Homework 4

Benjamin Noland

Ergodic Theorem

Exercise 1

Since $\mathrm{P}(A) = \mathrm{P}(B) = 1$, $\mathrm{P}(A^c) = \mathrm{P}(B^c) = 0$, and hence

$$\mathrm{P}(A^c \cup B^c) \leq \mathrm{P}(A^c) + \mathrm{P}(B^c) = 0.$$

Thus, upon taking complements, we find

$$\mathrm{P}(A \cap B) = 1 - \mathrm{P}(A^c \cup B^c) \geq 1 - 0 = 1,$$

so that $\mathrm{P}(A^c \cup B^c) = 1$.

Exercise 2

Define a map $h : \cup_{m=0}^{\infty} \mathcal{X}^{m+1} \to \mathbb{R}$ by

$$h(x_0, \ldots, x_t) = \sum_{i=0}^{t} f(x_i) \quad \text{for any } x_0, \ldots, x_t \in \cup_{m=0}^{\infty} \mathcal{X}^{m+1}. \tag{1}$$

Then for any $k \geq 0$, we can write $S_k = h(X_{\tau_x^k}, X_{\tau_x^k+1}, \ldots, X_{\tau_x^{k+1}-1})$. Roughly speaking, the fact that the S_k's are IID follows from the fact that the chain starts afresh at each return time τ_x^k. This can probably be formalized using Exercises 7 and 8 of the Strong Markov Property question, along with the map defined in (1) (but I don't have time to figure this out).

Exercise 3

First, note that $\mathrm{E}_x(|S_0|) < \infty$. To see this, note that since the state space $\mathcal{X}$ is finite, we can set

$$M = \max_{x \in \mathcal{X}} |f(x)|.$$

Thus,

$$\mathrm{E}_x(|S_0|) = \mathrm{E}_x\left(\sum_{t=0}^{\tau_x^1-1} f(X_t)\right) \le \mathrm{E}_x\left(\sum_{t=0}^{\tau_x^1-1} M\right) = \mathrm{E}_x(M\tau)x^1) = M\mathrm{E}_x(\tau_x^1) < \infty$$

where the fact that $\mathrm{E}_x(\tau_x^1) < \infty$ follows from Lemma 1.13, since the chain is irreducible. Note that for any $n \ge 0$ we have the relation

$$\sum_{k=0}^{n-1} S_k = \sum_{k=0}^{n-1} \sum_{t=\tau_x^k}^{\tau_x^{k+1}-1} f(X_t) = \sum_{t=0}^{\tau_x^n-1} f(X_t).$$

Thus, since the S_k's are IID, the Strong Law of Large Numbers yields

$$\mathrm{P}_x\left\{\lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{\tau_x^n-1} f(X_t) = \mathrm{E}_x(S_0)\right\} = 1.$$

In addition, note that for any $k \ge 0$,

$$\tau_x^{k+1} - \tau_x^k = \inf\{t - \tau_x^k > 0 : X_{t-\tau_x^k} = x\},$$

so that $(\tau_x^{k+1} - \tau_x^k)$ has the same distribution as τ_x^1, since the chain starts afresh at time τ_x^k. In addition, this implies that the differences $(\tau_x^{k+1} - \tau_x^k)$ are mutually independent, and hence IID. Moreover, for any $n \ge 1$,

$$\tau_x^n = \sum_{k=0}^{n-1} (\tau_x^{k+1} - \tau_x^k),$$

and that $\mathrm{E}_x(\tau_x^1 - \tau_x^0) = \mathrm{E}_x(\tau_x^1) < \infty$ by Lemma 1.13, since the chain is irreducible. So the Strong Law of Large Numbers yields

$$\mathrm{P}_x\left\{\lim_{n\to\infty} \frac{1}{n}\tau_x^n = \mathrm{E}_x(\tau_x^1)\right\} = 1.$$

