Nicolas Privault

Discrete Stochastic Processes

Tools for Machine Learning and Data Science

Solutions Manual

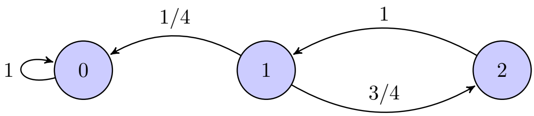
October 12, 2024

Contents

Chapter 1 - A Summary of Markov Chains

Exercise 1.1

a) The chain has the following graph:



Noting that state (0) is absorbing, by first step analysis we have

$$\begin{cases} g_0(0) = 1 \\ \\ g_0(1) = \dfrac{1}{4} g_0(0) + \dfrac{3}{4} g_0(2) \\ \\ g_0(2) = g_0(1), \end{cases}$$

which has for solution

$$g_0(0) = g_0(1) = g_0(2) = 1$$

as illustrated in the following R code.

```
install.packages("devtools"); library(devtools) # Install RTools as well
devtools::install_github('spedygiorgio/markovchain')
install.packages("igraph"); library(igraph); library(markovchain)
P<-matrix(c(1,0,0,1/4,0,3/4,0,1,0),nrow=3,byrow=TRUE);
      MC<-new("markovchain",transitionMatrix=P)
graph <- as(MC, "igraph")
plot(graph,vertex.size=50,edge.label.cex=2,edge.label=E(graph)$prob,
      edge.color='black', vertex.color='dodgerblue',vertex.label.cex=3)
hittingProbabilities(object = MC)
  1 2 3
1 1 0.00 0.00
2 1 0.75 0.75
3 1 1.00 0.75
```

b) By first step analysis, we have

$$\begin{cases} h_0(0) = 0 \\ \\ h_0(1) = 1 + \dfrac{1}{4} h_0(0) + \dfrac{3}{4} h_0(2) \\ \\ h_0(2) = 1 + h_0(1), \end{cases}$$

which has for solution

$$h_0(0) = 0, \quad h_0(1) = 7, \quad h_0(2) = 8,$$

as illustrated in the following R code.

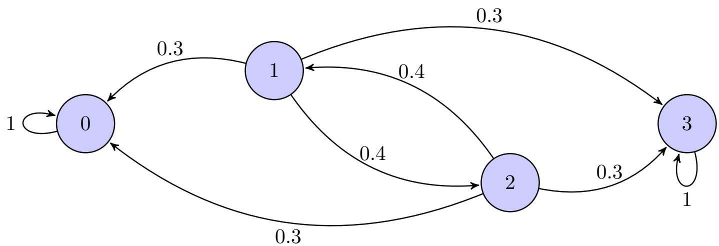
```
meanAbsorptionTime(object = MC)
7 8
```

Exercise 1.2 (Steele (2001), page 3). For all $k = 0, 1, \ldots, L$ and $n \geq 1$, we have

$$\begin{aligned}\mathbb{P}(T_{0,L} = \infty \mid S_0 = k) &\leq \mathbb{P}(T_{0,L} > nL \mid S_0 = k)\\ &\leq \mathbb{P}\left(\bigcap_{k=0}^{n-1}\{X_{kL+1} = 1, \ldots, X_{(k+1)L} = 1\}^c\right)\\ &= (1 - p^k)^n,\end{aligned}$$

from which we obtain $\mathbb{P}(T_{0,L} = \infty \mid S_0 = k) = 0$ after letting n tend to infinity when $p \in [0, 1)$, hence $\mathbb{P}(T_{0,L} < \infty \mid S_0 = k) = 1$. In case $p = 1$, we clearly have $\mathbb{P}(T_{0,L} < \infty \mid S_0 = k) = 1$.

Exercise 1.3 The chain has the following graph:



a) The absorbing states are 0 and 3.
b) By the example page 119 of Privault (2018) we have $g_0(1) = g_3(1) = 1/2$. On the other hand, we clearly have $g_1(0) = g_1(3) = 0$ and $g_1(1) = 1$, hence

$$g_1(2) = 0.3 \times g_1(0) + 0.4 \times g_1(1) + 0.3 \times g_1(3) = 0.4.$$

c) We clearly have $p_1(0) = p_1(3) = 0$, and

$$\begin{cases} p_1(1) = 0.3 \times p_1(0) + 0.4 \times p_1(2) + 0.3 \times p_1(3) = 0.4 \times p_1(2) \\ p_1(2) = 0.3 \times p_1(0) + 0.4 + 0.3 \times p_1(3) = 0.4, \end{cases}$$

hence $p_1(1) = 0.16$.

d) We have $h_1(1) = 0$ by construction and $h_1(0) = h_1(3) = +\infty$ because states (0) and (3) are absorbing, and $h_1(2) = +\infty$ because $g_0(2) \geq 0.3 > 0$. Regarding mean return times, we have $\mu_1(0) = \mu_1(1) = \mu_1(2) = \mu_1(3) = +\infty$ because states (1) and (2) communicate while states (0) and (3) are absorbing.

Exercise 1.4

a) The boundary conditions are given by

$$f(x,0) = -x \qquad \text{and} \qquad f(0,y) = y, \qquad x,y \geq 0.$$

b) The finite difference equation satisfied by $f(x,y)$ is given by

$$f(x,y) = \frac{x}{x+y}(f(x-1,y)-1) + \frac{y}{x+y}(f(x,y-1)+1), \quad x,y \geq 1.$$

c) We have

$$\begin{cases} f(1,1) = \frac{1}{2}(f(0,1)-1) + \frac{1}{2}(f(1,0)+1) = 0, \\ \\ f(1,2) = \frac{1}{3}(f(0,2)-1) + \frac{2}{3}(f(1,1)+1) = 1, \\ \\ f(2,2) = \frac{1}{2}(f(1,2)-1) + \frac{1}{2}(f(2,1)+1) = 0, \\ \\ f(1,3) = \frac{1}{4}(f(0,3)-1) + \frac{3}{4}(f(1,2)+1) = 2, \\ \\ f(2,3) = \frac{2}{5}(f(1,3)-1) + \frac{3}{5}(f(2,2)+1) = 1, \\ \\ f(3,3) = \frac{1}{2}(f(2,3)-1) + \frac{1}{2}(f(3,2)+1) = 0. \end{cases}$$

d) We check that $f(x,y) := y - x$ solves the finite difference equation

$$\begin{aligned} &\frac{x}{x+y}(f(x-1,y)-1) + \frac{y}{x+y}(f(x,y-1)+1) \\ &= \frac{x}{x+y}(y-(x-1)-1) + \frac{y}{x+y}(y-1-x+1) \\ &= \frac{x}{x+y}(y-x) + \frac{y}{x+y}(y-x) \\ &= y-x \\ &= f(x,y), \end{aligned}$$

with the correct boundary conditions.

Exercise 1.5

a) It clearly takes S steps for Buffalo A to travel up from 0 to S, and for Buffalo B to travel down from S to 0 ?
b) After the buffalos collide they can be assumed to both continue their way without any impact on their travel times to the boundary $\{0, S\}$, therefore the answer is S steps in this case as well.

Exercise 1.6

a) By a recurrence using Pascal's identity

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k},$$

we find

$$[P^n]_{i,j} = \begin{cases} p^{j-i} q^{n-(j-i)} \binom{n}{j-i}, & 0 \le j-i \le n, \\ 0, & n < j-i, \\ 0, & i > j. \end{cases}$$

b) We have

$$\begin{aligned} 0 &\le \lim_{n\to\infty} [P^n]_{i,j} \\ &= \frac{p^{j-i} q^{-(j-i)}}{(j-i)!} \lim_{n\to\infty} q^n \frac{n!}{(n-(j-i))!} \\ &= \lim_{n\to\infty} q^n n(n-1)\cdots(n-(j-i)+1) \\ &\le \lim_{n\to\infty} q^n n^{j-i} \\ &= \lim_{n\to\infty} e^{\log(q^n n^{j-i})} \\ &= \lim_{n\to\infty} e^{n\log q + (j-i)\log n} \\ &= 0, \qquad 0 \le j-i. \end{aligned}$$

c) We have

$$\sum_{n\ge 0} [P^n]_{i,j} = \begin{cases} \displaystyle\sum_{n\ge j-i} p^{j-i} q^{n-(j-i)} \binom{n}{j-i}, & i \le j, \\ 0, & i > j, \end{cases}$$

$$= \begin{cases} \dfrac{p^{j-i}}{(j-i)!} \sum_{n \geq 0} q^n \dfrac{(n+j-i)!}{n!}, & i \leq j, \\ 0, & i > j, \end{cases}$$

$$= \begin{cases} \dfrac{p^{j-i}}{(j-i)!} \sum_{n \geq 0} q^n \dfrac{(n+j-i)!}{n!}, & i \leq j, \\ 0, & i > j, \end{cases}$$

$$= \begin{cases} \dfrac{p^{j-i}}{(j-i)!} \dfrac{\partial^{j-i}}{\partial q^{j-i}} \dfrac{1}{1-q}, & i \leq j, \\ 0, & i > j, \end{cases}$$

$$= \begin{cases} \dfrac{p^{j-i}}{(1-q)^{j-i+1}}, & i \leq j, \\ 0, & i > j, \end{cases}$$

$$= \begin{cases} \dfrac{1}{p}, & i \leq j, \\ 0, & i > j. \end{cases}$$

d) We have

$$p_{i,j} = \mathbb{P}(T_j < \infty \mid X_0 = i) = \begin{cases} 1, & i < j, \\ q < 1, & i = j, \\ 0, & i > j. \end{cases}$$

e) Since $p_{i,i} = q < 1$ for all $i \geq 0$, the chain $(X_n)_{n \geq 0}$ is transient as all of its states are transient.

f) As in Proposition 1.7, the mean number of returns from state i to state j is given by

$$\sum_{n \geq 1} [P^n]_{i,j} = \mathbb{E}[R_j \mid X_0 = i] = \begin{cases} p \sum_{n \geq 1} n q^{n-1} = \dfrac{1}{p} = \dfrac{p_{i,j}}{1 - p_{j,j}}, & i < j, \\ qp \sum_{n \geq 1} n q^{n-1} = \dfrac{q}{p} = \dfrac{p_{i,i}}{1 - p_{i,i}}, & i = j, \\ 0 = \dfrac{p_{i,j}}{1 - p_{j,j}}, & i > j. \end{cases}$$

g) The matrix

$$I - P = \begin{bmatrix} 1-q & -p & 0 & 0 & \cdots \\ 0 & 1-q & -p & 0 & \cdots \\ 0 & 0 & 1-q & -p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} p & -p & 0 & 0 & \cdots \\ 0 & p & -p & 0 & \cdots \\ 0 & 0 & p & -p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

is invertible, and as in (1.38), its inverse can be expressed as

$$\begin{aligned} (I-P)^{-1} &= \left[\sum_{n\geq 0} [P^n]_{i,j} \right]_{i,j\in\mathbb{N}} \\ &= \left[\mathbf{1}_{\{i=j\}} + \mathbb{E}[R_j \mid X_0 = i] \right]_{i,j\in\mathbb{N}} \\ &= \begin{bmatrix} \frac{1}{p} & \frac{1}{p} & \frac{1}{p} & \frac{1}{p} & \frac{1}{p} & \cdots \\ 0 & \frac{1}{p} & \frac{1}{p} & \frac{1}{p} & \frac{1}{p} & \cdots \\ 0 & 0 & \frac{1}{p} & \frac{1}{p} & \frac{1}{p} & \cdots \\ 0 & 0 & 0 & \frac{1}{p} & \frac{1}{p} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \end{aligned}$$

Note that although the vector $e = (1, 1, 1, \ldots)$ satisfies $(I - P)e = 0$ it does not belong to $\ell^1(\mathbb{N})$, and $I - P$ is invertible as an operator from $\ell^1(\mathbb{N})$ into

$$\left\{ (u_k)_{k\geq 0} \, : \, \sum_{n\geq 0} \Big| \sum_{k\geq n} u_k \Big| < \infty \right\}.$$

Exercise 1.7

a) We have $\mu_A(x, y) = 0$ for all $(x, y) \in A$.
b) For all $0 \leq x, y \leq 3$ we have

$$\mu_A(x,y) = 1 + \frac{1}{2}\mu_A(x+1,y) + \frac{1}{2}\mu_A(x,y+1). \tag{S.1.1}$$

c) We have

$$
\begin{cases}
\mu_A(2,2) = 1 + \dfrac{1}{2}\mu_A(3,2) + \dfrac{1}{2}\mu_A(2,3) = 1, \\
\mu_A(1,2) = 1 + \dfrac{1}{2}\mu_A(2,2) + \dfrac{1}{2}\mu_A(1,3) = \dfrac{3}{2}, \\
\mu_A(2,1) = 1 + \dfrac{1}{2}\mu_A(2,2) + \dfrac{1}{2}\mu_A(3,1) = \dfrac{3}{2}, \\
\mu_A(0,2) = 1 + \dfrac{1}{2}\mu_A(1,2) + \dfrac{1}{2}\mu_A(0,3) = \dfrac{7}{4}, \\
\mu_A(2,0) = 1 + \dfrac{1}{2}\mu_A(2,1) + \dfrac{1}{2}\mu_A(3,0) = \dfrac{7}{4}, \\
\mu_A(1,1) = 1 + \dfrac{1}{2}\mu_A(2,1) + \dfrac{1}{2}\mu_A(1,2) = \dfrac{5}{2}, \\
\mu_A(0,1) = 1 + \dfrac{1}{2}\mu_A(1,1) + \dfrac{1}{2}\mu_A(0,2) = \dfrac{25}{8}, \\
\mu_A(1,0) = 1 + \dfrac{1}{2}\mu_A(1,1) + \dfrac{1}{2}\mu_A(2,0) = \dfrac{25}{8}, \\
\mu_A(0,0) = 1 + \dfrac{1}{2}\mu_A(1,0) + \dfrac{1}{2}\mu_A(0,1) = \dfrac{33}{8}.
\end{cases}
$$

4	0	0	0	0	0
3	0	0	0	0	0
2	7/4	3/2	1	0	0
1	25/8	5/2	3/2	0	0
0	33/8	25/8	7/4	0	0
	0	1	2	3	4

Table 1.1: Values of $\mu_A(x, y)$ with $N = 3$ and the set A in blue.

d) The mean number of rounds is $\mu_A(0,0) = 33/8 = 4.125$.

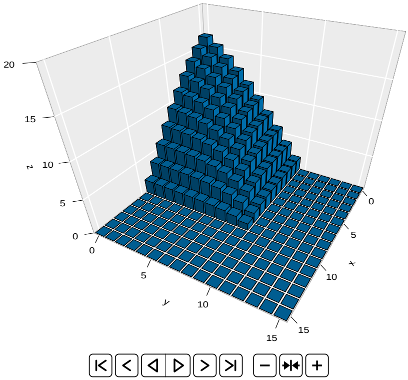


Fig. S.1: Backward solution of Equation (S.1.1) for $\mu_A(x,y)$ with $N = 10$.*

The following R code can be used to generate Figure S.1.

```
install.packages("plot3D"); require(plot3D);N=10;M=15
X=array(1:2,c(M+1,M+1));
for (i in seq(1,M+1)) {for (j in seq(1,M+1)) X[i,j]=0;}
par(mar=c(1,2,0,0)+0.01)
for (k in seq(N,-N)) {for (i in seq(k,N)) {
if (i>=1 && N+k-i>=1)
       {X[i,N+k-i]=1+(X[i+1,N+k-i]+X[i,N+k-i+1])/2.0;dev.hold();
hist3D(x=0:M, y=0:M, z=X, scale=T, bty="g", phi=35, theta=120,
       border="black", zlim=c(0,20), shade=0.3, space=0.15, col="#0072B2",
       colkey=F, ticktype="detailed"); dev.flush();}}}
```

Exercise 1.8

a) When $X_0 = x \geq 2$ and $Y_0 = y \geq 2$ we have $T_A = 0$, hence

$$\mu_A(x,y) := \mathbb{E}[T_A < \infty \mid X_0 = x,\ Y_0 = y] = 0, \qquad x \geq 2, \quad y \geq 2.$$

b) This equation is obtained by first step analysis, noting that we can only move up to to the right with probability 1/2 in both cases.

c) We note that $\mu_A(x,y) = \mu_A(x,y+1)$ for $y \geq 2$, and

$$\mu_A(1,y) = 1 + \frac{1}{2}\mu_A(2,y) + \frac{1}{2}\mu_A(1,y+1) = 1 + \frac{1}{2}\mu_A(1,y), \quad y \geq 2,$$

hence $\mu_A(1,y) = 2$ for all $y \geq 2$. We also have

* Animated figure (works in Acrobat Reader).

$$\mu_A(0,y) = 1 + \frac{1}{2}\mu_A(1,y) + \frac{1}{2}\mu_A(0,y+1) = 2 + \frac{1}{2}\mu_A(0,y), \quad y \geq 2,$$

hence $\mu_A(0,y) = 4$, $y \geq 2$. By symmetry we also have $\mu_A(x,1) = 2$ and $\mu_A(x,0) = 4$ for all $x \geq 2$.

These results can also be recovered using pathwise analysis as

$$\mu_A(1,y) = \sum_{k\geq 1} \frac{k}{2^k} = \frac{1}{2}\sum_{k\geq 0} \frac{k}{2^{k-1}} = \frac{1}{2(1-1/2)^2} = 2, \quad y \geq 2,$$

which yields similarly $\mu_A(x,1) = 2$ for all $x \geq 2$. Repeating this argument once also leads to $\mu_A(x,0) = \mu_A(0,y) = 4$ for all $x, y \geq 2$.

d) We have

$$\begin{cases} \mu_A(1,1) = 1 + \frac{1}{2}\mu_A(2,1) + \frac{1}{2}\mu_A(1,2) = 3, \\ \\ \mu_A(0,1) = 1 + \frac{1}{2}\mu_A(1,1) + \frac{1}{2}\mu_A(0,2) = \frac{9}{2}, \\ \\ \mu_A(1,0) = 1 + \frac{1}{2}\mu_A(2,0) + \frac{1}{2}\mu_A(1,1) = \frac{9}{2}, \\ \\ \mu_A(0,0) = 1 + \frac{1}{2}\mu_A(1,0) + \frac{1}{2}\mu_A(0,1) = \frac{11}{2}, \end{cases}$$

hence the mean time it takes until both cans contain at least \$2 is $\mu_A(0,0) = 11/2$.

4	4	2	0	0	0
3	4	2	0	0	0
2	4	2	0	0	0
1	9/2	3	2	2	2
0	11/2	9/2	4	4	4
	0	1	2	3	4

Table 1.2: Values of $\mu_A(x,y)$ with $N = 2$ and the set A in blue.

