lim_{t \to \infty} \frac{\mathrm{E}_z(\tau)}{t} = 0,$$

and therefore $\mathrm{P}_z\{\tau < \infty\} = 1$ upon taking the complement. Thus

$$\mathrm{P}_z\{\tau < \infty, X_\tau = z\} = \mathrm{P}_z\{X_\tau = z\} = 1.$$