Thus, applying Exercise 1 of this problem, we get

$$\begin{aligned}
1 &= \mathrm{P}_x\left\{\lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{\tau_x^n-1} f(X_t) = \mathrm{E}_x(S_0), \lim_{n\to\infty} \frac{1}{n}\tau_x^n = \mathrm{E}_x(\tau_x^1)\right\} \\
&\le \mathrm{P}_x\left\{\lim_{n\to\infty} \frac{1}{\tau_x^n} \sum_{t=0}^{\tau_x^n-1} f(X_t) = \frac{\mathrm{E}_x(S_0)}{\mathrm{E}_x(\tau_x^1)}\right\},
\end{aligned}$$

so that

$$\mathrm{P}_x\left\{\lim_{n\to\infty} \frac{1}{\tau_x^n} \sum_{t=0}^{\tau_x^n-1} f(X_t) = L\right\} = 1 \quad \text{where} \quad L = \frac{\mathrm{E}_x(S_0)}{\mathrm{E}_x(\tau_x^1)}$$

Exercise 4

First, note that

$$\begin{aligned}
\mathrm{E}_x(S_0) &= \mathrm{E}_x\left(\sum_{t=0}^{\tau_x^1-1} f(X_t)\right) = \mathrm{E}_x\left(\sum_{t=0}^{\tau_x^1-1}\sum_{y\in\mathcal{X}} I\{X_t=y\}f(y)\right) \\
&= \mathrm{E}_x\left(\sum_{y\in\mathcal{X}} f(y)\sum_{t=0}^{\tau_x^1-1} I\{X_t=y\}\right) = \sum_{y\in\mathcal{X}} f(y)\mathrm{E}_x\left(\sum_{t=0}^{\tau_x^1-1} I\{X_t=y\}\right).
\end{aligned}$$

But note that $\sum_{t=0}^{\tau_x^1-1} I\{X_t=y\}$ is the number of visits to state y before returning to state x. Thus,

$$\tilde{\pi}(y) = \mathrm{E}_x\left(\sum_{t=0}^{\tau_x^1-1} I\{X_t=y\}\right),$$

where $\tilde{\pi}$ is the distribution defined on page 11 of the book. Therefore, since $\mathrm{E}_x(\tau_x^1) < \infty$ by Lemma 1.13, and since π is the unique stationary distribution of the chain by Corollary 1.17, Proposition 1.14(ii) implies that

$$\pi = \frac{\tilde{\pi}}{\mathrm{E}_x(\tau_x^1)}.$$

Thus,

$$\mathrm{E}_x(S_0) = \sum_{y\in\mathcal{X}} f(y)\tilde{\pi}(y) = \sum_{y\in\mathcal{X}} f(y)\pi(y)\mathrm{E}_x(\tau_x^1),$$

so that

$$L = \frac{\mathrm{E}_x(S_0)}{\mathrm{E}_x(\tau_x^1)} = \sum_{y\in\mathcal{X}} f(y)\pi(y) = E_\pi(f).$$

Exercise 5

For every integer $T \geq 1$, define

$$A_T = \frac{1}{T}\sum_{i=1}^{T} a_i.$$

Then for any choice of $T \geq 1$, since $T_k \to \infty$ as $k \to \infty$, there exists $k \geq 1$ satisfying $T_k \leq T \leq T_{k+1}$. Thus we can write

$$A_T = \frac{1}{T}\sum_{i=1}^{T} a_i = \frac{T_k}{T}\left(\frac{1}{T_k}\sum_{i=1}^{T_k} a_i\right) + \frac{1}{T}\sum_{i=T_k+1}^{T} a_i = \frac{T_k}{T}A_{T_k} + \frac{1}{T}\sum_{i=T_k+1}^{T} a_i.$$

I will consider each term in this expression separately. Since $T_k \leq T \leq T_{k+1}$,

$$A_{T_k} = \frac{T_k}{T_k} A_{T_k} \leq \frac{T_k}{T} A_{T_k} \leq \frac{T_k}{T_{k+1}} A_{T_k}.$$