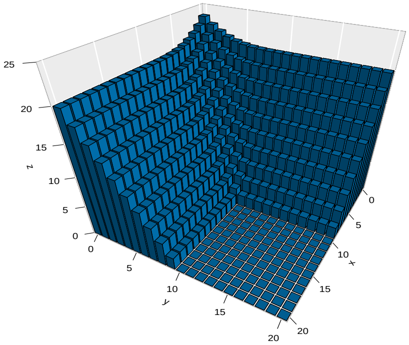


Fig. S.2: Backward solution of (1.51) for $\mu_A(x, y)$ with $N = 10$.*

The following R code can be used to generate Figure S.2.

```
require(plot3D);N=10;M=20;X=array(1:2,c(M+1,M+1));
for (i in seq(N+2,M+1)) {for (j in seq(N+2,M+1)) X[i,j]=0;}
for (i in seq(N+1,M+1)) {for (j in seq(1,N+1)) X[i,j]=2*(N+1-j);}
for (i in seq(1,N+1)) {for (j in seq(N+1,M+1)) X[i,j]=2*(N+1-i);}
for (k in seq(N,-N)) {for (i in seq(k,N)) {if (i>=1 && N+k-i>=1)
      X[i,N+k-i]=1+(X[i+1,N+k-i]+X[i,N+k-i+1])/2.0;}}
hist3D(x=1:21, y=1:21, z=X, scale=T, bty="g", phi=35, theta=120,
      border="black", zlim=c(0,25), shade=0.3, space=0.15, col="#0072B2",
      colkey=F, ticktype="detailed")
```

Problem 1.9

a) We have $f_{i,j}^{(1)} = P_{i,j}$, $i, j \in \mathbb{S}$.
b) We have

$$\begin{aligned}
f_{i,j}^{(n+1)} &= \mathbb{P}(X_{n+1} = j, X_n \neq j, \ldots, X_1 \neq j \mid X_0 = i) \\
&= \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} \mathbb{P}(X_{n+1} = j, X_n \neq j, \ldots, X_2 \neq j \mid X_1 = k) \\
&= \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} \mathbb{P}(X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j \mid X_0 = k) \\
&= \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} f_{k,j}^{(n)}, \qquad i, j \in \mathbb{S}, \ n \geq 1.
\end{aligned}$$

* Animated figure (works in Acrobat Reader).

c) By summing (1.53) over $n \geq 1$, we find

$$\begin{aligned}
f_{i,j} &= \sum_{n\geq 1} f_{i,j}^{(n)} \\
&= f_{i,j}^{(1)} + \sum_{n\geq 2} f_{i,j}^{(n)} \\
&= P_{i,j} + \sum_{n\geq 1} f_{i,j}^{(n+1)} \\
&= P_{i,j} + \sum_{n\geq 1} \sum_{\substack{k\in\mathbb{S} \\ k\neq j}} P_{i,k} f_{k,j}^{(n)} \\
&= P_{i,j} + \sum_{\substack{k\in\mathbb{S} \\ k\neq j}} P_{i,k} f_{k,j}, \qquad i,j \in \mathbb{S}.
\end{aligned}$$

d) Let $\tilde{f}$ denote another solution of (1.54). We have $\tilde{f}_{i,j} \geq P_{i,j} = f_{i,j}^{(1)}$, and if $\tilde{f}_{i,j} \geq \sum_{l=1}^{n} f_{i,j}^{(l)}$ then by (1.53) and (1.54) we have

$$\begin{aligned}
\tilde{f}_{i,j} &= P_{i,j} + \sum_{\substack{k\in\mathbb{S} \\ k\neq j}} P_{i,k} \tilde{f}_{k,j} \\
&\geq P_{i,j} + \sum_{\substack{k\in\mathbb{S} \\ k\neq j}} P_{i,k} \sum_{l=1}^{n} f_{k,j}^{(l)} \\
&= P_{i,j} + \sum_{l=1}^{n} \sum_{\substack{k\in\mathbb{S} \\ k\neq j}} P_{i,k} f_{k,j}^{(l)} \\
&= P_{i,j} + \sum_{l=1}^{n} f_{i,j}^{(l+1)} \\
&= P_{i,j} + \sum_{l=2}^{n+1} f_{i,j}^{(l)} \\
&= \sum_{l=1}^{n+1} f_{i,j}^{(l)}
\end{aligned}$$

hence by induction we obtain

$$\tilde{f}_{i,j} \geq \sum_{l=1}^{n} f_{i,j}^{(l)}, \quad i,j \in \mathbb{S}, \quad n \geq 1,$$

and letting n tend to infinity, we find

$$\tilde{f}_{i,j} \geq \sum_{l=1}^{\infty} f_{i,j}^{(l)} = f_{i,j}, \quad i,j \in \mathbb{S}.$$

Finally, we check that if f and g are two minimal solutions then $f \geq g$ and $g \geq f$, hence $f = g$ and the minimal solution is unique.

e) The condition $g_{i,j}^{(1)} = f_{i,j}^{(1)}$ is satisfied by construction, for $i,j \in \mathbb{S}$. Next, assuming that $g_{i,j}^{(n)} = n f_{i,j}^{(n)}$, $i,j \in \mathbb{S}$, we have

$$\begin{aligned} g_{i,j}^{(n+1)} &= f_{i,j}^{(n+1)} + n \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} f_{k,j}^{(n)} \\ &= f_{i,j}^{(n+1)} + n f_{k,j}^{(n+1)} \\ &= (n+1) f_{i,j}^{(n+1)}, \quad i,j \in \mathbb{S}, \quad n \geq 1. \end{aligned}$$

f) We have

$$\begin{aligned} h_{i,j} &= \sum_{n \geq 1} g_{i,j}^{(n)} \\ &= g_{i,j}^{(1)} + \sum_{n \geq 1} g_{i,j}^{(n+1)} \\ &= f_{i,j}^{(1)} + \sum_{n \geq 1} \left(f_{i,j}^{(n+1)} + n \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} f_{k,j}^{(n)} \right) \\ &= \sum_{n \geq 1} f_{i,j}^{(n)} + \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} \sum_{n \geq 1} n f_{k,j}^{(n)} \\ &= f_{i,j} + \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} h_{k,j}, \qquad i,j \in \mathbb{S}. \end{aligned}$$

g) By (1.55), for $n = 1$ we have

$$\begin{aligned} \tilde{h}_{i,j} &= f_{i,j} + \sum_{\substack{k \in \mathbb{S} \\ k \neq j}} P_{i,k} \tilde{h}_{k,j} \\ &\geq f_{i,j} \\ &\geq f_{i,j}^{(1)} \\ &= g_{i,j}^{(1)}. \end{aligned}$$

Next, assuming that

$$\tilde{h}_{i,j} \geq \sum_{l=1}^{n} g_{i,j}^{(l)}, \qquad i,j \in \mathbb{S},$$

holds at the rank $n \geq 1$, we have

$$\begin{aligned}
\tilde{h}_{i,j} &= f_{i,j} + \sum_{\substack{k\in\mathbb{S}\\k\neq j}} P_{i,k}\tilde{h}_{k,j} \\
&\geq f_{i,j} + \sum_{\substack{k\in\mathbb{S}\\k\neq j}} P_{i,k} \sum_{l=1}^{n} g_{k,j}^{(l)} \\
&= f_{i,j} + \sum_{l=1}^{n} \sum_{\substack{k\in\mathbb{S}\\k\neq j}} P_{i,k} g_{k,j}^{(l)} \\
&= f_{i,j} + \sum_{l=1}^{n} l \sum_{\substack{k\in\mathbb{S}\\k\neq j}} P_{i,k} f_{k,j}^{(l)} \\
&= f_{i,j} + \sum_{l=1}^{n} l f_{i,j}^{(l+1)} \\
&= f_{i,j} + \sum_{l=2}^{n+1} (l-1) f_{i,j}^{(l)} \\
&\geq \sum_{l=1}^{n+1} f_{i,j}^{(l)} + \sum_{l=2}^{n+1} (l-1) f_{i,j}^{(l)} \\
&= f_{i,j}^{(1)} + \sum_{l=2}^{n+1} l f_{i,j}^{(l)} \\
&= \sum_{l=1}^{n+1} g_{i,j}^{(l)}, \qquad i,j \in \mathbb{S}.
\end{aligned}$$

Letting n tend to infinity, we find

$$\tilde{h}_{i,j} \geq \sum_{l=1}^{\infty} g_{i,j}^{(l)} = h_{i,j}, \qquad i,j \in \mathbb{S},$$

proving that $h_{i,j}$ is a minimal solution to (1.55). Finally, we check that if f and g are two minimal solutions then $f \geq g$ and $g \geq f$, hence $f = g$ and the minimal solution is unique.

Chapter 2 - Phase-Type Distributions

Exercise 2.1

a) We have $h_3(3) = 0$, and

$$\begin{cases} h_3(1) = 1 + (1-p)h_3(1) + ph_3(2) \\ \\ h_3(2) = 1 + (1-q)h_3(1), \end{cases}$$

hence

$$\begin{cases} h_3(1) = 1 + (1-p)h_3(1) + ph_3(2) = 1 + p + ((1-p) + (1-q)p)h_3(1) \\ \\ h_3(2) = 1 + (1-q)h_3(1), \end{cases}$$

hence

$$\begin{cases} h_3(1) = \dfrac{1+p}{1-(1-p)-(1-q)p} = \dfrac{1+p}{pq} \\ \\ h_3(2) = 1 + \dfrac{(1+p)(1-q)}{1-((1-p)+(1-q)p)} = \dfrac{1+p-q}{pq}. \end{cases}$$

b) We have

$$\begin{cases} G_1(s) = (1-p)\,\mathbb{E}\left[s^{1+T_3} \,\middle|\, X_0 = 1\right] + p\,\mathbb{E}\left[s^{1+T_3} \,\middle|\, X_0 = 2\right] \\ \\ G_2(s) = (1-q)\,\mathbb{E}\left[s^{1+T_3} \,\middle|\, X_0 = 1\right] + qs, \end{cases}$$

i.e.

$$\begin{cases} G_1(s) = (1-p)s\,\mathbb{E}\left[s^{T_3} \,\middle|\, X_0 = 1\right] + ps\,\mathbb{E}\left[s^{T_3} \,\middle|\, X_0 = 2\right] \\ \\ G_2(s) = (1-q)s\,\mathbb{E}\left[s^{T_3} \,\middle|\, X_0 = 1\right] + qs, \end{cases}$$

hence

$$\begin{cases} G_1(s) = (1-p)sG_1(s) + psG_2(s) \\ \\ G_2(s) = (1-q)sG_1(s) + qs \end{cases}$$

or

$$\begin{cases} G_1(s) = (1+ps)G_2(s) - qs \\ \\ G_2(s) = (1-q)sG_1(s) + qs \end{cases}$$

i.e.

$$\begin{cases} G_1(s) = (1+ps)(1-q)sG_1(s) + qs(1+ps) - qs \\ \\ G_2(s) = (1-q)s(1+ps)G_2(s) - q(1-p)s^2 + qs, \end{cases}$$

hence

$$\begin{cases} G_1(s) = \dfrac{pqs^2}{1-(1-p)s-p(1-q)s^2} \\ \\ G_2(s) = \dfrac{-q(1-p)s^2+qs}{1-(1-q)s(1+ps)} \end{cases}$$

c) Using the identity

$$\frac{\sqrt{(1-p)^2+4(1-q)p}}{1-(1-p)s-p(1-q)s^2} = \sum_{n=0}^{\infty} \frac{s^n}{z_+^{n+1}} - \sum_{n=0}^{\infty} \frac{s^n}{z_-^{n+1}},$$

we find

$$\begin{aligned} G_1(s) &= \frac{pqs^2}{1-(1-p)s-p(1-q)s^2} \\ &= \frac{pqs^2}{\sqrt{(1-p)^2+4(1-q)p}} \sum_{n=0}^{\infty} \left(\frac{s^n}{z_+^{n+1}} - \frac{s^n}{z_-^{n+1}} \right) \\ &= \frac{pq}{\sqrt{(1-p)^2+4(1-q)p}} \sum_{n=2}^{\infty} \left(\frac{s^n}{z_+^{n-1}} - \frac{s^n}{z_-^{n-1}} \right) \\ &= \sum_{n=0}^{\infty} s^n \mathbb{P}(T_3 = n \mid X_0 = 1), \qquad -1 \leq s \leq 1, \end{aligned}$$

hence by identification we find $\mathbb{P}(T_3 = n \mid X_0 = 1) = 0$, $n = 0, 1$, and

$$\mathbb{P}(T_3 = n \mid X_0 = 1) = \frac{pq}{\sqrt{(1-p)^2+4(1-q)p}} \left(\frac{1}{z_+^{n-1}} - \frac{1}{z_-^{n-1}} \right), \quad n \geq 2.$$

In particular, this recovers

$$\begin{aligned} \mathbb{P}(T_3 = 2 \mid X_0 = 1) &= \frac{pq}{\sqrt{(1-p)^2+4(1-q)p}} \left(\frac{1}{z_+} - \frac{1}{z_-} \right) \\ &= \frac{pq}{\sqrt{(1-p)^2+4(1-q)p}} \frac{z_- - z_+}{z_- z_+} \\ &= pq. \end{aligned}$$

d) We note that the hitting time is *a.s.*[*] finite, *i.e.* $\mathbb{P}(T_3 < \infty \mid X_0 = 1) = 1$, hence the mean hitting time $\mathbb{E}[T_3 \mid X_0 = 1]$ is given from (A.14) as

$$\mathbb{E}[T_3 \mid X_0 = 1] = G_1'(1)$$

[*] Almost surely, *i.e.* with probability one.

$$
\begin{aligned}
&= \frac{2pqs}{1-(1-p)s-p(1-q)s^2}_{|s=1} \\
&\quad + \frac{pqs^2(1-p)+2p(1-q)s}{(1-(1-p)s-p(1-q)s^2)^2}_{|s=1} \\
&= \frac{1+p}{pq}.
\end{aligned}
$$

Chapter 3 - Synchronizing Automata

Exercise 3.1

a) We have

$$
\mathbb{E}[T^{(m)}] = mp^m + \sum_{k=0}^{m-1} p^k q(k+1+\mathbb{E}[T^{(m)}]).
$$

b) We find

$$
\begin{aligned}
\mathbb{E}[T^{(m)}] &= \frac{mp^m + q\displaystyle\sum_{k=0}^{m-1} p^k(k+1)}{1-q\displaystyle\sum_{k=0}^{m-1} p^k} \\
&= \frac{mp^m + \dfrac{1-(m+1)p^m+mp^{m+1}}{1-p}.}{p^m} \\
&= \frac{m(1-p)p^m + 1-(m+1)p^m+mp^{m+1}}{(1-p)p^m} \\
&= \frac{1/p^m-1}{1-p} \\
&= \sum_{k=1}^{m} \frac{1}{p^k}.
\end{aligned} \tag{S.3.2}
$$

Alternative solution: We note the recurrence relation

$$
\mathbb{E}[T^{(m)}] = \mathbb{E}[T^{(m-1)}] + p\times 1 + (1-p)(1+\mathbb{E}[T^{(m)}]), \qquad m \geq 2,
$$

which rewrites as

$$
\mathbb{E}[T^{(m)}] = \frac{\mathbb{E}[T^{(m-1)}]+1}{p}, \qquad m \geq 2,
$$

and also recovers (S.3.2) from $\mathbb{E}[T^{(0)}] = 0$.

Exercise 3.2

a) The sequence $(Z_n)_{n\geq 0}$ is a Markov chain since every new transition is determined by the current state, and its transition matrix P is given by

$$P = \begin{bmatrix} q & p & 0 & \cdots & \cdots & 0 & 0 \\ q & 0 & p & \cdots & \cdots & 0 & 0 \\ q & 0 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ q & 0 & \cdots & \ddots & p & 0 & 0 \\ q & 0 & \cdots & \cdots & 0 & p & 0 \\ q & 0 & \cdots & \cdots & 0 & 0 & p \end{bmatrix},$$

b) By first step analysis, the mean hitting times $\mathbb{E}\left[T^{(m)} \mid Z_0 = l\right]$, $l = 0, 1, \ldots, m$, satisfy the equations

$$\begin{cases} \mathbb{E}\left[T^{(m)} \mid Z_0 = 0\right] = 1 + (1-p)\,\mathbb{E}\left[T^{(m)} \mid Z_0 = 0\right] + p\,\mathbb{E}\left[T^{(m)} \mid Z_0 = 1\right] \\ \mathbb{E}\left[T^{(m)} \mid Z_0 = 1\right] = 1 + (1-p)\,\mathbb{E}\left[T^{(m)} \mid Z_0 = 0\right] + p\,\mathbb{E}\left[T^{(m)} \mid Z_0 = 2\right] \\ \qquad \vdots \\ \mathbb{E}\left[T^{(m)} \mid Z_0 = m-1\right] = 1 + (1-p)\,\mathbb{E}\left[T^{(m)} \mid Z_0 = 0\right] + p\,\mathbb{E}\left[T^{(m)} \mid Z_0 = m\right] \\ \mathbb{E}\left[T^{(m)} \mid Z_0 = m\right] = 0, \end{cases}$$

i.e.

$$\begin{cases} \mathbb{E}[T_m \mid Z_0 = 0] = \dfrac{1}{p} + \mathbb{E}[T_m \mid Z_0 = 1] \\ p\,\mathbb{E}[T_m \mid Z_0 = 1] = p\,\mathbb{E}[T_m \mid Z_0 = 2] + \mathbb{E}[T_m \mid Z_0 = 0] - \mathbb{E}[T_m \mid Z_0 = 1] \\ \qquad \vdots \\ p\,\mathbb{E}[T_m \mid Z_0 = m-1] = p\,\mathbb{E}[T_m \mid Z_0 = m] \\ \qquad\qquad + \mathbb{E}[T_m \mid Z_0 = m-2] - \mathbb{E}[T_m \mid Z_0 = m-1] \\ \mathbb{E}[T_m \mid Z_0 = m] = 0, \end{cases}$$

or

$$\begin{cases} \mathbb{E}\left[T^{(m)} \mid Z_0 = 0\right] = \dfrac{1}{p} + \mathbb{E}\left[T^{(m)} \mid Z_0 = 1\right] \\ \mathbb{E}\left[T^{(m)} \mid Z_0 = 1\right] = \dfrac{1}{p} + \mathbb{E}\left[T^{(m)} \mid Z_0 = 2\right] \\ \qquad \vdots \\ \mathbb{E}\left[T^{(m)} \mid Z_0 = m-1\right] = \dfrac{1}{p} + \mathbb{E}\left[T^{(m)} \mid Z_0 = m\right], \\ \mathbb{E}\left[T^{(m)} \mid Z_0 = m\right] = 0, \end{cases}$$

with solution

$$\begin{aligned}
\mathbb{E}\left[T^{(m)} \mid Z_0 = k\right] &= \sum_{l=k+1}^{m} \frac{1}{p^l} \\
&= \frac{1}{p^{k+1}} \sum_{l=0}^{m-k-1} \frac{1}{p^l} \\
&= \frac{1-(1/p)^{m-k}}{(1-1/p)p^{k+1}} \\
&= \frac{1-p^{m-k}}{(1-p)p^m}, \qquad k = 0, 1, \ldots, m.
\end{aligned}$$

c) We have

$$\begin{aligned}
\mathbb{E}\left[T^{(m)}\right] &= \mathbb{E}\left[T^{(m)} \mid Z_0 = 0\right] \\
&= \sum_{l=1}^{m} \frac{1}{p^l} \\
&= \frac{1-(1/p)^m}{(1-1/p)p} \\
&= \frac{1-p^m}{(1-p)p^m}.
\end{aligned}$$