Thus, since $A_{T_k} \to L$ and $T_k/T_{k+1} \to 1$ as $k \to \infty$, we see that

$$\lim_{k\to\infty} \frac{T_k}{T} A_{T_k} = L,$$

or equivalently, since $T_k \leq T \leq T_{k+1}$,

$$\lim_{T\to\infty} \frac{T_k}{T} A_{T_k} = L.$$

As for the second term, since $a_i \in [-M, M]$ for every $i \geq 1$,

$$\begin{aligned} \left| \frac{1}{T} \sum_{i=T_k+1}^{T} a_i \right| &\leq \frac{1}{T} \sum_{i=T_k+1}^{T} |a_i| \leq \frac{1}{T} \sum_{i=T_k+1}^{T} M = \frac{M}{T}(T - T_k) \leq \frac{M}{T}(T_{k+1} - T_k) \\ &\leq \frac{M}{T_k}(T_{k+1} - T_k) = M\left(\frac{T_{k+1}}{T_k} - 1\right) \to 0 \quad \text{as} \quad T \to \infty \end{aligned}$$

since because $T_k/T_{k+1} \to 1$ as $T \to \infty$, $T_{k+1}/T_k \to 1$ as $T \to \infty$ as well. Therefore

$$\lim_{T\to\infty} \frac{1}{T} \sum_{i=T_k+1}^{T} a_i = 0,$$

so that

$$\lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{T} a_i = L.$$

Exercise 6

To see that $\tau_x^n \geq n$ for every $n \geq 0$, I will proceed by induction on n. When $n = 0$, we have

$$\tau_x^0 = 0 \geq 0$$

by definition. Now let $n > 0$, and suppose that $\tau_x^n \geq n$. Then

$$\tau_x^{n+1} = \inf\{t > \tau_x^n : X_t = x\} = \inf\{t \geq \tau_x^n + 1 : X_t = x\} \geq \tau_x^n + 1 \geq n + 1$$

by the inductive hypotheses. Hence $\tau_x^n \geq n$ for every $n \geq 0$ by induction. In particular, this result implies that

$$\mathrm{P}_x \left\{ \lim_{n\to\infty} \tau_x^n = \infty \right\} = 1.$$

Moreover, note that for every $n \geq 0$,

$$\frac{\tau_x^n}{\tau_x^{n+1}} < \frac{\tau_x^n + \tau_x^1}{\tau_x^{n+1} + \tau_x^1}, \tag{1}$$

where the ratio $(\tau_x^n + \tau_x^1)/(\tau_x^{n+1} + \tau_x^1)$ has the same distribution as $\tau_x^{n+1}/\tau_x^{n+2}$, since the chain starts afresh at each of the times τ_x^n and τ_x^{n+1}. Thus, since $\tau_x^n/\tau_x^{n+1} \leq 1$ by definition, it follows from the inequality (1) that the ratio τ_x^n/τ_x^{n+1} can be made arbitrarily close to 1 if n is chosen large enough. Therefore,

$$\mathrm{P}_x \left\{ \lim_{n\to\infty} \frac{\tau_x^n}{\tau_x^{n+1}} = 1 \right\} = 1.$$

Since the function f is bounded (specifically, $f : \mathcal{X} \to [-M, M]$ for some $M > 0$), it follows from Exercise 5 of this problem that

$$1 = \mathrm{P}_x \left\{ \lim_{n\to\infty} \frac{1}{\tau_x^n} \sum_{t=0}^{\tau_x^n - 1} f(X_t) = E_\pi(f) \right\} \leq \mathrm{P}_x \left\{ \lim_{n\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} f(X_t) = E_\pi(f) \right\},$$

and the result follows.

Exercise 7

By Exercise 6 of this problem,

$$\begin{aligned}
\mathrm{P}_\mu \left\{ \lim_{n\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} f(X_t) = E_\pi(f) \right\} &= \sum_{x\in\mathcal{X}} \mathrm{P} \left\{ \lim_{n\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} f(X_t) = E_\pi(f) \,\middle|\, X_0 = x \right\} \mathrm{P}\{X_0 = x\} \\
&= \sum_{x\in\mathcal{X}} \mathrm{P}_x \left\{ \lim_{n\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} f(X_t) = E_\pi(f) \right\} \mu(x) \\
&= \sum_{x\in\mathcal{X}} \mu(x) = 1.
\end{aligned}$$

Exercise 8

The map $f : \mathcal{X} \to \{0, 1\}$ is given by

$$f(y) = I\{y = x_0\} \quad \text{for any } y \in \mathcal{X}.$$

Thus,

$$E_\pi(f) = \sum_{y\in\mathcal{X}} \pi(y) f(y) = \sum_{y\in\mathcal{X}} \pi(y) I\{y = x_0\} = \pi(x_0).$$

So in this special case the result becomes

$$\mathrm{P}_x \left\{ \lim_{n\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} I\{X_t = x_0\} = \pi(x_0) \right\} = 1$$

(i.e., we expect the long run proportion of time spent in state x_0 to be $\pi(x_0)$).

Random mapping representation

Exercises from Homework 2

Exercise 5

Using the given random mapping representation, we get

$$\begin{aligned}
&\mathrm{P}_\mu\{f(X_t,\dots,X_{t+k})=1|X_t=x_t,p(X_0,\dots,X_t)=1\}\\
&\qquad=\mathrm{P}_\mu\{f(x_t,\dots,X_{t+k})=1|p(X_0,\dots,x_t)=1\}\\
&\qquad=\mathrm{P}_\mu\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k})=1|p(r_0(Z_0),r(X_0,Z_0),\dots,x_t)=1\}.
\end{aligned}$$

But notice that the event $\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k})=1\}$ depends solely upon $Z_{t+1},\dots,Z_{t+k}$, and that the event $\{p(r_0(Z_0),r(X_0,Z_0),\dots,x_t)=1\}$ depends solely upon $Z_0,\dots,Z_{t-1}$. Hence these two events are independent, and so we can write

$$\begin{aligned}
&\mathrm{P}_\mu\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k})=1|p(r_0(Z_0),r(X_0,Z_0),\dots,x_t)=1\}\\
&\qquad=\mathrm{P}_\mu\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k})=1\}\\
&\qquad=\mathrm{P}_\mu\{f(X_t,r(X_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k})=1|X_t=x_t\}\\
&\qquad=\mathrm{P}_\mu\{f(X_t,\dots,X_{t+k})=1|X_t=x_t\}\\
&\qquad=\mathrm{P}_{x_t}\{f(X_0,\dots,X_k)=1\},
\end{aligned}$$

where the final equality follows from the time-homogeneity of the chain.

Exercise 6

Using the given random mapping representation, we get

$$\begin{aligned}
&\mathrm{P}_\mu\{f(X_t,\dots,X_{t+k})=1,p(X_0,\dots,X_t)=1|X_t=x_t\}\\
&\qquad=\mathrm{P}_\mu\{f(x_t,\dots,X_{t+k})=1,p(X_0,\dots,x_t)=1\}\\
&\qquad=\mathrm{P}_\mu\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k}))=1,p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1\}.
\end{aligned}$$