Problem 3.3

a) The transition matrix is given by

$$\begin{array}{c c}
 & \begin{array}{cccc} aa & ab & ba & bb \end{array} \\
\begin{array}{c} aa \\ ab \\ ba \\ bb \end{array} & \begin{bmatrix} p & q & 0 & 0 \\ 0 & 0 & p & q \\ p & q & 0 & 0 \\ 0 & 0 & p & q \end{bmatrix}
\end{array}.$$

b) We have $\tau_{ab} = 1$ with probability one, hence

$$G_{ab}(s) = \mathbb{E}[s \mid Z_1 = (a, b)] = s.$$

c) We find

$$\begin{cases} G_{aa}(s) = psG_{aa}(s) + qsG_{ab}(s), \\ G_{ba}(s) = psG_{aa}(s) + qsG_{ab}(s). \end{cases}$$

d) We have

$$\begin{cases} G_{aa}(s) = psG_{aa}(s) + qs^2, \\ \\ G_{ba}(s) = psG_{aa}(s) + qs^2, \end{cases}$$

hence

$$G_{aa}(s) = G_{ba}(s) = \frac{pqs^3}{1-ps} + qs^2 = \frac{qs^2}{1-ps}, \qquad s \in (-1, 1).$$

We note that

$$\begin{aligned} \mathbb{P}(\tau_{ab} < \infty \mid Z_1 = (a, a)) &= \mathbb{P}(\tau_{ab} < \infty \mid Z_1 = (b, a)) \\ &= G_{ba}(1^-) \\ &= \lim_{s \nearrow 1} G_{ba}(s) \\ &= \lim_{s \nearrow 1} \frac{qs^2}{1-ps} \\ &= \frac{q}{1-p} \\ &= 1. \end{aligned}$$

e) We have

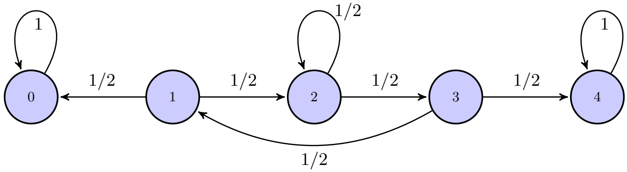
$$\begin{aligned} \mathbb{E}[\tau_{ab} \mid Z_1 = (a, a)] &= \mathbb{E}[\tau_{ab} \mid Z_1 = (b, a)] \\ &= G'_{ba}(1) = G'_{aa}(1) \\ &= \frac{2q}{1-p} + \frac{pq}{(1-p)^2} = 2 + \frac{p}{q}. \end{aligned}$$

f) This average time is

$$p\,\mathbb{E}[\tau_{ab} \mid Z_1 = (a, a)] + q\,\mathbb{E}[\tau_{ab} \mid Z_1 = (a, b)] = p\left(2 + \frac{p}{q}\right) + q = 1 + \frac{p}{q}.$$

Exercise 3.4

a) The word "abb" synchronizes to state 4 starting from states 1 and 2. However, the unique shortest word that synchronizes to state 4 starting from all states 1, 2 and 3 is "*aabb*".
b) The process $(Z_k)_{k \geq 0}$ is a Markov chain on the state space $\{0, 1, 2, 3, 4\}$, with the following graph:



The transition matrix of the chain $(Z_k)_{k\geq 0}$ is

$$\left[\, P_{i,j} \, \right]_{0\leq i,j\leq 4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

c) Denoting by $g_4(k)$ the probability that state (4) is reached first starting from state $k = 0, 1, 2, 3, 4$, we have the equations

$$\begin{cases} g_4(0) = 0 \\ \\ g_4(1) = \dfrac{1}{2}g_4(0) + \dfrac{1}{2}g_4(2) = \dfrac{1}{2}g_4(2) \\ \\ g_4(2) = \dfrac{1}{2}g_4(2) + \dfrac{1}{2}g_4(3) \\ \\ g_4(3) = \dfrac{1}{2}g_4(1) + \dfrac{1}{2}g_4(4) = \dfrac{1}{2}g_4(1) + \dfrac{1}{2} \\ \\ g_4(4) = 1, \end{cases}$$

with the solution

$$\begin{cases} g_4(0) = 0 \\ \\ g_4(1) = \dfrac{1}{3} \\ \\ g_4(2) = \dfrac{2}{3} \\ \\ g_4(3) = \dfrac{2}{3} \\ \\ g_4(4) = 1. \end{cases}$$

Hence the probability that the first synchronized word is "aabb" when the automaton is started from state ① is 1/3.

Exercise 3.5

a) The unique shortest word that synchronizes to state ④ starting from all states ①, ② and ③ is "*aba*".
b) By the same analysis as in Exercise 3.4-(c), the probability that the first synchronized word is "aba" when the automaton is started from state ① is 1/3.

Exercise 3.6 Denoting by $\lfloor x \rfloor = \max\{n \in \mathbb{Z} \ : \ n \leq x\}$ the integer *floor* of $x \in \mathbb{R}$, we have

$$\begin{aligned}
G_{T^{(m)}}(s) &= p^m s^m \frac{1-ps}{1-s+qp^m s^{m+1}} \\
&= p^m s^m (1-ps) \sum_{k\geq 0} s^k (1-qp^m s^m)^k \\
&= p^m s^m (1-ps) \sum_{k\geq 0} s^k \sum_{l=0}^{k} \binom{k}{l} (-qp^m s^m)^l \\
&= p^m s^m (1-ps) \sum_{n\geq 0} s^n \sum_{l=0}^{\lfloor n/(m+1)\rfloor} \binom{n-ml}{l} (-q)^l p^{ml} \\
&= p^m s^m \sum_{n\geq 0} s^n \left(\sum_{l=0}^{\lfloor n/(m+1)\rfloor} \binom{n-ml}{l} (-q)^l p^{ml} - p \sum_{l=0}^{\lfloor (n-1)/(m+1)\rfloor} \binom{n-1-ml}{l} (-q)^l p^{ml} \right),
\end{aligned}$$

$-1 \leq s \leq 1$, which shows that

$$\begin{aligned}
&\mathbb{P}\big(T^{(m)} = m+n\big) \\
&= p^m \left(\sum_{l=0}^{\lfloor n/m\rfloor} \binom{n-ml}{l} (-q)^l p^{ml} - p \sum_{l=0}^{\lfloor (n-1)/m\rfloor} \binom{n-1-ml}{l} (-q)^l p^{ml} \right) \\
&= p^m \sum_{l=0}^{\lfloor (n-1)/m\rfloor} \left(\binom{n-ml}{l} - p \binom{n-1-m[n/m]}{l} \right) (-q)^l p^{ml} \\
&\quad + p^m \mathbb{1}_{\{[n/m] > \lfloor (n-1)/(m+1)\rfloor\}} \binom{n-m\lfloor n/m\rfloor}{\lfloor n/m\rfloor} (-q)^{\lfloor n/m\rfloor} p^{m\lfloor n/m\rfloor}, \text{(S.3.3)}
\end{aligned}$$

and recovers in particular $\mathbb{P}\big(T^{(m)} = m\big) = p^m$ and

$$\mathbb{P}\big(T^{(m)} = m+n\big) = qp^m, \qquad n = 1, 2, \ldots, m,$$

and yields

$$\mathbb{P}\big(T^{(m)} = 2m+1\big) = (1-p^m)qp^m.$$

For $m = 1$ we also have

$$G_{T^{(m)}}(1) = ps\frac{1-ps}{1-s+qps^2} = \frac{ps}{1-qs} = \sum_{k\geq 1} s^k pq^{k-1},$$

and

$$\mathbb{P}\big(T^{(1)} = n\big) = pq^{n-1}, \qquad n \geq 1.$$

Chapter 4 - Random Walks and Recurrence

Exercise 4.1

a) By independence of the sequence $(X_k)_{1\leq k\leq n}$, we have

$$\begin{aligned}\mathbb{E}\left[\exp\left(t\sum_{k=1}^n X_k\right)\right] &= \prod_{k=1}^n \mathbb{E}[\mathrm{e}^{tX_k}] \\ &= (q+p\mathrm{e}^t)^n, \quad n\geq 0, \quad t\in\mathbb{R}.\end{aligned}$$

b) By the classical Markov or Chernoff bound argument, we have

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n (X_k - p) \geq z\right) &= \mathbb{P}\left(\exp\left(t\sum_{k=1}^n X_k\right) \geq \mathrm{e}^{ntz+npt}\right) \\ &= \mathrm{e}^{-ntz-npt}\,\mathbb{E}\left[\exp\left(t\sum_{k=1}^n X_k\right)\right] \\ &= \mathrm{e}^{-ntz-npt}(q+p\mathrm{e}^t)^n \\ &= \mathrm{e}^{-n(t(p+z)-\log(q+p\mathrm{e}^t))}, \qquad t>0.\end{aligned}$$

c) By differentiating $t\mapsto xt - \log(q+p\mathrm{e}^t)$ with respect to $t>0$, we find that the maximizing value $t(x)$ is given by

$$t(x) = \log\frac{qx}{(1-x)p}, \qquad x\in(0,1).$$

d) We have

$$\begin{aligned}&\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n (X_k - p) \geq z\right) \leq \mathrm{e}^{-n((p+z)t(x)-\log(q+p\mathrm{e}^{t(x)}))} \\ &= \exp\left(-n\left((p+z)\log\frac{(p+z)q}{(q-z)p} - \log\frac{q}{q-z}\right)\right), \qquad 0\leq z<q.\end{aligned}$$

e) Applying Taylor's formula with remainder

$$f(t) = f(0) + tf'(0) + \frac{t^2}{2} f''(\theta t)$$

to the function $f(t) := \log(q + p\mathrm{e}^t)$ with $f(0) = 0$, $f'(t) = p\mathrm{e}^t/(q + p\mathrm{e}^t)$, and $f''(t) = pq\mathrm{e}^t/(q + p\mathrm{e}^t)^2$, hence $f'(0) = p$ and

$$f''(\theta t) = \frac{pq\mathrm{e}^{\theta t}}{(q + p\mathrm{e}^{\theta t})^2} \leq \frac{1}{4},$$

we obtain

$$\log(q + p\mathrm{e}^t) = pt + \frac{t^2}{2} f''(\theta t) \leq pt + \frac{t^2}{8}, \qquad t \in \mathbb{R}.$$

The inequality $4pq\mathrm{e}^{\theta t} \leq (q + p\mathrm{e}^{\theta t})^2$ can be proved by noting that it is equivalent to $(q - p\mathrm{e}^{\theta t})^2 \geq 0$.

f) By differentiating $t \mapsto zt - t^2/8$ with respect to $t > 0$ we find that the maximizing value $t(z)$ is given by $t(z) = 4z$, $z \in (0, 1)$.

g) We have

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n}(X_k - p) \geq z\right) &\leq \mathrm{e}^{-n(t(p+z) - \log(q + p\mathrm{e}^t))} \\ &\leq \mathrm{e}^{-n(zt(z) - t(z)^2/8)} \\ &\leq \mathrm{e}^{-2nz^2}, \qquad z \geq 0.\end{aligned}$$

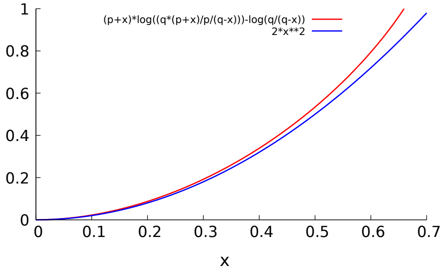


Fig. S.3: Comparison of rate functions.

Problem 4.2

a) If none of the stated conditions, hold, *i.e.* if

$$\widehat{m}_{n-1}^{(N,\alpha^*)}+\sqrt{\frac{2\log n}{T_{n-1}^{(N,\alpha^*)}}} > p_N,\;\; \widehat{m}_{n-1}^{(i,\alpha^*)} \le p_i+\sqrt{\frac{2\log n}{T_{n-1}^{(i,\alpha^*)}}},\;\; T_{n-1}^{(i,\alpha^*)} \ge \frac{2\log n}{(p_N-p_i)^2},$$

then we have

$$\begin{aligned}\widehat{m}_{n-1}^{(N,\alpha^*)} + \sqrt{\frac{2\log n}{T_{n-1}^{(N,\alpha^*)}}} &> p_N \\ &= p_i + p_N - p_i \\ &\ge p_i + \sqrt{\frac{2\log n}{T_{n-1}^{(i,\alpha^*)}}} \\ &\ge \widehat{m}_{n-1}^{(i,\alpha^*)},\end{aligned}$$

which implies $\alpha_n^* \neq i$.

b) We have

$$\begin{aligned}&\mathbb{E}[T_n^{(i,\alpha^*)}] = \mathbb{E}\left[\sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}}\right] \\ &\le \mathbb{E}\left[\sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} < \frac{2\log n}{(p_N-p_i)^2}\}}\right] + \mathbb{E}\left[\sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} \ge \frac{2\log n}{(p_N-p_i)^2}\}}\right] \\ &\le \widehat{n}_i + \mathbb{E}\left[\sum_{\widehat{n}_i<k\le n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} \ge \frac{2\log n}{(p_N-p_i)^2}\}}\right] \\ &\le \widehat{n}_i + \sum_{\widehat{n}_i<k\le n} \mathbb{P}\left(\widehat{m}_{k-1}^{(N,\alpha^*)} + \sqrt{\frac{2\log k}{T_{k-1}^{(N,\alpha^*)}}} \le p_N\right) \\ &\quad + \sum_{\widehat{n}_i<k\le n} \mathbb{P}\left(\widehat{m}_{k-1}^{(N,\alpha^*)} > p_i + \sqrt{\frac{2\log k}{T_{k-1}^{(i,\alpha^*)}}}\right).\end{aligned}$$

c) We have

$$\begin{aligned}&\mathbb{P}\left(\widehat{m}_{k-1}^{(N,\alpha^*)} + \sqrt{\frac{2\log k}{T_{k-1}^{(N,\alpha^*)}}} \le p_N\right) \\ &\le \mathbb{P}\left(\exists l \in \{1,\dots,k\} \;:\; \frac{1}{l}\sum_{j=1}^{l}(X_j^{(N,\alpha^*)} - p_N) + \sqrt{\frac{2\log k}{l}} \le p_N\right) \\ &\le \sum_{l=1}^{k} \mathbb{P}\left(\frac{1}{l}\sum_{j=1}^{l}(X_j^{(N,\alpha^*)} - p_N) + \sqrt{\frac{2\log k}{l}} \le p_N\right)\end{aligned}$$

$$\leq \sum_{l=1}^{k} \mathbb{P}\left(\frac{1}{l}\sum_{j=1}^{l}(1 - X_j^{(N,\alpha^*)} - (1-p_N)) \geq \sqrt{\frac{2\log k}{l}}\right)$$
$$\leq \sum_{l=1}^{k} \mathrm{e}^{-4\log k} = \sum_{l=1}^{k} \frac{1}{k^4} = \frac{1}{k^3}.$$

The argument is similar for

$$\mathbb{P}\left(\widehat{m}_{k-1}^{(i,\alpha^*)} > p_i + \sqrt{\frac{2\log k}{T_{k-1}^{(i,\alpha^*)}}}\right) \leq \frac{1}{k^3}, \quad i = 1, \ldots, N, \quad k > N.$$

d) We have

$$\begin{aligned}
\mathbb{E}[T_n^{(i)}] &\leq \widehat{n}_i + \sum_{k=1}^{n} \frac{2}{k^3} \\
&= \frac{8\log n}{(p_N - p_i)^2} + \sum_{k=1}^{n} \frac{2}{k^3} \\
&\leq \frac{8\log n}{(p_N - p_i)^2} + \int_1^n \frac{2}{t^3} dt \\
&\leq \frac{8\log n}{(p_N - p_i)^2} + \left(1 - \frac{1}{n^2}\right),
\end{aligned}$$

hence

$$\begin{aligned}
\overline{\mathcal{R}}_n^{\alpha^*} &= np_N - \mathbb{E}\left[\sum_{k=1}^{n} p_{\alpha_k^*}\right] \\
&= \sum_{k=1}^{n} \mathbb{E}[p_N - p_{\alpha_k^*}] \\
&= np_N - \sum_{i=1}^{N} p_i \, \mathbb{E}[T_n^{i,\alpha^*}] \\
&= \sum_{i=1}^{N} (p_N - p_i) \, \mathbb{E}[T_n^{i,\alpha^*}] \\
&\leq 8 \sum_{i=1}^{N-1} \frac{\log n}{p_N - p_i} + \sum_{i=1}^{N-1} (p_N - p_i).
\end{aligned}$$

Problem 4.3

a) i) By first step analysis, the probability generating function

$$G_i(s) := \mathbb{E}\left[s^{T_{0,L}} \,\middle|\, S_0 = i\right], \qquad s \in [-1, 1],$$

of $T_{0,L}$ satisfies the equation

$$G_i(s) = psG_{i+1}(s) + qsG_{i-1}(s), \qquad i = 1, \ldots, L-1,$$

with the boundary conditions $G_0(s) = G_L(s) = 1$. This equation can be solved as

$$G_i(s) = C_+(s) \left(\frac{1+\sqrt{1-4pqs^2}}{2ps} \right)^i + C_-(s) \left(\frac{1-\sqrt{1-4pqs^2}}{2ps} \right)^i,$$