But notice that the event $\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k}))=1\}$ depends solely upon $Z_{t+1},\dots,Z_{t+k}$, and that the event $\{p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1\}$ depends solely upon $Z_0,\dots,Z_{t-1}$. Hence these two events are independent, and so we can write

$$\begin{aligned}
&\mathrm{P}_\mu\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k}))=1,p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1\}\\
&\qquad=\mathrm{P}_\mu\{f(x_t,r(x_t,Z_{t+1}),\dots,r(X_{t+k-1},Z_{t+k}))=1\}\mathrm{P}_\mu\{p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1\}\\
&\qquad=\mathrm{P}_\mu\{f(x_t,\dots,X_{t+k})=1\}\mathrm{P}_\mu\{p(X_0,\dots,x_t)=1\}\\
&\qquad=\mathrm{P}_\mu\{f(X_t,\dots,X_{t+k})=1|X_t=x_t\}\mathrm{P}_\mu\{p(X_0,\dots,X_t)=1|X_t=x_t\}\\
&\qquad=\mathrm{P}_{x_t}\{f(X_0,\dots,X_k)=1\}\mathrm{P}_\mu\{p(X_0,\dots,X_t)=1|X_t=x_t\},
\end{aligned}$$

where the final equality follows from the time-homogeneity of the chain.

Exercise 7

Using the given random mapping representation, we get

$$\begin{aligned}&\mathrm{P}_\mu\{f(X_{t+m},\dots,X_{t+m+k})=1,p(X_0,\dots,X_t)=1|E\}\\&\quad=\mathrm{P}_\mu\{f(x_{t+m},r(x_{t+m},Z_{t+m+1}),\dots,r(X_{t+m+k-1},Z_{t+m+k}))=1,\\&\qquad\quad p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1|E\}.\end{aligned}$$

But notice that the event $\{f(x_{t+m},r(x_{t+m},Z_{t+m+1}),\dots,r(X_{t+m+k-1},Z_{t+m+k}))=1\}$ depends solely upon $Z_{t+m+1},\dots,Z_{t+m+k}$, and that the event $\{p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1\}$ depends solely upon $Z_0,\dots,Z_{t-1}$. Hence these two events are independent under conditioning on E, and so we can write

$$\begin{aligned}&\mathrm{P}_\mu\{f(x_{t+m},r(x_{t+m},Z_{t+m+1}),\dots,r(X_{t+m+k-1},Z_{t+m+k}))=1,\\&\qquad p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1|E\}\\&\quad=\mathrm{P}_\mu\{f(x_{t+m},r(x_{t+m},Z_{t+m+1}),\dots,r(X_{t+m+k-1},Z_{t+m+k}))=1|E\}\\&\qquad\times\mathrm{P}_\mu\{p(r_0(Z_0),r(X_0,Z_1),\dots,x_t)=1|E\}\\&\quad=\mathrm{P}_\mu\{f(X_{t+m},\dots,X_{t+m+k})=1|E\}\mathrm{P}_\mu\{p(X_0,\dots,X_t)=1|E\}\\&\quad=\mathrm{P}_\mu\{f(X_{t+m},\dots,X_{t+m+k})=1|X_{t+m}=x_{t+m}\}\mathrm{P}_\mu\{p(X_0,\dots,X_t)=1|X_t=x_t\}\\&\quad=\mathrm{P}_{x_{t+m}}\{f(X_0,\dots,X_k)=1\}\mathrm{P}_\mu\{p(X_0,\dots,X_t)=1|X_t=x_t\},\end{aligned}$$

where the final equality follows from the time-homogeneity of the chain. As before, the result still holds if E is replaced by $\{X_t=x_t,X_{t+m}=x_{t+m}\}$.

Exercises from Homework 3

Exercise 5

(Ran out of time).

Exercise 7

(Ran out of time).