$i = 0, \ldots, L$, where

$$\begin{cases} C_+(s) := \dfrac{(2ps)^L - (1-\sqrt{1-4pqs^2})^L}{(1+\sqrt{1-4pqs^2})^L - (1-\sqrt{1-4pqs^2})^L} \\ \\ C_-(s) := \dfrac{(1-\sqrt{1+4pqs^2})^L - (2ps)^L}{(1+\sqrt{1-4pqs^2})^L - (1-\sqrt{1-4pqs^2})^L}. \end{cases}$$

ii) The Laplace transform

$$L_i(\lambda) := \mathbb{E}\left[\mathrm{e}^{-\lambda T_{0,L}} \mid S_0 = i \right], \qquad i = 0, 1, \ldots, L, \quad \lambda \geq 0.$$

of $T_{0,L}$ is then evaluated as

$$\begin{aligned} L_i(\lambda) &= G_i(\mathrm{e}^{-\lambda}) \\ &= C_+(\mathrm{e}^{-\lambda}) \left(\frac{1+\sqrt{1-4pq\mathrm{e}^{-2\lambda}}}{2p\mathrm{e}^{-\lambda}} \right)^i + C_-(\mathrm{e}^{-\lambda}) \left(\frac{1-\sqrt{1-4pq\mathrm{e}^{-2\lambda}}}{2p\mathrm{e}^{-\lambda}} \right)^i, \end{aligned}$$

$i = 0, \ldots, L$.

b) i) When $\mu = 0$, taking the limit as ε tends to zero yields the Laplace transform

$$L_x(\lambda) := \frac{\sinh(x\sqrt{2\lambda}) + \sinh((y-x)\sqrt{2\lambda})}{\sinh(y\sqrt{2\lambda})},$$

$x \in [0, y]$, $\lambda \geq 0$, of the first hitting time of the boundary $\{0, y\}$ by a standard Brownian motion $(B_t)_{t \in \mathbb{R}_+}$ started at $x \in [0, y]$, which recovers Equation (3) in Antal and Redner (2005), see also Equation (2.2.10) in Redner (2001), Theorem 1 in Williams (1992), and Relation (2.12) in Borodin (2017).

ii) When $\mu \neq 0$, we find the Laplace transform

$$L_x(\lambda) = C_1(\lambda)\mathrm{e}^{\mu+\sqrt{2\lambda+\mu^2}} + C_2(\lambda)\mathrm{e}^{\mu-\sqrt{2\lambda+\mu^2}}$$

$$= \frac{\mathrm{e}^{(x-y)\mu}\sinh(x\sqrt{2\lambda+\mu^2}) + \mathrm{e}^{\mu x}\sinh((y-x)\sqrt{2\lambda+\mu^2})}{\sinh(y\sqrt{2\lambda+\mu^2})},$$

$x \in [0, y]$, $\lambda \geq 0$, of the first hitting time of the boundary $\{0, y\}$ by a Brownian motion $(B_t + \mu t)_{t\in\mathbb{R}_+}$ with drift $\mu \in \mathbb{R}$ and started at $x \in [0, y]$, which recovers Equation (3), where

$$\begin{cases} C_1(s) := \dfrac{1 - \mathrm{e}^{(\mu-\sqrt{2\lambda+\mu^2})y}}{\mathrm{e}^{(\mu+\sqrt{2\lambda+\mu^2})y} - \mathrm{e}^{(\mu-\sqrt{2\lambda+\mu^2})y}} \\ \\ C_2(s) := \dfrac{\mathrm{e}^{(\mu+\sqrt{2\lambda+\mu^2})y}}{\mathrm{e}^{(\mu+\sqrt{2\lambda+\mu^2})y} - \mathrm{e}^{(\mu-\sqrt{2\lambda+\mu^2})y}}, \end{cases}$$

see Theorem 1 in Williams (1992) in the case $x = 0$, by taking $\alpha = 0$ and $C = -1$ therein.

c) i) By first step analysis, the probability generating function

$$G_i(s) := \mathbb{E}\left[s^{T_{0,L}} \,\middle|\, S_0 = i\right], \qquad s \in [-1, 1],$$

of $T_{0,L}$ satisfies the same equation

$$G_i(s) = psG_{i+1}(s) + qsG_{i-1}(s), \qquad i = 1, \ldots, L-1,$$

as above. However, the boundary conditions are modified into $G_0(s) = psG_1(s) + qsG_0(s)$, with $G_L(s) = 1$. The finite difference equation can now be solved as

$$G_i(s) = C_+(s)\left(\frac{1+\sqrt{1-4pqs^2}}{2ps}\right)^i + C_-(s)\left(\frac{1-\sqrt{1-4pqs^2}}{2ps}\right)^i,$$

$i = 0, \ldots, L$, where

$$\begin{cases} C_+(s) := \dfrac{ps\alpha_-(s) + qs - 1}{(1-qs)(\alpha_-^L(s) - \alpha_+^L(s)) - ps(\alpha_+(s)\alpha_-(s)^L - \alpha_+^L(s)\alpha_-(s))} \\ \\ C_-(s) := \dfrac{ps\alpha_+(s) + qs - 1}{(qs-1)(\alpha_-^L(s) - \alpha_+^L(s)) + ps(\alpha_+(s)\alpha_-(s)^L - \alpha_+^L(s)\alpha_-(s))} \end{cases}$$

and

$$\alpha_+(s) = \frac{1+\sqrt{1-4pqs^2}}{2ps}, \quad \alpha_-(s) = \frac{1-\sqrt{1-4pqs^2}}{2ps}.$$

ii) The Laplace transform is then evaluated as

$$L_i(\lambda) = G_i(\mathrm{e}^{-\lambda})$$
$$= C_+(\mathrm{e}^{-\lambda})\left(\frac{1+\sqrt{1-4pq\mathrm{e}^{-2\lambda}}}{2p\mathrm{e}^{-\lambda}}\right)^i + C_-(\mathrm{e}^{-\lambda})\left(\frac{1-\sqrt{1-4pq\mathrm{e}^{-2\lambda}}}{2p\mathrm{e}^{-\lambda}}\right)^i,$$

$i = 0, \ldots, L$.

i) When $\mu = 0$, taking the limit as ε tends to zero yields the Laplace transform

$$L_x(\lambda) := \frac{\cosh(x\sqrt{2\lambda})}{\cosh(y\sqrt{\lambda})}, \qquad x \in [0, y], \quad \lambda \geq 0,$$

of the first hitting time of the boundary $\{y\}$ by a standard Brownian motion reflected at 0, which recovers Equation (5) in Antal and Redner (2005), see also Equation (2.2.21) in Redner (2001).*

ii) When $\mu \neq 0$ we find the Laplace transform

$$L_x(\lambda) := \mathrm{e}^{(x-y)\mu}\frac{\mu\sinh(x\sqrt{2\lambda+\mu^2}) - \sqrt{2\lambda+\mu^2}\cosh(x\sqrt{2\lambda+\mu^2})}{\mu\sinh(y\sqrt{2\lambda+\mu^2}) - \sqrt{2\lambda+\mu^2}\cosh(y\sqrt{2\lambda+\mu^2})},$$

$x \in [0, y]$, $\lambda \geq 0$, of the first hitting time of the boundary $\{y\}$ by a Brownian motion $(B_t + \mu t)_{t\in\mathbb{R}_+}$ with drift $\mu \in \mathbb{R}$ reflected at 0 and started at $x \in [0, y]$.

Problem 4.4

a) By first step analysis, we have

$$H_i(s) = psH_{i+1}(s) + qsH_{i-1}(s), \qquad -1 \leq s \leq 1, \quad i \leq -2, \ i \geq 2,$$

and

$$H_1(s) = psH_2(s) + qs(1 + H_0(s)), \quad H_{-1}(s) = psH_{-2}(s) + qs(1 + H_0(s)),$$

and

$$H_0(s) = psH_1(s) + qsH_{-1}(s), \qquad -1 \leq s \leq 1.$$

b) Letting

* Equation (2.2.21) in Redner (2001) is stated for a reflecting boundary at $x = L$ ("Reflection mode" page 48), however in Antal and Redner (2005) the reflecting boundary is at $x = 0$, and therefore (5) therein has to be corrected accordingly.

$$
H_i(s) := \begin{cases} \dfrac{1}{\sqrt{1-4pqs^2}} \left(\dfrac{1-\sqrt{1-4pqs^2}}{2ps} \right)^i, & i \geq 1, \\[2ex] \dfrac{1-\sqrt{1-4pqs^2}}{\sqrt{1-4pqs^2}}, & i = 0, \\[2ex] \dfrac{1}{\sqrt{1-4pqs^2}} \left(\dfrac{1-\sqrt{1-4pqs^2}}{2qs} \right)^{-i}, & i \leq -1, \end{cases}
$$

we check that

$$
\begin{aligned}
& psH_{i+1}(s) + qsH_{i-1}(s) \\
&= \frac{ps}{\sqrt{1-4pqs^2}} \left(\frac{1-\sqrt{1-4pqs^2}}{2ps} \right)^{i+1} + \frac{qs}{\sqrt{1-4pqs^2}} \left(\frac{1-\sqrt{1-4pqs^2}}{2ps} \right)^{i-1} \\
&= \frac{1}{\sqrt{1-4pqs^2}} \left(\frac{1-\sqrt{1-4pqs^2}}{2ps} \right)^{i} \left(\frac{1-\sqrt{1-4pqs^2}}{2} + \frac{2pqs^2}{1-\sqrt{1-4pqs^2}} \right) \\
&= \frac{1}{\sqrt{1-4pqs^2}} \left(\frac{1-\sqrt{1-4pqs^2}}{2ps} \right)^{i}, \qquad i \geq 1.
\end{aligned}
$$

c) We have

$$
H_i(s) = (1 + H_0(s))G_i(s), \qquad i \in \mathbb{Z}, \quad -1 \leq s \leq 1.
$$

d) As a direct consequence of the answers to Questions (b) and (c), we have

$$
G_i(s) := \begin{cases} \left(\dfrac{1-\sqrt{1-4pqs^2}}{2ps} \right)^i, & i \geq 1, \\[2ex] 1-\sqrt{1-4pqs^2}, & i = 0, \\[2ex] \left(\dfrac{1-\sqrt{1-4pqs^2}}{2qs} \right)^{-i}, & i \leq -1. \end{cases}
$$

e) We find

$$
\mathbb{P}(T_0 < \infty \mid S_0 = i) = G_i(1) = \begin{cases} \min\left(1, \left(\dfrac{q}{p}\right)^i\right), & i \neq 0, \\[2ex] 1 - |p-q|, & i = 0, \end{cases}
$$

see (4.6) and (4.11).

f) Using the relations $\mathbb{E}[T_0^r \mid S_0 = i] = G_i'(1)$ when $\mathbb{P}(T_0^r \mid S_0 = i) = 1$, see (A.14), and $\mathbb{E}[T_0^r \mid S_0 = i] = +\infty$ when $\mathbb{P}(T_0^r \mid S_0 = i) < 1$, We find

$$\mathbb{E}[T_0^r \mid S_0 = i] = \begin{cases} \dfrac{i}{q-p}, & i \geq 1, \quad q > p, \\ +\infty, & i \geq 1, \quad q \leq p, \\ +\infty, & i = 0, \\ \dfrac{i}{q-p}, & i \leq -1, \quad p > q \\ +\infty, & i \leq -1, \quad p \leq q, \end{cases}$$

see (4.8).

Problem 4.5

a) We have

$$\mathbb{P}(S_{2n} = 2k) = \binom{2n}{n+k} p^{n+k} q^{n-k}, \quad -n \leq k \leq n.$$

b) We partition the event $\{S_{2n} = 0\}$ into

$$\{S_{2n} = 0\} = \bigcup_{k=1}^{2n} \{S_1 \neq 0, \ldots, S_{2k-1} \neq 0, \ S_{2k} = 0\}, \qquad n \geq 1,$$

according to all possible times $2k = 2, 4, \ldots, 2n$ of *first* return to state 0 *before* time $2n$, see Figure S.4.

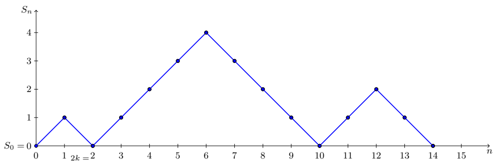


Fig. S.4: Last return to state 0 at time $k = 10$.

Then we have

$$\begin{aligned}
\mathbb{P}(S_{2n}=0) &= \sum_{r=1}^{n} \mathbb{P}\big(S_2 \neq 0, \ldots, S_{2r-1} \neq 0,\ S_{2r}=0,\ S_{2n}=0\big) \\
&= \sum_{r=1}^{n} \mathbb{P}\big(S_{2n}=0 \mid S_{2r}=0,\ S_{2r-1}\neq 0, \ldots,\ S_2 \neq 0\big) \\
&\qquad \times \mathbb{P}\big(S_2 \neq 0, \ldots, S_{2r-1}\neq 0,\ S_{2r}=0\big) \\
&= \sum_{r=1}^{n} \mathbb{P}\big(S_{2n}=0 \mid S_{2r}=0\big)\mathbb{P}\big(T_0=2r\big) \\
&= \sum_{k=1}^{n} \mathbb{P}\big(S_{2n-2r}=0\big)\mathbb{P}\big(T_0=2r\big), \qquad n \geq 1.
\end{aligned}$$

c) The idea of the proof is to note that after starting from $S_0=0$, one may move up with probability $1/2$, in which case $T_0=2r$ time steps strictly above 0 will be counted from time 0 until time T_0, after which the remaining $2r-2k$ time steps will be counted from time T_0 until time $2n$. On the other hand, if one moves down with probability $1/2$, zero time step strictly above 0 will be counted from time 0 until time $T_0=2r$, after which the remaining $2k$ time steps strictly above zero will be counted from time $T_0=2r$ until time $2n$. Hence we have

$$\begin{aligned}
\mathbb{P}(T_{2n}^{+}=2k) &= \sum_{r=1}^{n} \mathbb{P}\big(S_0=0, T_0=2r, T_{2n}^{+}=2k\big) \\
&= \sum_{r=1}^{n} \mathbb{P}\big(S_0=0, S_1=1, T_0=2r, T_{2n}^{+}=2k\big) \\
&\quad + \sum_{r=1}^{n} \mathbb{P}\big(S_0=0, S_1=-1, T_0=2r, T_{2n}^{+}=2k\big) \\
&= \sum_{r=1}^{k} \mathbb{P}\big(S_0=0, S_1=1, T_0=2r\big)\mathbb{P}\big(T_{2n}^{+}=2k \mid S_1=1, T_0=2r\big) \\
&\quad + \sum_{r=1}^{n-k} \mathbb{P}\big(S_0=0, S_1=-1, T_0=2r\big)\mathbb{P}\big(T_{2n}^{+}=2k \mid S_1=-1, T_0=2r\big) \\
&= \sum_{r=1}^{k} \mathbb{P}\big(S_0=0, S_1=1, T_0=2r\big)\mathbb{P}\big(T_{2n-2r}^{+}=2k-2r\big) \\
&\quad + \sum_{r=1}^{n-k} \mathbb{P}\big(S_0=0, S_1=-1, T_0=2r\big)\mathbb{P}\big(T_{2n-2r}^{+}=2k\big) \\
&= \frac{1}{2}\sum_{r=1}^{k} \mathbb{P}\big(T_0=2r\big)\mathbb{P}\big(T_{2n-2r}^{+}=2k-2r\big) + \frac{1}{2}\sum_{r=1}^{n-k} \mathbb{P}\big(T_0=2r\big)\mathbb{P}\big(T_{2n-2r}^{+}=2k\big),
\end{aligned}$$

$n \geq 1$.

d) We check that, when

$$\mathbb{P}\big(T^+_{2n-2r} = 2k-2r\big) = 2^{-(2n-2r)}\binom{2k-2r}{k-r}\binom{2n-2k}{n-k}$$

and

$$\mathbb{P}\big(T^+_{2n-2r} = 2k\big) = 2^{-(2n-2r)}\binom{2k}{k}\binom{2n-2r-2k}{n-r-k},$$

we have

$$\begin{aligned}
&\frac{1}{2}\sum_{r=1}^{k}\mathbb{P}\big(T_0 = 2r\big)\mathbb{P}\big(T^+_{2n-2r} = 2k-2r\big) + \frac{1}{2}\sum_{r=1}^{n-k}\mathbb{P}\big(T_0 = 2r\big)\mathbb{P}\big(T^+_{2n-2r} = 2k\big) \\
&\quad = \frac{1}{2}\sum_{r=1}^{k}\mathbb{P}\big(T_0 = 2r\big)2^{-2n+2r}\binom{2k-2r}{k-r}\binom{2n-2k}{n-k} \\
&\qquad + \frac{1}{2}\sum_{r=1}^{n-k}2^{-2n+2r}\mathbb{P}\big(T_0 = 2r\big)\binom{2k}{k}\binom{2n-2r-2k}{n-r-k} \\
&\quad = \frac{1}{2}2^{-2n}\binom{2n-2k}{n-k}2^{2k}\sum_{r=1}^{k}\mathbb{P}\big(T_0 = 2r\big)\frac{1}{2^{2(k-r)}}\binom{2k-2r}{k-r} \\
&\qquad + \frac{1}{2}2^{-2n}\binom{2k}{k}2^{2(n-k)}\sum_{r=1}^{n-k}\mathbb{P}\big(T_0 = 2r\big)\frac{1}{2^{2(n-k-r)}}\binom{2n-2r-2k}{n-r-k} \\
&\quad = \frac{1}{2}2^{-2(n-k)}\binom{2n-2k}{n-k}\sum_{r=1}^{k}\mathbb{P}\big(T_0 = 2r\big)\mathbb{P}(S_{2k-2r} = 0) \\
&\qquad + \frac{1}{2}\binom{2k}{k}2^{-2k}\sum_{r=1}^{n-k}\mathbb{P}\big(T_0 = 2r\big)\mathbb{P}(S_{2n-2k+2r} = 0) \\
&\quad = \frac{1}{2}2^{-2(n-k)}\binom{2n-2k}{n-k}\mathbb{P}(S_{2k} = 0) + \frac{1}{2}2^{-2k}\binom{2k}{k}\mathbb{P}(S_{2n-2k} = 0) \\
&\quad = \frac{1}{2}2^{-2n}\binom{2n-2k}{n-k}\binom{2k}{k} + \frac{1}{2}2^{-2n}\binom{2k}{k}\binom{2n-2k}{n-k} \\
&\quad = 2^{-2n}\binom{2k}{k}\binom{2n-2k}{n-k} \\
&\quad = \mathbb{P}\big(T^+_{2n} = 2k\big), \qquad n \geq 1.
\end{aligned}$$

e) We have

$$\mathbb{P}\big(T^+_{2n} = 2k\big) = 2^{-2n}\binom{2k}{k}\binom{2n-2k}{n-k}$$

$$
\begin{aligned}
&= 2^{-2n} \frac{(2k)!}{k!^2} \frac{(2n-2k)!}{(n-k)!^2} \\
&\simeq 2^{-2n} \frac{(2k/e)^{2k}\sqrt{4\pi k}}{(k/e)^{2k} 2\pi k} \frac{((2n-2k)/e)^{(2n-2k)}\sqrt{2\pi(2n-2k)}}{((n-k)/e)^{(2n-2k)} 2\pi(n-k)} \\
&= \frac{1}{\pi\sqrt{k(n-k)}}, \qquad k, n-k \to \infty.
\end{aligned}
$$

Next, we compute the limit

$$
\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\big(T_{2n}^+/2n \le x\big) &= \lim_{n\to\infty} \sum_{k=0}^{nx} \mathbb{P}\big(T_{2n}^+/2n = k/n\big) \\
&= \lim_{n\to\infty} \sum_{0\le k/n\le x} 2^{-2n} \binom{2k}{k}\binom{2n-2k}{n-k} \\
&\simeq \frac{1}{\pi} \lim_{n\to\infty} \frac{1}{n} \sum_{0\le k/n\le x} \frac{1}{\sqrt{k(1-k/n)/n}} \\
&= \frac{1}{\pi}\int_0^x \frac{1}{\sqrt{t(1-t)}} dt \\
&= \frac{1}{2} + \frac{\arcsin(2x-1)}{\pi} \\
&= \frac{2}{\pi}\arcsin\sqrt{x}, \qquad x \in [0,1],
\end{aligned}
$$

which yields the arcsine distribution.

Problem 4.6

a) We have

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(\alpha\sum_{l=1}^{n} f(X_l)\right)\right] &= \prod_{l=1}^{n} \mathbb{E}\left[e^{\alpha f(l)}\right] \\
&= \left(\mathbb{E}\left[e^{\alpha f(l)}\right]\right)^n \\
&= (\lambda_0(\alpha))^n, \qquad n \ge 1.
\end{aligned}
$$

b) For any $\alpha \in \mathbb{R}$ and $\gamma > 0$, we have

$$
\begin{aligned}
e^{\alpha\gamma n}\mathbb{P}\left(\sum_{l=1}^{n} f(X_l) \ge n\gamma\right) &= e^{\alpha\gamma n}\,\mathbb{E}\left[\mathbf{1}_{\left\{\sum_{l=1}^{n} f(X_l)\ge n\gamma\right\}}\right] \\
&\le \mathbb{E}\left[\exp\left(\alpha\sum_{l=1}^{n} f(X_l)\right)\right] \\
&= e^{-\alpha\gamma n}(\lambda_0(\alpha))^n
\end{aligned}
$$

$$= e^{-n(\alpha\gamma - \log \lambda_0(\alpha))}, \qquad n \geq 1,$$

hence

$$\mathbb{P}\left(\sum_{l=1}^{n} f(X_l) \geq n\gamma\right) = e^{-n(\alpha\gamma - \log \lambda_0(\alpha))}, \qquad n \geq 1. \qquad \text{(S.4.4)}$$

c) Since

$$\sum_{l=1}^{d} \pi_l f(l) = \mathbb{E}[f(X_1)] = 0,$$

we have

$$\begin{aligned}
\lambda_0(\alpha) &= \sum_{l=1}^{d} \pi_l e^{\alpha f(l)} \\
&= \sum_{l=1}^{d} \pi_l + \alpha \sum_{l=1}^{d} \pi_l f(l) + \sum_{l=1}^{d} \pi_l (e^{\alpha f(l)} - \alpha f(l) - 1) \\
&= 1 + \sum_{l=1}^{d} \pi_l (e^{\alpha f(l)} - \alpha f(l) - 1), \qquad \alpha \geq 1.
\end{aligned}$$

d) We have

$$\begin{aligned}
\lambda_0(\alpha) &= 1 + \sum_{l=1}^{d} \pi_l (e^{\alpha f(l)} - \alpha f(l) - 1), \\
&= 1 + \sum_{k=2}^{\infty} \sum_{l=1}^{d} \pi_l \frac{(\alpha f(l))^n}{n!} \\
&\leq 1 + \sum_{k=2}^{\infty} \sum_{l=1}^{d} \pi_l \alpha^n \\
&= 1 + \sum_{k=2}^{\infty} \alpha^n \\
&= 1 + \frac{\alpha^2}{1-\alpha}, \qquad \alpha \in [0, 1).
\end{aligned}$$

e) By (S.4.4) and Question (d), for any $\alpha \in [0, 1)$ and $\gamma > 0$ we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{l=1}^{n} f(X_l) \geq \gamma\right) \leq e^{-n(\alpha\gamma - \frac{\alpha^2}{1-\alpha})}, \qquad n \geq 1.$$

f) The value of $\alpha \in [0, 1)$ which maximizes $\alpha\gamma - \alpha^2/(1-\alpha)$ satisfies

$$\gamma - 2\frac{\alpha}{1-\alpha} - \frac{\alpha^2}{(1-\alpha)^2} = 0$$

i.e.

$$\alpha = \frac{\gamma}{\gamma + 1 + \sqrt{\gamma+1}} < 1$$

and

$$1 - \alpha = \frac{1+\sqrt{\gamma+1}}{\gamma+1+\sqrt{\gamma+1}} = \frac{1}{\sqrt{\gamma+1}}.$$

g) We have

$$\begin{aligned}
\alpha\gamma - \frac{\alpha^2}{1-\alpha} &= \frac{\gamma^2}{\gamma+1+\sqrt{\gamma+1}} - \frac{\gamma^2}{\gamma+1+\sqrt{\gamma+1}(1+\sqrt{\gamma+1})} \\
&= \frac{\gamma^2\sqrt{\gamma+1}}{(\gamma+1+\sqrt{\gamma+1})(1+\sqrt{\gamma+1})} \\
&= \frac{\gamma^2}{(1+\sqrt{\gamma+1})^2} \\
&\geq \frac{\gamma^2}{(1+\sqrt{2})^2} \\
&\geq \frac{\gamma^2}{6},
\end{aligned}$$

hence for all $\gamma \in [0,1)$ and $n \geq 0$ we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{l=1}^{n} f(X_l) \geq \gamma\right) \leq e^{-n\gamma^2/6}.$$

We note that this bound is better than the upper bound $e^{-(1-\lambda_1)n\gamma^2/12}$ where λ_1 is the second largest eigenvalue of P, since $0 \leq 1-\lambda_1 \leq 2$.

Problem 4.7

a) For all $i = 1, \ldots, d$, we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{j=1}^{d}\left|\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{\{X_k=j\}} - \pi_j\right|\right] &= \frac{1}{n}\sum_{j=1}^{d}\mathbb{E}\left[\left|\sum_{k=1}^{n}(\mathbf{1}_{\{X_k=j\}} - \pi_j)\right|\right] \\
&\leq \frac{1}{n}\sum_{j=1}^{d}\sqrt{\mathbb{E}\left[\left|\sum_{k=1}^{n}(\mathbf{1}_{\{X_k=j\}} - \pi_j)\right|^2\right]} \\
&= \frac{1}{n}\sum_{j=1}^{d}\sqrt{\mathbb{E}\left[\sum_{k=1}^{n}\left|\mathbf{1}_{\{X_k=j\}} - \pi_j\right|^2\right]}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{n}} \sum_{j=1}^{d} \sqrt{\mathbb{E}[\left|\mathbf{1}_{\{X_k=j\}} - \pi_j\right|^2]} \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^{d} \sqrt{\pi_j(1-\pi_j)} \\
&\leq \frac{1}{\sqrt{n}} \sum_{j=1}^{d} \sqrt{\pi_j} \\
&\leq \frac{\sqrt{d}}{\sqrt{n}} \sqrt{\sum_{j=1}^{d} \pi_j} \\
&= \sqrt{\frac{d}{n}}.
\end{aligned}
$$

b) We have

$$
\begin{aligned}
&\sup_{y\in\mathbb{R}} |f(x_1,\ldots,x_n) - f(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)| \\
&\quad = \sup_{y\in\mathbb{R}} \sum_{j=1}^{d} \left| \frac{1}{n}(\mathbf{1}_{\{x_i=j\}} - \mathbf{1}_{\{y=j\}}) \right| \\
&\quad \leq \sup_{y\in\mathbb{R}} \sum_{j=1}^{d} \frac{1}{n} \left| \mathbf{1}_{\{x_i=j\}} + \mathbf{1}_{\{y=j\}} \right| \\
&\quad \leq \frac{2}{n},
\end{aligned}
$$

$x_1,\ldots,x_n \in \mathbb{R}$, $i = 1,\ldots,n$.

c) For all $i = 1,\ldots,d$ we have

$$
\begin{aligned}
&\mathbb{P}\left(\sum_{j=1}^{d} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k=j\}} - \pi_j \right| > \varepsilon \right) \\
&= \mathbb{P}\left(\sum_{j=1}^{d} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k=j\}} - \pi_j \right| - \mathbb{E}\left[\sum_{j=1}^{d} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k=j\}} - \pi_j \right| \right] \right. \\
&\qquad \left. > \varepsilon - \mathbb{E}\left[\sum_{j=1}^{d} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k=j\}} - \pi_j \right| \right] \right) \\
&\leq \mathbb{P}\left(\sum_{j=1}^{d} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k=j\}} - \pi_j \right| - \mathbb{E}\left[\sum_{j=1}^{d} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{X_k=j\}} - \pi_j \right| \right] > \varepsilon - \sqrt{\frac{d}{n}} \right)
\end{aligned}
$$

$$\leq \exp\left(-\frac{n}{2}\left(\varepsilon - \sqrt{\frac{d}{n}}\right)^2\right),$$

provided that $\varepsilon - \sqrt{d/n} > 0$, which implies

$$\mathbb{P}\left(\sum_{j=1}^{d}\left|\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}_{\{X_k=j\}} - \pi_j\right| > \varepsilon\right) \leq \exp\left(-\frac{n}{2}\max\left(0, \varepsilon - \sqrt{\frac{d}{n}}\right)^2\right).$$

d) When $n \geq 4d/\varepsilon^2$, *i.e.* $\varepsilon \geq 2\sqrt{d/n}$, we have

$$\begin{aligned}\mathbb{P}\left(\sum_{j=1}^{d}\left|\widetilde{\pi}_j(n) - \pi_j\right| > \varepsilon\right) &\leq \exp\left(-\frac{n}{2}\max\left(0, \varepsilon - \sqrt{\frac{d}{n}}\right)^2\right) \\ &= e^{-n\varepsilon^2/8}.\end{aligned}$$

e) Setting $n > -8(\log \delta)/\varepsilon^2$, we have

$$\mathbb{P}\left(\sum_{j=1}^{d}\left|\widetilde{\pi}_j(n) - \pi_j\right| > \varepsilon\right) \leq e^{-n\varepsilon^2/8} < \delta,$$

which allows us to conclude by taking $c = 8$.

Chapter 5 - Cookie-Excited Random Walks

Exercise 5.1

a) The number of cookies present in the considered region is kL.
b) The number of time steps is kL.
c) Let N denote the average number of time steps needed. From the relation $N(\widetilde{p} - \widetilde{q}) = L$ we deduce $N = L/(\widetilde{p} - \widetilde{q})$.
d) The condition is $kL \leq N = L/(\widetilde{p} - \widetilde{q})$, or $k \leq 1/(\widetilde{p} - \widetilde{q})$, which yields

$$\frac{1}{2} < \widetilde{p} \leq \frac{1}{2}\left(1 + \frac{1}{k}\right).$$

e) Under the condition

$$\widetilde{p} > \frac{1}{2}\left(1 + \frac{1}{k}\right)$$

the amount of cookies consumed will remain strictly lower than the number of available cookies, thus ensuring the transience of the random walk.

Exercise 5.2

a) The probability $\mathbb{P}(X=0)$ that the random walk eats no cookies before hitting the origin is the probability of going directly from ⓪ to ⓪ in one time step, which is $1/2$.

The probability $\mathbb{P}(X=1)$ that the random walk eats exactly *one* cookie before hitting the origin is the probability of first moving from ⓪ to ① in one time step and then back to ⓪ in one time step, that is $q \times (1/2) = q/2$.

In general, we have

$$\begin{aligned}\mathbb{P}(X=x) &= \mathbb{P}(\tau_x < \tau_0 \mid S_0 = 0) - \mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = 0)\\ &= \frac{1}{2}\prod_{l=2}^{x}\left(1-\frac{2q}{l}\right) - \frac{1}{2}\prod_{l=2}^{x+1}\left(1-\frac{2q}{l}\right),\\ &= \frac{1}{2}\left(1-\left(1-\frac{2q}{x+1}\right)\right)\prod_{l=2}^{x}\left(1-\frac{2q}{l}\right)\\ &= \frac{q}{x+1}\prod_{l=2}^{x}\left(1-\frac{2q}{l}\right).\end{aligned}$$

b) We have

$$\mathbb{E}[X] = \sum_{x\geq 0} x\mathbb{P}(X=x) = q\sum_{x\geq 0}\frac{x}{x+1}\prod_{l=2}^{x}\left(1-\frac{2q}{l}\right),$$

hence

$$qc_q\sum_{x\geq 0}\frac{x}{(x+1)x^{2q}} \leq \mathbb{E}[X] \leq qC_q\sum_{x\geq 0}\frac{x}{(x+1)x^{2q}},$$

and $\mathbb{E}[X]$ is finite if and only if $2q > 1$.

Remark. One could show in addition that the mean return time to ⓪ is always infinite, see Antal and Redner (2005).

Chapter 6 - Convergence to Equilibrium

Exercise 6.1 The limiting distribution of the chain $(Y_k)_{k\geq 0}$ is $(0,0,0,0,0,1)$ independently of the initial state because the states $\{0,1,2,3,4\}$ are transient and state ⑤ is absorbing. This means that

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

which would be difficult to recover by a direct computation of P^n. The equation $\pi = \pi P$ which determines the stationary distribution π reads

$$\begin{cases} \pi_0 = q\pi_0 + q\pi_1 + q\pi_2 \\ \pi_1 = p\pi_0 + p\pi_4 \\ \pi_2 = p\pi_1 \\ \pi_3 = p\pi_2 + p\pi_3 \\ \pi_4 = p\pi_3 \\ \pi_5 = q\pi_4 + \pi_5, \end{cases}$$

i.e.

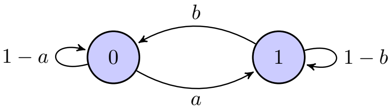
$$\begin{cases} p\pi_0 = q\pi_1 + q\pi_2 \\ \pi_1 = p\pi_0 + p\pi_4 \\ \pi_2 = p\pi_1 \\ q\pi_3 = p\pi_2 \\ \pi_4 = p\pi_3 \\ \pi_4 = 0, \end{cases}$$

hence $(\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5) = (0,0,0,0,0,1)$, which coincides with the limiting distribution. Note that the relation $\pi_i = 1/\mu_i(i)$ still holds for $i = 0,1,2,3,4,5$, although not all of the assumptions of Theorems 6.2, 6.6 and 6.6 (notably the irreducibility condition) are satisfied here.

Exercise 6.2 Writing the condition $\pi P = \pi$ leads to the equations

$$\begin{cases} \dfrac{\pi_0}{3} + 2\dfrac{\pi_1}{3} = \pi_0 \\[2ex] 2\dfrac{\pi_0}{3} + \dfrac{\pi_1}{3} = \pi_1 \end{cases}$$

i.e. $\pi_0 = \pi_1$. Combining this relation with the condition $\pi_0 + \pi_1 = 1$ shows that $\pi_0 = \pi_1 = 1/2$.



Using the general relation

$$[\pi_0, \pi_1] = \left[\frac{b}{a+b}, \frac{a}{a+b}\right],$$

with $(a,b) \neq (0,0)$ and $(a,b) \neq (1,1)$ for the two-state chain with transition matrix

$$P = \begin{bmatrix} 1-a & a \\ \\ b & 1-b \end{bmatrix}$$

yields the same answer

$$[\pi_0, \pi_1] = \left[\frac{1}{2}, \frac{1}{2}\right]$$

when $a = b$, in which case the matrix P is also *column-stochastic*, as illustrated in the following R code.

```
library("igraph"); library(markovchain)
P<-matrix(c(1/3,2/3,2/3,1/3),nrow=2,byrow=TRUE);MC
        <-new("markovchain",transitionMatrix=P)
graph <- as(MC, "igraph")
plot(graph,vertex.size=50,edge.label.cex=2,edge.label=E(graph)$prob,edge.color='black',
        vertex.color='dodgerblue',vertex.label.cex=3)
steadyStates(object = MC)
        1 2
[1,] 0.5 0.5
```

Exercise 6.3

a) The chain is reducible and its communicating classes are $\{0,1,2,3,4\}$ and $\{5\}$.
b) The limiting distribution is $(0,0,0,0,0,1)$ independently of the initial state because the states $\{0,1,2,3,4\}$ are transient (cf. Proposition 7.4 in Privault (2018)) and state 5 is absorbing. This means that

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

which would be difficult to recover by a direct computation of P^n.

For the stationary distribution, the equation $\pi = \pi P$ reads

$$\begin{cases} \pi_0 = q\pi_0 + q\pi_1 + q\pi_2 \\ \pi_1 = p\pi_0 + p\pi_4 \\ \pi_2 = p\pi_1 \\ \pi_3 = p\pi_2 + p\pi_3 \\ \pi_4 = p\pi_3 \\ \pi_5 = q\pi_4 + \pi_5, \end{cases}$$

i.e.

$$\begin{cases} p\pi_0 = q\pi_1 + q\pi_2 \\ \pi_1 = p\pi_0 + p\pi_4 \\ \pi_2 = p\pi_1 \\ q\pi_3 = p\pi_2 \\ \pi_4 = p\pi_3 \\ \pi_4 = 0, \end{cases}$$

hence $(\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5) = (0, 0, 0, 0, 0, 1)$, which coincides with the limiting distribution.

Note that the relation $\pi_i = 1/\mu_i(i)$ still holds for $i = 0, 1, 2, 3, 4, 5$, although not all of the assumptions of Theorems 6.2, 6.6 and 6.6 (notably the irreducibility condition) are satisfied here.