Exercises from the book

Exercise 3.1

We want to show that

$$\pi(x)P(x,y)=\pi(y)P(y,x)\quad\text{for every } x,y\in\mathcal{X}.$$

Let $x, y \in \mathcal{X}$. When $x = y$ the relation holds trivially, so assume $x \neq y$. Then

$$\begin{aligned}\pi(x)P(x,y) &= \pi(x)\Psi(x,y)\left[\frac{\pi(y)\Psi(y,x)}{\pi(x)\Psi(x,y)} \wedge 1\right] = [\pi(y)\Psi(y,x) \wedge \pi(x)\Psi(x,y)] \\ &= \pi(y)\Psi(y,x)\left[1 \wedge \frac{\pi(x)\Psi(x,y)}{\pi(y)\Psi(y,x)}\right] = \pi(y)P(y,x).\end{aligned}$$

Hence the chain is reversible, and π is stationary for P by Proposition 1.20.

Exercise 3.2

Let $x, y \in \mathcal{X}$. We want to show that

$$\pi(x)P(x,y) = \pi(y)P(y,x) \quad \text{for every } x, y \in \mathcal{X}. \tag{1}$$

Note that $P(x,y) > 0$ if and only if there exists a unique $v \in V$ with $y \in \mathcal{X}(x,v)$. In addition, note that $y \in \mathcal{X}(x,v)$ if and only if $x \in \mathcal{X}(y,v)$. Thus if such a unique $v \in V$ does not exist, then $P(x,y) = P(y,x) = 0$, and the relation (1) holds trivially. On the other hand, if such a unique $v \in V$ does exist, then

$$P(x,y) = \frac{1}{|V|}\pi^{x,v}(y) = \frac{1}{|V|}\frac{\pi(y)}{k_x},$$

where $k_x = \sum_{z \in \mathcal{X}(x,v)} \pi(z)$. In addition, noting that $\mathcal{X}(x,v) = \mathcal{X}(y,v)$, we see that

$$k_y = \sum_{z \in \mathcal{X}(y,v)} \pi(z) = \sum_{z \in \mathcal{X}(x,v)} \pi(z) = k_x,$$

and thus

$$P(y,x) = \frac{1}{|V|}\pi^{y,v}(x) = \frac{1}{|V|}\frac{\pi(x)}{k_y} = \frac{1}{|V|}\frac{\pi(x)}{k_x}.$$

Therefore,

$$\pi(x)P(x,y) = \pi(x)\left[\frac{1}{|V|}\frac{\pi(y)}{k_x}\right] = \pi(y)\left[\frac{1}{|V|}\frac{\pi(x)}{k_x}\right] = \pi(y)P(y,x).$$

Hence the chain is reversible, and π is stationary for P by Proposition 1.20.

Metropolis on graphs

Exercises 1-2

See the file `Metropolis_on_graphs_1.ipynb`.

Exercise 3

See the file `Metropolis_on_graphs_3.ipynb`.

Glauber dynamics on hardcore configurations

Exercises 1-3

See the file `Glauber_hardcore.ipynb`.

Exercise 4

I do this in the code. See the file `Glauber_hardcore.ipynb` for details of the implementation. To justify this theoretically, let $f : \mathcal{X} \to \mathbb{R}$ be a function defined as follows:

$$f(x) = \text{the number of particles in the configuration } x$$

for any $x \in \mathcal{X}$. Then f is a bounded function; specifically, $0 \leq f(x) \leq |V|$ for any $x \in \mathcal{X}$, where V denotes the vertex set of the graph. Since the uniform distribution π, defined by

$$\pi(x) = \frac{1}{|\mathcal{X}|} \quad \text{for any } x \in \mathcal{X}$$

is stationary for the chain, the Ergodic Theorem implies that

$$\mathrm{P}_{x_0} \left\{ \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} f(X_t) = E_\pi(f) \right\} = 1, \tag{1}$$

where x_0 denotes the starting configuration of the chain. But note that

$$E_\pi(f) = \sum_{x \in \mathcal{X}} \pi(x) f(x) = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} f(x),$$

the average number of particles over all hardcore configurations. Thus if we use the code to compute an observed value of the average $\frac{1}{T} \sum_{t=0}^{T-1} f(X_t)$ in (1) for a large value of T, then we expect this value to be close to the true average $E_\pi(f)$.