Exercise 6.4

a) We have

$$(\pi_0, \pi_1) = \left(\frac{b}{a+b}, \frac{a}{a+b} \right).$$

b) We have

$$\mu_0(0) = 1 + \frac{a}{b}, \quad \mu_1(1) = 1 + \frac{b}{a}, \quad h_0(1) = \frac{1}{b}, \quad h_1(0) = \frac{1}{a}.$$

c) We have

$$\begin{aligned} \mathbb{E}[\tau - 1 \mid X_0 = 0] &= a\mu_1(1) + (1-a)\mu_0(0) \\ &= a\left(1 + \frac{b}{a}\right) + (1-a)\left(1 + \frac{a}{b}\right) \\ &= (1 + b - a)\frac{a+b}{b} \\ &= \frac{1 + b - a}{\pi_0}, \end{aligned}$$

and

$$
\begin{aligned}
\mathbb{E}[\tau - 1 \mid X_0 = 1] &= (1-b)\mu_1(1) + b\mu_0(0) \\
&= (1-b)\left(1 + \frac{b}{a}\right) + b\left(1 + \frac{a}{b}\right) \\
&= (1 + a - b)\frac{a+b}{a} \\
&= \frac{1+a-b}{\pi_1}.
\end{aligned}
$$

d) We have

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=1\}} \,\middle|\, X_0 = 1\right] = b(\mu_0(0) - 1) + (1-b) = 1 + a - b, \\
&\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=1\}} \,\middle|\, X_0 = 0\right] = a + (1-a)(\mu_0(0) - 1) = a + (1-a)\frac{a}{b} = (1 + b - a)\frac{\pi_1}{\pi_0}, \\
&\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=0\}} \,\middle|\, X_0 = 1\right] = b + (1-b)(\mu_1(1) - 1) = b + (1-b)\frac{b}{a} = (1 + a - b)\frac{\pi_0}{\pi_1}, \\
&\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=0\}} \,\middle|\, X_0 = 0\right] = a(\mu_1(1) - 1) + (1-a) = 1 + b - a.
\end{aligned}
$$

e) We note that

$$
\begin{cases}
\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=1\}} \,\middle|\, X_0 = 1\right] = \mathbb{E}[\tau - 1 \mid X_0 = 1]\pi_1, \\[2ex]
\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=1\}} \,\middle|\, X_0 = 0\right] = \mathbb{E}[\tau - 1 \mid X_0 = 0]\pi_1, \\[2ex]
\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=0\}} \,\middle|\, X_0 = 1\right] = \mathbb{E}[\tau - 1 \mid X_0 = 1]\pi_0, \\[2ex]
\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=0\}} \,\middle|\, X_0 = 0\right] = \mathbb{E}[\tau - 1 \mid X_0 = 0]\pi_0,
\end{cases}
$$

hence for any initial distribution $(\mathbb{P}(X_0 = 0), \mathbb{P}(X_0 = 1))$ we have

$$
\begin{aligned}
&\frac{\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=i\}}\right]}{\mathbb{E}[\tau - 1]} \\
&= \frac{\mathbb{P}(X_0 = 0)\,\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=i\}} \,\middle|\, X_0 = 0\right] + \mathbb{P}(X_0 = 1)\,\mathbb{E}\left[\sum_{l=1}^{\tau-1} \mathbb{1}_{\{X_l=i\}} \,\middle|\, X_0 = 1\right]}{\mathbb{E}[\tau - 1]}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\mathbb{P}(X_0 = 0)\,\mathbb{E}[\tau - 1 \mid X_0 = 0]\pi_i + \mathbb{P}(X_0 = 1)\,\mathbb{E}[\tau - 1 \mid X_0 = 0]\pi_i}{\mathbb{E}[\tau - 1]} \\
&= \pi_i \mathbb{P}(X_0 = 0) + \pi_i \mathbb{P}(X_0 = 1) \\
&= \pi_i, \qquad i = 0, 1.
\end{aligned}
$$

Exercise 6.5

a) This inequality follows from the definitions of $\widehat{d}(n)$ and $d(n)$, $n \geq 0$.
b) We have

$$
\begin{aligned}
d(n) &= \max_{\mu \in \mathcal{P}_N} \|\mu P^n - \pi\|_1 \\
&= \max_{\mu \in \mathcal{P}_N} \sum_{l=1}^{N} \big|[\mu P^n]_l - \pi_l\big| \\
&= \max_{\mu \in \mathcal{P}_N} \sum_{l=1}^{N} \left|\sum_{k=1}^{N} \mu_k [P^n]_{k,l} - \pi_l\right| \\
&= \max_{\mu \in \mathcal{P}_N} \sum_{l=1}^{N} \left|\sum_{k=1}^{N} \mu_k \big([P^n]_{k,l} - \pi_l\big)\right| \\
&\leq \max_{\mu \in \mathcal{P}_N} \sum_{l=1}^{N} \sum_{k=1}^{N} \big|\mu_k \big([P^n]_{k,l} - \pi_l\big)\big| \\
&= \max_{\mu \in \mathcal{P}_N} \sum_{l=1}^{N} \sum_{k=1}^{N} \mu_k \big|[P^n]_{k,l} - \pi_l\big| \\
&= \max_{\mu \in \mathcal{P}_N} \sum_{k=1}^{N} \mu_k \sum_{l=1}^{N} \big|[P^n]_{k,l} - \pi_l\big| \\
&= \max_{\mu \in \mathcal{P}_N} \sum_{k=1}^{N} \mu_k \big\|[P^n]_{k,\cdot} - \pi\big\|_1 \\
&\leq \max_{\mu \in \mathcal{P}_N} \sum_{k=1}^{N} \mu_k \max_{j=1,2,\ldots,N} \big\|[P^n]_{j,\cdot} - \pi\big\|_1 \\
&= \widehat{d}(n) \max_{\mu \in \mathcal{P}_N} \sum_{k=1}^{N} \mu_k \\
&= \widehat{d}(n).
\end{aligned}
$$

Alternatively, we can note that

$$
\mu \mapsto \|\mu P^n - \pi\|_1
$$

is a convex function on the polyhedron

$$\Delta_N := \{\mu \in [0,1]^N \ : \ \mu_1 + \cdots + \mu_N = 1\},$$

and therefore it reaches its maximum on an extremal vertex on Δ_N, *i.e.* there exists some $k_0 \in \{1, \ldots, N\}$ such that

$$\begin{aligned} d(n) &:= \max_{\mu \in \mathcal{P}_N} \|\mu P^n - \pi\|_1 \\ &= \big\| [P^n]_{k_0, \cdot} - \pi \big\|_1 \\ &\leq \max_{k=1,2,\ldots,N} \big\| [P^n]_{k, \cdot} - \pi \big\|_1 \\ &= \widehat{d}(n), \quad n \geq 0. \end{aligned}$$

Exercise 6.6

a) We have

$$\begin{aligned} \mathbb{P}(X_n \in A) &= \mathbb{P}(X_n \in A \text{ and } \tau \leq n) + \mathbb{P}(X_n \in A \text{ and } \tau > n) \\ &= \mathbb{P}(X_n \in A \mid \tau \leq n)\mathbb{P}(\tau \leq n) + \mathbb{P}(X_n \in A \mid \tau > n)\mathbb{P}(\tau > n) \\ &= \pi(A)\mathbb{P}(\tau \leq n) + \mathbb{P}(X_n \in A \mid \tau > n)\mathbb{P}(\tau > n) \\ &= \pi(A) + (\mathbb{P}(X_n \in A \mid \tau > n) - \pi(A))\mathbb{P}(\tau > n). \end{aligned}$$

b) We have

$$\begin{aligned} \mid \mathbb{P}(X_n \in A) - \pi(A) \mid &= \mid (\mathbb{P}(X_n \in A \mid \tau > n) - \pi(A)) \mid \mathbb{P}(\tau > n) \\ &\leq \mathbb{P}(\tau > n), \end{aligned}$$

since for any $a, b \in [0,1]$ we have $|a - b| \leq 1$ due to the inequalities

$$-1 \leq a - 1 \leq a - b \leq 1 - b \leq 1.$$

c) Such an example can be constructed as the hitting time τ of a domain inside $\mathbb{S}$, by freezing X_n as $X_n = X_{\min(\tau,n)}$ after time τ.

Exercise 6.7

a) Since M has positive entries and is column-stochastic, $P := M^\top$ is the transition probability matrix of an aperiodic irreducible Markov chain with finite state space $\mathbb{S} = \{1, 2, \ldots, n\}$. By Corollary 6.7, the chain admits a unique stationary distribution π such that $\pi = \pi P$, *i.e.* $\pi^\top = (\pi P)^\top = P^\top \pi^\top = M\pi^\top$, *i.e.* π^T is the only eigenvector of M with eigenvalue 1 under the normalization condition $\|\pi\|_1 = 1$.

b) The first statement follows as in Question (a) above from Corollary 6.7, by letting $\pi = q^\top$. The second statement also follows from Corollary 6.7, which states that

$$q = \pi^\top = \lim_{k\to\infty} (e_j P^k)^\top = \lim_{k\to\infty} (P^\top)^k e_j^\top = \lim_{k\to\infty} M^k e_j^\top = \lim_{k\to\infty} [M^k]_{\cdot,j}$$

for any $e_j = \mathbf{1}_{\{j\}}$, $j \in \mathbb{S}$. Therefore, decomposing x_0 as $x_0 = \sum_{j\in\mathbb{S}} x_0^j e_j^\top$, we have

$$q = q \sum_{j\in\mathbb{S}} x_0^j = \sum_{j\in\mathbb{S}} x_0^j \lim_{k\to\infty} M^k e_j^\top = \lim_{k\to\infty} M^k \sum_{j\in\mathbb{S}} x_0^j e_j^\top = \lim_{k\to\infty} M^k x_0.$$

Exercise 6.8

a) We have

$$\begin{aligned}
\lim_{n\to\infty} \frac{\mathbb{E}[R_n^i]}{n} &= \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[\sum_{j=1}^n \mathbb{1}_{\{X_j=i\}}\right] \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n \mathbb{E}[\mathbb{1}_{\{X_j=i\}}] \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n \mathbb{P}(X_j = i) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n \sum_{l\in\mathbb{S}} \mathbb{P}(X_j = i \mid X_0 = l)\mathbb{P}(X_0 = l) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n \sum_{l\in\mathbb{S}} [P^j]_{l,i}\mathbb{P}(X_0 = l) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \sum_{l\in\mathbb{S}} [P^{j+1}]_{l,i}\mathbb{P}(X_0 = l) \\
&= \sum_{k\in\mathbb{S}} P_{k,i} \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n \sum_{l\in\mathbb{S}} [P^j]_{l,k}\mathbb{P}(X_0 = l) \\
&= \sum_{k\in\mathbb{S}} P_{k,i} \lim_{n\to\infty} \frac{\mathbb{E}[R_n^k]}{n},
\end{aligned}$$

hence $\eta_i := \lim_{n\to\infty} \mathbb{E}[R_n^i]/n$, $i \in \mathbb{S}$, satisfies the equation $\eta = \eta P$ and we conclude by uniqueness of the stationary distribution $(\pi_i)_{i\in\mathbb{S}}$ as the solution to that equation.

b) Letting $\tau_x^{(0)} := 0$ and letting $\tau_x^{(k)}$ denote the time of the k-*th* visit to state x, the sequence $\big(\tau_x^{(k+1)} - \tau_x^{(k)}\big)_{k\geq 0}$, resp. $\big(R^y_{\tau_x^{(k+1)}} - R^y_{\tau_x^{(k)}}\big)_{k\geq 0}$, is made of independent random variables, $i \in \mathbb{S}$, hence by the law of large numbers for renewal processes, see Corollary 14 page 106 of Serfozo (2009), we have

$$\pi_y = \lim_{n\to\infty} \frac{\mathbb{E}[R_n^y]}{n} = \frac{\mathbb{E}\left[R^y_{\tau_x^{(1)}} \mid X_0 = x\right]}{\mathbb{E}\left[\tau_x^{(1)} \mid X_0 = x\right]} = \frac{\mathbb{E}[N_{x,y} \mid X_0 = x]}{\mathbb{E}[\tau_x \mid X_0 = x]}, \qquad x, y \in \mathbb{S}.$$

c) We have

$$\mathbb{P}(N_{x,y} = 0 \mid X_0 = x) = 1 - \mathbb{P}(N_{x,y} \geq 1 \mid X_0 = x) = 1 - \alpha_{x,y}$$

and

$$\begin{aligned}
&\mathbb{P}(N_{x,y} = k \mid X_0 = x) \\
&= \mathbb{P}(N_{x,y} \geq 1 \mid X_0 = x)\big(\mathbb{P}(N_{y,x} = 0 \mid X_0 = y)\big)^{k-1}\mathbb{P}(N_{y,x} \geq 1 \mid X_0 = y) \\
&= \alpha_{x,y}(1 - \alpha_{y,x})^{k-1}\alpha_{y,x}, \qquad k \geq 1,
\end{aligned}$$

and we check that

$$\begin{aligned}
\mathbb{P}(N_{x,y} \geq 0 \mid X_0 = x) &= \mathbb{P}(N_{x,y} = 0 \mid X_0 = x) + \mathbb{P}(N_{x,y} \geq 1 \mid X_0 = x) \\
&= 1 - \alpha_{x,y} + \sum_{k\geq 1} \mathbb{P}(N_{x,y} = k \mid X_0 = x) \\
&= 1 - \alpha_{x,y} + \alpha_{x,y}\alpha_{y,x} \sum_{k\geq 1} (1 - \alpha_{y,x})^{k-1} \\
&= 1, \qquad x, y \in \mathbb{S}.
\end{aligned}$$

d) We have

$$\begin{aligned}
\frac{\pi_y}{\pi_x} &= \pi_y\, \mathbb{E}[\tau_x \mid X_0 = x] \\
&= \mathbb{E}[N_{x,y} \mid X_0 = x] \\
&= \sum_{k=1}^{\infty} k\mathbb{P}(N_{x,y} = k \mid X_0 = x) \\
&= \alpha_{x,y}\alpha_{y,x} \sum_{k=1}^{\infty} k(1 - \alpha_{y,x})^{k-1} \\
&= \frac{\alpha_{x,y}\alpha_{y,x}}{\alpha_{y,x}^2} \\
&= \frac{\alpha_{x,y}}{\alpha_{y,x}}, \qquad x, y \in \mathbb{S}.
\end{aligned}$$

Problem 6.9

a) The computation of eigenvalues shows that the two eigenvalues are $\lambda = 1 - a - b$ and 1.
b) Solving the equation $\pi = \pi P$ for π shows that the stationary distribution is given by $(\pi_0, \pi_1) = (b/(a + b), a/(a + b))$.

c) The relation is clearly verified for $n = 0$. Next, assuming that it holds at the rank n, we have

$$\begin{aligned}
&\begin{bmatrix} \mathbb{E}\left[\exp\left(t\sum_{k=1}^{n+1} X_k\right) \,\middle|\, X_0 = 0\right] \\ \mathbb{E}\left[\exp\left(t\sum_{k=1}^{n+1} X_k\right) \,\middle|\, X_0 = 1\right] \end{bmatrix} \\
&= \begin{bmatrix} (1-a)\,\mathbb{E}\left[\exp\left(t\sum_{k=2}^{n+1} X_k\right) \,\middle|\, X_1 = 0\right] + a\mathrm{e}^t\,\mathbb{E}\left[\exp\left(t\sum_{k=2}^{n+1} X_k\right) \,\middle|\, X_1 = 1\right] \\ b\,\mathbb{E}\left[\exp\left(t\sum_{k=2}^{n+1} X_k\right) \,\middle|\, X_1 = 0\right] + (1-b)\mathrm{e}^t\,\mathbb{E}\left[\exp\left(t\sum_{k=2}^{n+1} X_k\right) \,\middle|\, X_1 = 1\right] \end{bmatrix} \\
&= \begin{bmatrix} 1-a & a\mathrm{e}^t \\ b & (1-b)\mathrm{e}^t \end{bmatrix} \begin{bmatrix} \mathbb{E}\left[\exp\left(t\sum_{k=2}^{n+1} X_k\right) \,\middle|\, X_1 = 0\right] \\ \mathbb{E}\left[\exp\left(t\sum_{k=2}^{n+1} X_k\right) \,\middle|\, X_1 = 1\right] \end{bmatrix} \\
&= \begin{bmatrix} 1-a & a\mathrm{e}^t \\ b & (1-b)\mathrm{e}^t \end{bmatrix} \begin{bmatrix} \mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} X_k\right) \,\middle|\, X_0 = 0\right] \\ \mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} X_k\right) \,\middle|\, X_0 = 1\right] \end{bmatrix} \\
&= \begin{bmatrix} 1-a & a\mathrm{e}^t \\ b & (1-b)\mathrm{e}^t \end{bmatrix} \left(\begin{bmatrix} 1-a & a\mathrm{e}^t \\ b & (1-b)\mathrm{e}^t \end{bmatrix}\right)^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= \left(\begin{bmatrix} 1-a & a\mathrm{e}^t \\ b & (1-b)\mathrm{e}^t \end{bmatrix}\right)^{n+1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad t \in \mathbb{R}.
\end{aligned}$$

d) By diagonalizing P as

$$\begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\pi_0}} & 0 \\ 0 & \frac{1}{\sqrt{\pi_1}} \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & -\sqrt{\pi_1} \\ \sqrt{\pi_1} & \sqrt{\pi_0} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & \sqrt{\pi_1} \\ -\sqrt{\pi_1} & \sqrt{\pi_0} \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & 0 \\ 0 & \sqrt{\pi_1} \end{bmatrix},$$

we have

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} X_k\right)\right] = [\pi_0, \pi_1] \begin{bmatrix} \mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} X_k\right) \,\middle|\, X_0 = 0\right] \\ \mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} X_k\right) \,\middle|\, X_0 = 1\right] \end{bmatrix} \\
&= [\pi_0, \pi_1] \left(\begin{bmatrix} 1-a & a\mathrm{e}^t \\ b & (1-b)\mathrm{e}^t \end{bmatrix}\right)^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= [\pi_0, \pi_1] \left(\begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix}\right)^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\end{aligned}$$

$$
\begin{aligned}
&= [\pi_0, \pi_1] \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix} \\
&\quad \times \left(\begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix} \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix} \right)^{n-1} \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= [\pi_0, \pi_1] \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix} \\
&\quad \times \left(\begin{bmatrix} \frac{1}{\sqrt{\pi_0}} & 0 \\ 0 & \frac{\mathrm{e}^{t/2}}{\sqrt{\pi_1}} \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & -\sqrt{\pi_1} \\ \sqrt{\pi_1} & \sqrt{\pi_0} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & \sqrt{\pi_1} \\ -\sqrt{\pi_1} & \sqrt{\pi_0} \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & 0 \\ 0 & \mathrm{e}^{t/2}\sqrt{\pi_1} \end{bmatrix} \right)^{n-1} \\
&\quad \times \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{e}^{t/2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= [\pi_0, \pi_1 \mathrm{e}^{t/2}] \\
&\times \left(\begin{bmatrix} \frac{1}{\sqrt{\pi_0}} & 0 \\ 0 & \frac{1}{\sqrt{\pi_1}} \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & -\sqrt{\pi_1} \\ \mathrm{e}^{t/2}\sqrt{\pi_1} & \mathrm{e}^{t/2}\sqrt{\pi_0} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & \mathrm{e}^{t/2}\sqrt{\pi_1} \\ -\sqrt{\pi_1} & \mathrm{e}^{t/2}\sqrt{\pi_0} \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & 0 \\ 0 & \sqrt{\pi_1} \end{bmatrix} \right)^{n-1} \\
&\times \begin{bmatrix} 1 \\ \mathrm{e}^{t/2} \end{bmatrix} \\
&= [\pi_0, \pi_1 \mathrm{e}^{t/2}] \begin{bmatrix} \frac{1}{\sqrt{\pi_0}} & 0 \\ 0 & \frac{1}{\sqrt{\pi_1}} \end{bmatrix} \left(\begin{bmatrix} \sqrt{\pi_0} & -\sqrt{\pi_1} \\ \mathrm{e}^{t/2}\sqrt{\pi_1} & \mathrm{e}^{t/2}\sqrt{\pi_0} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} \sqrt{\pi_0} & \mathrm{e}^{t/2}\sqrt{\pi_1} \\ -\sqrt{\pi_1} & \mathrm{e}^{t/2}\sqrt{\pi_0} \end{bmatrix} \right)^{n-1} \\
&\times \begin{bmatrix} \sqrt{\pi_0} & 0 \\ 0 & \sqrt{\pi_1} \end{bmatrix} \begin{bmatrix} 1 \\ \mathrm{e}^{t/2} \end{bmatrix} \\
&= [\sqrt{\pi_0}, \sqrt{\pi_1}\mathrm{e}^{t/2}] \left(\begin{bmatrix} \lambda + (1-\lambda)\pi_0 & (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} \\ (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} & (\lambda + (1-\lambda)\pi_1)\mathrm{e}^{t} \end{bmatrix} \right)^{n-1} \begin{bmatrix} \sqrt{\pi_0} \\ \sqrt{\pi_1} e^{t/2} \end{bmatrix},
\end{aligned}
$$

$t \in \mathbb{R}$.

e) Taking

$$
M(t) = \begin{bmatrix} \lambda + (1-\lambda)\pi_0 & (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} \\ (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} & (\lambda + (1-\lambda)\pi_1)\mathrm{e}^{t} \end{bmatrix}
$$

$$= \begin{bmatrix} \pi_0 + \lambda\pi_1 & (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} \\ (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} & \mathrm{e}^t(\pi_1 + \lambda\pi_0) \end{bmatrix},$$

We have

$$\mu(t) = \frac{1}{2}\big(\mathrm{Tr}(M(t)) + \sqrt{(\mathrm{Tr}(M(t)))^2 - 4\lambda\mathrm{e}^t}\big),$$

where

$$\mathrm{Tr}(M(t)) = \lambda + (1-\lambda)\pi_0 + (\lambda + (1-\lambda)\pi_1)\mathrm{e}^t$$

f) Since the matrix $M(t)$ is symmetric, by Proposition 9 in Foucart (2010) we have

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(t\sum_{k=1}^n X_k\right)\right] \\
&\leq \|[\sqrt{\pi_0}, \sqrt{\pi_1}\mathrm{e}^{t/2}]\|_2 \\
&\quad \times \left\|\left(\begin{bmatrix} \lambda + (1-\lambda)\pi_0 & (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} \\ (1-\lambda)\mathrm{e}^{t/2}\sqrt{\pi_0\pi_1} & (\lambda + (1-\lambda)\pi_1)\mathrm{e}^t \end{bmatrix}\right)^{n-1}\right\|_2 \left\|\begin{bmatrix} \sqrt{\pi_0} \\ \sqrt{\pi_1}e^{t/2} \end{bmatrix}\right\|_2 \\
&= (\mu(t))^{n-1}\|[\sqrt{\pi_0}, \sqrt{\pi_1}\mathrm{e}^{t/2}]\|_2^2 \\
&= (\pi_0 + \pi_1\mathrm{e}^t)(\mu(t))^{n-1}.
\end{aligned}$$

Next, applying again Proposition 9 in Foucart (2010) to $A := \sqrt{M(t)}$, we have

$$\begin{aligned}
\mu(t) &\geq \frac{1}{\|[\sqrt{\pi_0}, \mathrm{e}^{t/2}\sqrt{\pi_1}]\|_2^2}\left\|\sqrt{M(t)}[\sqrt{\pi_0}, \mathrm{e}^{t/2}\sqrt{\pi_1}]^\top\right\|_2^2 \\
&= \frac{1}{\pi_0 + \pi_1\mathrm{e}^t}\langle[\sqrt{\pi_0}, \mathrm{e}^{t/2}\sqrt{\pi_1}], M(t)[\sqrt{\pi_0}, \mathrm{e}^{t/2}\sqrt{\pi_1}]^\top\rangle \\
&= \frac{1}{\pi_0 + \pi_1\mathrm{e}^t}\left\langle[\sqrt{\pi_0}, \mathrm{e}^{t/2}\sqrt{\pi_1}], \begin{bmatrix} \pi_0\sqrt{\pi_0} + \lambda\pi_1\sqrt{\pi_0} + (1-\lambda)\mathrm{e}^t\pi_1\sqrt{\pi_0} \\ (1-\lambda)\mathrm{e}^{t/2}\pi_0\sqrt{\pi_1} + \mathrm{e}^{3t/2}\pi_1\sqrt{\pi_1} + \lambda\mathrm{e}^{3t/2}\pi_0\sqrt{\pi_1} \end{bmatrix}\right\rangle \\
&= \frac{\pi_0^2 + 2\mathrm{e}^t\pi_0\pi_1 + \mathrm{e}^{2t}\pi_1^2 + \lambda(\pi_0\pi_1 - 2\mathrm{e}^t\pi_0\pi_1 + \mathrm{e}^{2t}\pi_0\pi_1)}{\pi_0 + \pi_1\mathrm{e}^t} \\
&= \pi_0 + \pi_1 e^t + \lambda\frac{(\pi_0 - \mathrm{e}^t\pi_1)^2}{\pi_0 + \pi_1\mathrm{e}^t} \\
&\geq \pi_0 + \pi_1 e^t
\end{aligned}$$

since $\lambda \geq 0$, which shows that

$$\mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} X_k\right)\right] \leq (\mu(t))^n, \qquad t \in \mathbb{R}_+.$$

g) By the classical Markov or Chernoff bound argument, we have

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n}(X_k - \pi_1) \geq z\right) &= \mathbb{P}\left(\exp\left(t\sum_{k=1}^{n} X_k\right) \geq \mathrm{e}^{ntz+nt\pi_1}\right) \\
&= \mathrm{e}^{-ntz-nt\pi_1}\,\mathbb{E}\left[\exp\left(t\sum_{k=1}^{n} X_k\right)\right] \\
&= \mathrm{e}^{-ntz-nt\pi_1}(\mu(t))^n \\
&= \mathrm{e}^{-n(t(\pi_1+z)-\log\mu(t))}, \qquad t > 0.
\end{aligned}$$

h) This section only sketches the solution argument, see Appendices A and B in Léon and Perron (2004) for the full proof details. By differentiating

$$\begin{aligned}
&t \mapsto xt - \log\mu(t) \\
&= xt \\
&- \log\left(\frac{1}{2}(\lambda + (1-\lambda)\pi_0 + (\lambda + (1-\lambda)\pi_1)\mathrm{e}^t\right. \\
&\qquad \left. + \sqrt{(\lambda + (1-\lambda)\pi_0 + (\lambda + (1-\lambda)\pi_1)\mathrm{e}^t)^2 - 4\lambda\mathrm{e}^t}\right)
\end{aligned}$$

with respect to $t > 0$, we find that the maximizing value $t(x)$ satisfies

$$\begin{aligned}
x &= \frac{\mu'(t)}{\mu(t)} \\
&= \frac{\mathrm{Tr}(M'(t)) + (2\mathrm{Tr}(M'(t))\mathrm{Tr}(M(t)) - 4\lambda\mathrm{e}^t)/2/\sqrt{(\mathrm{Tr}(M(t)))^2 - 4\lambda\mathrm{e}^t}}{\mathrm{Tr}(M(t)) + \sqrt{(\mathrm{Tr}(M(t)))^2 - 4\lambda\mathrm{e}^t}},
\end{aligned}$$

After multiplying the numerator and denominator by

$$\mathrm{Tr}(M(t)) - \sqrt{(\mathrm{Tr}(M(t)))^2 - 4\lambda\mathrm{e}^t}$$

and simplifying, we obtain

$$(2x-1)\sqrt{(\mathrm{Tr}(M(t)))^2 - 4\lambda\mathrm{e}^t} = (\pi_1 + \lambda\pi_0)\mathrm{e}^t - (\pi_0 + \lambda\pi_1).$$

This relation can be used to derive a quadratic equation for $\mathrm{e}^{t(x)}$, with solution

$$\mathrm{e}^{t(x)} = \frac{(\pi_0 + \lambda\pi_1)\big(2x - 1 + \sqrt{\Delta(x)}\big)}{(\pi_1 + \lambda\pi_0)\big(1 - 2x + \sqrt{\Delta(x)}\big)},$$

where

$$\Delta(x) := 1 + \frac{4\lambda(1-x)x}{\pi_0\pi_1(1-\lambda)^2},$$

which yields

$$\mu(t(x)) = \frac{(\pi_0 + \lambda\pi_1)(1 + \sqrt{\Delta})}{1 - 2x + \sqrt{\Delta}}.$$

Letting

$$g(x) := \frac{xt(x) - \log \mu(t(x))}{(x - \pi_1)^2}, \qquad x \in (0,1),$$

we check that $g'(\pi_0) = 0$ and $g(x)$ admits a global minimum at $x = \pi_0$. Then, we have

$$\Delta(\pi_0) := 1 + \frac{4\lambda}{(1-\lambda)^2} = \frac{(1+\lambda)^2}{(1-\lambda)^2},$$

$$t(\pi_0) = \log \frac{(\pi_0 + \lambda\pi_1)\big(\pi_0 - \pi_1 + \frac{1+\lambda}{1-\lambda}\big)}{(\pi_1 + \lambda\pi_0)\big(\pi_1 - \pi_0 + \frac{1+\lambda}{1-\lambda}\big)},$$

$$\mu(t(\pi_0)) = \frac{(\pi_0 + \lambda\pi_1)(1 + \frac{1+\lambda}{1-\lambda})}{\pi_1 - \pi_0 + \frac{1+\lambda}{1-\lambda}},$$

and letting $r := (b-a)/(2-a-b)$, we have

$$\begin{aligned}
g(\pi_0) &= \frac{1}{\pi_0 - \pi_1} \log \frac{1 - (1-\lambda)\pi_1}{1 - (1-\lambda)\pi_0} \\
&= \frac{a+b}{b-a} \log \frac{1-a}{1-b} \\
&= \frac{1-\lambda}{1+\lambda}\frac{1}{r} \log \frac{1+r}{1-r} \\
&= \frac{1-\lambda}{1+\lambda}\frac{1}{r}(\log(1+r) - \log(1-r)) \\
&= \frac{1-\lambda}{1+\lambda}\frac{1}{r}\left(\sum_{n\geq 1}(-1)^{n+1}\frac{r^n}{n} + \sum_{n\geq 1}\frac{r^n}{n}\right) \\
&= \frac{1-\lambda}{1+\lambda}\frac{1}{r}\sum_{n\geq 0}\frac{r^{2n+1}}{2n+1} \\
&\geq 2\frac{1-\lambda}{1+\lambda},
\end{aligned}$$

hence for $z \in [0, 1-\pi_1]$ we have

$$\log \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n (X_k-\pi_1)\geq z\right) \leq -nz^2 g(\pi_1+z)$$
$$\leq -nz^2 g(\pi_0)$$
$$\leq -2nz^2\frac{1-\lambda}{1+\lambda},$$

while

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n (X_k-\pi_1)\geq z\right)=0$$

for $z>1-\pi_1$.

Problem 6.10

a) Theorem 31 page 15 of Freedman (1983) shows that letting $\tau_0 := 0$, the sequence $(\tau_{k+1}-1-\tau_k)_{k\geq 0}$, resp. $(R^i_{\tau_{k+1}-1}-R^i_{\tau_k})_{k\geq 0}$, is made of independent random variables, $i\in\mathbb{S}$, hence by the law of large numbers for renewal processes, see Corollary 14 page 106 of Serfozo (2009), we have

$$\pi_i=\lim_{n\to\infty}\frac{\mathbb{E}[R^i_n]}{n}=\frac{\mathbb{E}[R^i_{\tau_1-1}]}{\mathbb{E}[\tau_1-1]}.$$

b) By the Wald identity, see e.g. Theorem 2 of Chewi (2017), we have

$$\mathbb{E}[T-1]=\mathbb{E}[\tau_1-1]\,\mathbb{E}[\kappa]$$

and

$$\mathbb{E}\left[\sum_{j=1}^{T-1}\mathbb{1}_{\{X_j=i\}}\right]=\mathbb{E}\left[\sum_{j=1}^{\tau_1-1}\mathbb{1}_{\{X_j=i\}}\right]\mathbb{E}[\kappa],$$

hence

$$\pi_i=\frac{\mathbb{E}\left[\sum_{j=1}^{\tau_1-1}\mathbb{1}_{\{X_j=i\}}\right]}{\mathbb{E}[\tau_1-1]}=\frac{\mathbb{E}\left[\sum_{j=1}^{T-1}\mathbb{1}_{\{X_j=i\}}\right]}{\mathbb{E}[T-1]},\qquad i\in\mathbb{S}.$$

Problem 6.11

a) Bounded regret.

i) Define the sequence $(\tau_k)_{k\geq 1}$ recursively as

$$\tau_1 := \inf\{l>1\ :\ X_l=X_1\},$$

and

$$\tau_k := \inf\{l>\tau_{k-1}\ :\ X_l=X_1\},\qquad k\geq 2,$$

and let

$$T := \inf\{l > \tau \ : \ X_l = X_1\}.$$

By Question (b) of Problem 6.10, we have

$$\pi_1^{(i)} \mathbb{E}[T-1] = \mathbb{E}\left[R_{T-1}^{(i)}\right], \qquad i \in \mathbb{S}.$$

Hence we have

$$R_{T-1}^{(i)} - (T-\tau) \le R_T^{(i)} - (T-\tau) \le R_\tau^{(i)} \le R_{T-1}^{(i)}$$

and

$$\pi_1^{(i)} \mathbb{E}[T-1] - \mathbb{E}[T-\tau] \le \mathbb{E}\left[R_\tau^{(i)}\right] \le \mathbb{E}\left[R_{T-1}^{(i)}\right] = \pi_1^{(i)} \mathbb{E}[T-1]$$

or

$$\pi_1^{(i)} \mathbb{E}[T-1] - \mathbb{E}[T-\tau] \le \mathbb{E}\left[R_\tau^{(i)}\right] \le \pi_1^{(i)} \mathbb{E}[\tau] + \mathbb{E}[T-\tau]$$

hence

$$\pi_1^{(i)} \mathbb{E}[\tau] - \mathbb{E}[T-\tau] \le \mathbb{E}\left[R_\tau^{(i)}\right] \le \pi_1^{(i)} \mathbb{E}[\tau] + \mathbb{E}[T-\tau],$$

and therefore

$$\left| \mathbb{E}\left[R_\tau^{(i)}\right] - \pi_1^{(i)} \mathbb{E}[\tau] \right| \le \mathbb{E}[T-\tau]. \tag{S.6.5}$$

ii) We have

$$\begin{aligned}
&\left| \mathbb{E}\left[\sum_{i=1}^N \sum_{k=1}^{T_n^{(i,\alpha)}} X_k^{(i)} - \sum_{i=1}^N \pi_1^{(i)} T_n^{(i,\alpha)} \right] \right| \le \sum_{i=1}^N \left| \mathbb{E}\left[\sum_{k=1}^{T_n^{(i,\alpha)}} X_k^{(i)} - \pi_1^{(i)} T_n^{(i,\alpha)} \right] \right| \\
&\le \sum_{i=1}^N \left| \mathbb{E}\left[R_{T_n^{(i,\alpha)}}^{(i)} - \pi_1^{(i)} T_n^{(i,\alpha)} \right] \right| \\
&\le \sum_{i=1}^N \mathbb{E}\left[\tau_\kappa^{(i)} - T_n^{(i,\alpha)} \right] \\
&= \sum_{i=1}^N \sum_{l,j\in\{0,1\}} \mathbb{E}\left[\tau_\kappa^{(i)} - T_n^{(i,\alpha)} \,\middle|\, X_{\tau_\kappa^{(i)}}^{(i)} = l, \ X_{T_n^{(i,\alpha)}}^{(i)} = j \right] \mathbb{P}\big(X_{\tau_\kappa^{(i)}}^{(i)} = l, \ X_{T_n^{(i,\alpha)}}^{(i)} = j \big) \\
&\le C, \quad n > N,
\end{aligned}$$

for some constant $C > 0$ independent of $n > N$, where we applied (S.6.5), see also Anantharam et al. (1987).

Remark 1.1. *Note that in general we do not have*

$$\mathbb{E}\left[\tau_\kappa^{(i)} - \tau\right] \leq \max_{j\in\mathbb{S}} \mu_j^{(i)}(j)$$

for any stopping time τ. For example, if τ is the first hitting time of state 0 by the two-state chain with transition matrix $P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$, we have

$$\begin{aligned}
\mathbb{E}\left[\tau_\kappa^{(i)} - \tau\right] &= \mu_0(0)\mathbb{P}(X_1^{(i)} = 0) + \mu_1(0)\mathbb{P}(X_1^{(i)} = 1) \\
&= \mathbb{P}(X_1^{(i)} = 0)\left(1 + \frac{a}{b}\right) + \frac{1}{a}\mathbb{P}(X_1^{(i)} = 1) \\
&= \big((1-a)\mathbb{P}(X_0^{(i)} = 0) + b\mathbb{P}(X_0^{(i)} = 1)\big)\left(1 + \frac{a}{b}\right) \\
&\quad + \frac{1}{a}\big(a\mathbb{P}(X_0^{(i)} = 0) + (1-b)\mathbb{P}(X_0^{(i)} = 1)\big).
\end{aligned}$$

In particular, when $a = b$ we find

$$\begin{aligned}
\mathbb{E}\left[\tau_\kappa^{(i)} - \tau\right] &= 2\big((1-a)\mathbb{P}(X_0^{(i)} = 0) + a\mathbb{P}(X_0^{(i)} = 1)\big) \\
&\quad + \mathbb{P}(X_0^{(i)} = 0) + \frac{1-a}{a}\mathbb{P}(X_0^{(i)} = 1),
\end{aligned}$$

which does not remain bounded as a tends to zero, whereas in this case

$$\max_{j\in\mathbb{S}} \mu_j^{(i)}(j) = \max\left(\frac{a+b}{a}, \frac{a+b}{b}\right) = 2.$$

iii) Letting

$$K := 2\sum_{i=1}^N \max_{l,j\in\mathbb{S}} \mu_l^{(i)}(j),$$

we have

$$\begin{aligned}
\mathcal{R}_n^\alpha &= n\pi_1^{(N)} - \mathbb{E}\left[\sum_{k=1}^n X_k^{(\alpha_k)}\right] \\
&\leq K + n\pi_1^{(N)} - \sum_{i=1}^N \pi_1^{(i)}\,\mathbb{E}[T_n^{(i,\alpha)}], \qquad n > N.
\end{aligned}$$

b) Bounding the modified regret.

i) If none of the stated conditions, hold, *i.e.* if

$$\begin{cases} \widehat{m}_{n-1}^{(N,\alpha^*)} + \sqrt{\dfrac{L\log n}{T_{n-1}^{(N,\alpha^*)}}} > \pi_1^{(N)}, \\ \widehat{m}_{n-1}^{(i,\alpha^*)} \leq \pi_1^{(i)} + \sqrt{\dfrac{L\log n}{T_{n-1}^{(i,\alpha^*)}}}, \\ T_{n-1}^{(i,\alpha^*)} \geq \dfrac{4L\log n}{(\pi_1^{(N)} - \pi_1^{(i)})^2}, \end{cases}$$

then we have

$$\begin{aligned} \widehat{m}_{n-1}^{(N,\alpha^*)} + \sqrt{\frac{L\log n}{T_{n-1}^{(N,\alpha^*)}}} > \pi_1^{(N)} & \\ &= \pi_1^{(i)} + \pi_1^{(N)} - \pi_1^{(i)} \\ &\geq \pi_1^{(i)} + 2\sqrt{\frac{L\log n}{T_{n-1}^{(i,\alpha^*)}}} \\ &\geq \widehat{m}_{n-1}^{(i,\alpha^*)} + \sqrt{\frac{L\log n}{T_{n-1}^{(i,\alpha^*)}}}, \end{aligned}$$

which implies $\alpha_n^* \neq i$.

ii) We have

$$\begin{aligned} T_n^{(i,\alpha^*)} &= \sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \\ &= \sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} < \widehat{n}_i\}} + \sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} \geq \widehat{n}_i\}} \\ &= \sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k}^{(i,\alpha^*)} \leq \widehat{n}_i\}} + \sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} \geq \widehat{n}_i\}} \\ &\leq \widehat{n}_i + \sum_{k=1}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} \geq \widehat{n}_i\}} \\ &\leq \widehat{n}_i + \sum_{k>\widehat{n}_i}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\{T_{k-1}^{(i,\alpha^*)} \geq \widehat{n}_i\}} \\ &\leq \widehat{n}_i + \sum_{k>\widehat{n}_i}^{n} \mathbb{1}_{\{\alpha_k^*=i\}} \mathbb{1}_{\left\{T_{k-1}^{(i,\alpha^*)} \geq \frac{4L\log k}{(\pi_1^{(N)} - \pi_1^{(i)})^2}\right\}} \\ &\leq \widehat{n}_i + \sum_{k=1+\widehat{n}_i}^{n} \mathbb{1}_{\left\{\widehat{m}_{k-1}^{(N,\alpha^*)} + \sqrt{(L\log k)/T_{k-1}^{(N,\alpha^*)}} \leq \pi_1^{(N)}\right\}} \end{aligned}$$

$$+\sum_{k=1+\widehat{n}_i}^{n} \mathbb{1}_{\left\{\widehat{m}_{k-1}^{(N,\alpha^*)}>\pi_1^{(i)}+\sqrt{(L\log k)/T_{k-1}^{(i,\alpha^*)}}\right\}},$$

hence

$$\begin{aligned}\mathbb{E}\left[T_n^{(i,\alpha^*)}\right] \leq \widehat{n}_i &+ \sum_{k=\widehat{n}_i+1}^{n} \mathbb{P}\left(\widehat{m}_{k-1}^{(N,\alpha^*)} + \sqrt{\frac{L\log k}{T_{k-1}^{(N,\alpha^*)}}} \leq \pi_1^{(N)}\right)\\ &+ \sum_{k=\widehat{n}_i+1}^{n} \mathbb{P}\left(\widehat{m}_{k-1}^{(N,\alpha^*)} > \pi_1^{(i)} + \sqrt{\frac{L\log k}{T_{k-1}^{(i,\alpha^*)}}}\right),\end{aligned}$$

see § 2.2 of Bubeck and Cesa-Bianchi (2012).

iii) By Question (h) of Problem 6.9, we have

$$\begin{aligned}&\mathbb{P}\left(\widehat{m}_{k-1}^{(N,\alpha^*)} + \sqrt{\frac{L\log k}{T_{k-1}^{(N,\alpha^*)}}} \leq \pi_1^{(N)}\right)\\ &\leq \mathbb{P}\left(\exists l \in \{1,\ldots,k\} \ : \ \frac{1}{l}\sum_{j=1}^{l}\left(X_j^{(N)} - \pi_1^{(N)}\right) + \sqrt{\frac{L\log k}{l}} \leq \pi_1^{(N)}\right)\\ &\leq \sum_{l=1}^{k}\mathbb{P}\left(\frac{1}{l}\sum_{j=1}^{l}\left(X_j^{(N)} - \pi_1^{(N)}\right) + \sqrt{\frac{L\log k}{l}} \leq \pi_1^{(N)}\right)\\ &\leq \sum_{l=1}^{k}\mathbb{P}\left(\frac{1}{l}\sum_{j=1}^{l}\left(1 - X_j^{(N)} - \left(1-\pi_1^{(N)}\right)\right) \geq \sqrt{\frac{L\log k}{l}}\right)\\ &\leq \sum_{l=1}^{k} \mathrm{e}^{-2(1-\lambda_N)(L\log k)/(1+\lambda_N)}\\ &= \sum_{l=1}^{k}\frac{1}{k^{2L(1-\lambda)/(1+\lambda)}}\\ &= \frac{1}{k^{2L(1-\lambda)/(1+\lambda)-1}},\end{aligned}$$

and similarly

$$\begin{aligned}&\mathbb{P}\left(\widehat{m}_{k-1}^{(i,\alpha^*)} > \pi_1^{(i)} + \sqrt{\frac{L\log k}{T_{k-1}^{(i,\alpha^*)}}}\right)\\ &\leq \mathbb{P}\left(\exists l \in \{1,\ldots,k\} \ : \ \frac{1}{l}\sum_{j=1}^{l} X_j^{(N)} > \pi_1^{(N)} + \sqrt{\frac{L\log k}{l}}\right)\end{aligned}$$

$$\leq \sum_{l=1}^{k} \mathbb{P}\left(\frac{1}{l}\sum_{j=1}^{l} \big(X_j^{(N)} - \pi_1^{(N)}\big) > \sqrt{\frac{L\log k}{l}} \right)$$

$$\leq \sum_{l=1}^{k} \mathrm{e}^{-2L(1-\lambda)(\log k)/(1+\lambda)}$$

$$= \frac{1}{k^{2L(1-\lambda)/(1+\lambda)-1}}.$$

iv) We have

$$\mathbb{E}\left[T_n^{(i,\alpha^*)}\right] \leq \frac{4L\log n}{(\pi_1^{(N)} - \pi_1^{(i)})^2} + \sum_{k=1}^{n} \frac{2}{k^{2L(1-\lambda)/(1+\lambda)-1}}$$

$$\leq \frac{4L\log n}{(\pi_1^{(N)} - \pi_1^{(i)})^2} + \int_1^n \frac{2}{t^{2L(1-\lambda)/(1+\lambda)-1}}dt$$

$$\leq \frac{4L\log n}{(\pi_1^{(N)} - \pi_1^{(i)})^2} + \frac{1}{L(1-\lambda)/(1+\lambda)-1}\left(1 - \frac{1}{n^{2L(1-\lambda)/(1+\lambda)-2}}\right),$$

hence

$$\begin{aligned}
\overline{\mathcal{R}}_n^{\alpha^*} &= n\pi_1^{(N)} - \mathbb{E}\left[\sum_{k=1}^{n} \pi_{\alpha_k^*}\right] \\
&= \sum_{k=1}^{n} \mathbb{E}\left[\pi_1^{(N)} - \pi_1^{(\alpha_k^*)}\right] \\
&= n\pi_1^{(N)} - \sum_{i=1}^{N} \pi_1^{(i)}\, \mathbb{E}[T_n^{(i,\alpha^*)}] \\
&= \sum_{i=1}^{N} \big(\pi_1^{(N)} - \pi_1^{(i)}\big)\, \mathbb{E}[T_n^{(i,\alpha^*)}] \\
&\leq (\log n)\sum_{i=1}^{N-1} \frac{4L}{\pi_1^{(N)} - \pi_1^{(i)}} + \sum_{i=1}^{N} \frac{\pi_1^{(N)} - \pi_1^{(i)}}{L(1-\lambda)/(1+\lambda)-1},
\end{aligned}$$

provided that $L > (1+\lambda)/(1-\lambda)$.

Problem 6.12

a) We have

$$\mathbb{P}(T_l - T_{l-1} = m) = \frac{l}{N}\left(1 - \frac{l}{N}\right)^{m-1}, \qquad m \geq 1, \quad l = 1, \ldots, N-1,$$

i.e. $T_l - T_{l-1}$ has a geometric distribution started at 1, with parameter $p_l := 1 - l/N$, $l = l, \ldots, N-1$.

b) We have

$$\mathbb{E}[T_k] = \sum_{l=1}^{k} \mathbb{E}[T_l - T_{l-1}] = \sum_{l=1}^{k} \frac{N}{l},$$

and in particular

$$\mathbb{E}[T_{N-1}] = \sum_{l=1}^{N-1} \frac{N}{l}.$$

c) We have

$$\mathrm{Var}[T_k] = \sum_{l=1}^{k} \mathrm{Var}[T_l - T_{l-1}] = \sum_{l=1}^{k} \frac{p_l}{(1-p_l)^2} = \sum_{l=1}^{k} \frac{N^2}{l^2} \left(1 - \frac{l}{N}\right),$$

and in particular

$$\mathrm{Var}[T_{N-1}] = \sum_{l=1}^{N-1} \frac{N^2}{l^2} \left(1 - \frac{l}{N}\right) \leq CN^2,$$

with

$$C := \sum_{l=1}^{\infty} \frac{1}{l^2} = \frac{\pi^2}{6} < \infty.$$

d) Since

$$\mathbb{E}[T_{N-1}] = \sum_{k=1}^{N-1} \frac{N}{k} \leq N(1 + \log N),$$

using Markov's inequality we have, for N large enough,

$$\begin{aligned}
&\mathbb{P}(T_{N-1} > (1+a)N\log N) \\
&\quad = \mathbb{P}(T_{N-1} - \mathbb{E}[T_{N-1}] > (1+a)N\log N - \mathbb{E}[T_{N-1}]) \\
&\quad \leq \mathbb{P}(T_{N-1} - \mathbb{E}[T_{N-1}] > (1+a)N\log N - N(1+\log N)) \\
&\quad \leq \mathbb{P}(T_{N-1} - \mathbb{E}[T_{N-1}] > aN\log N - N) \\
&\quad \leq \frac{\mathrm{Var}[T_{N-1}]}{(aN\log N - N)^2} \\
&\quad \leq \frac{CN^2}{(N(-1 + a\log N))^2} \\
&\quad = \frac{C}{(-1 + a\log N)^2}.
\end{aligned}$$

e) The distribution of X_n given that $1 + T_{N-1} \leq n$ is uniform on $\mathbb{S}$, because at time $1 + T_{N-1}$ all cards have been uniformly displaced, including the original bottom card after it reached the top position at time T_{N-1}.

f) Let $(Y_n)_{n\geq 0}$ denote a Markov chain with same transition matrix as $(X_n)_{n\geq 0}$, but started in the uniform stationary distribution. Since X_n has the uniform distribution π given that $1+T_{N-1}\leq n$, by the coupling argument of Proposition 6.24 and the answers to Questions (b) and (d), for N large enough we find the convergence rate in total variation to the uniform distribution

$$\begin{aligned}
&\big\|\mathbb{P}\big(X_{1+(1+a)N\log N}\in\cdot\big)-\pi\big\|_{\mathrm{TV}} = \sup_{A\subset\mathbb{S}}\big|\mathbb{P}\big(X_{1+(1+a)N\log N}\in A\big)-\pi(A)\big| \\
&= \sup_{A\subset\mathbb{S}}\big|\mathbb{P}\big(X_{1+(1+a)N\log N}\in A\big)-\mathbb{P}\big(Y_{1+(1+a)N\log N}\in A\big)\big| \\
&\leq \sup_{A\subset\mathbb{S}}\big|\mathbb{P}\big(X_{1+(1+a)N\log N}\in A \text{ and } T_{N-1}\leq (1+a)N\log N\big) \\
&\quad -\mathbb{P}\big(Y_{1+(1+a)N\log N}\in A \text{ and } T_{N-1}\leq (1+a)N\log N\big)\big| \\
&\quad +\sup_{A\subset\mathbb{S}}\big|\mathbb{P}\big(X_{1+(1+a)N\log N}\in A \text{ and } 1+T_{N-1}> 1+(1+a)N\log N\big) \\
&\quad -\mathbb{P}\big(Y_{1+(1+a)N\log N}\in A \text{ and } 1+T_{N-1}> 1+(1+a)N\log N\big)\big| \\
&= \sup_{A\subset\mathbb{S}}\big|\mathbb{P}\big(X_{1+(1+a)N\log N}\in A \text{ and } 1+T_{N-1}> 1+(1+a)N\log N\big) \\
&\quad -\mathbb{P}\big(Y_{1+(1+a)N\log N}\in A \text{ and } 1+T_{N-1}> 1+(1+a)N\log N\big)\big| \\
&\leq \mathbb{P}\big(1+T_{N-1}>1+(1+a)N\log N\big) \\
&\leq \frac{C}{(-1+a\log N)^2},
\end{aligned}$$

provided that $a>0$.

Remark. It can also be shown that

$$\liminf_{N\to\infty}\big\|\mathbb{P}\big(X_{(1+a)N\log N}\in\cdot\big)-\pi\big\|_{\mathrm{TV}}>0,$$

for all $a\in(-1,0)$, which shows that the speed $N\log N$ is optimal for the convergence of the random shuffling $(X_n)_{n\geq 0}$ to the uniform distribution on $\mathbb{S}$ in total variation distance as N tends to infinity.

In addition to the top-to-random shuffle, other types of shuffling include the random transpositions shuffle, the transposing neighbors shuffle, the overhand shuffle, the riffle shuffle, etc.

Problem 6.13 (cf. Levin et al. (2009)-§ 4.3-4.5)

a) For any two probability distributions $\mu=[\mu_1,\mu_2,\ldots,\mu_N]$ and $\nu=[\nu_1,\nu_2,\ldots,\nu_N]$ on $\{1,2,\ldots,N\}$ we have

$$
\begin{aligned}
\|\mu - \nu\|_{\mathrm{TV}} &= \frac{1}{2}\sum_{k=1}^{N} |\mu_k - \nu_k| \\
&\leq \frac{1}{2}\sum_{k=1}^{N} (\mu_k + \nu_k) \\
&= \frac{1}{2}\sum_{k=1}^{N} \mu_k + \frac{1}{2}\sum_{k=1}^{N} \nu_k \\
&= 1.
\end{aligned}
$$

b) We have

$$
\begin{aligned}
\|\mu P - \nu P\|_{\mathrm{TV}} &= \frac{1}{2}\sum_{j=1}^{N} |[\mu P]_j - [\nu P]_j| \\
&= \frac{1}{2}\sum_{j=1}^{N} \left| \sum_{i=1}^{n} \mu_i P_{i,j} - \sum_{i=1}^{n} \nu_i P_{i,j} \right| \\
&\leq \frac{1}{2}\sum_{j=1}^{N} \sum_{i=1}^{n} P_{i,j} |\mu_i - \nu_i| \\
&= \frac{1}{2}\sum_{i=1}^{n} |\mu_i - \nu_i| \sum_{j=1}^{N} P_{i,j} \\
&= \frac{1}{2}\sum_{i=1}^{n} |\mu_i - \nu_i|.
\end{aligned}
$$

c) Replacing μ and ν with μP^n and π in the result of Question (b) we find

$$
\begin{aligned}
\|\mu P^{n+1} - \pi\|_{\mathrm{TV}} &= \|(\mu P^n) P - \pi P\|_{\mathrm{TV}} \\
&\leq \|\mu P^n - \pi\|_{\mathrm{TV}}.
\end{aligned}
$$

d) Letting $k \in \{1, 2, \ldots, N\}$ and taking

$$
\mu := (0, \ldots, 0, \underset{\substack{\uparrow \\ k}}{1}, 0, \ldots, 0)
$$

we have $\mu P^{n+1} = [P^{n+1}]_{k,\cdot}$ and by Question (c) we find

$$
\begin{aligned}
\|[P^{n+1}]_{k,\cdot} - \pi\|_{\mathrm{TV}} &= \|\mu P^{n+1} - \pi P\|_{\mathrm{TV}} \\
&\leq \|\mu P^n - \pi\|_{\mathrm{TV}} \\
&= \|[P^n]_{k,\cdot} - \pi\|_{\mathrm{TV}}.
\end{aligned}
$$

Taking the maximum over $k = 1, 2, \ldots, N$ in the above inequality yields

$$\begin{aligned} d(n+1) &= \max_{k=1,2,\ldots,N} \|[P^{n+1}]_{k,\cdot} - \pi\|_{\mathrm{TV}} \\ &\leq \max_{k=1,2,\ldots,N} \|[P^{n}]_{k,\cdot} - \pi\|_{\mathrm{TV}} \\ &= d(n), \qquad n \in \mathbb{N}. \end{aligned}$$

e) The chain is irreducible because all states can communicate in one time step since $P_{i,j} > 0$, $1 \leq i, j \leq N$. In addition the chain is aperiodic as all states have period one, given that $P_{i,i} > 0$, $i = 1, 2, \ldots, N$. Since the state space is finite, Corollary 6.2 shows that all states are positive recurrent, hence by Corollary 6.7 the chain admits a limiting and a stationary distribution that are equal.

f) We note that Q_θ can be written as

$$\begin{aligned} Q_\theta &= [\ [Q_\theta]_{i,j}\ ]_{1\leq i,j\leq N} \\ &= \begin{bmatrix} [Q_\theta]_{1,1} & [Q_\theta]_{1,2} & \cdots & [Q_\theta]_{1,N} \\ [Q_\theta]_{2,1} & [Q_\theta]_{2,2} & \cdots & [Q_\theta]_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ [Q_\theta]_{N,1} & [Q_\theta]_{N,2} & \cdots & [Q_\theta]_{N,N} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{1-\theta}(P_{1,1} - \theta\pi_1) & \frac{1}{1-\theta}(P_{1,2} - \theta\pi_2) & \cdots & \frac{1}{1-\theta}(P_{1,N} - \theta\pi_N) \\ \frac{1}{1-\theta}(P_{2,1} - \theta\pi_1) & \frac{1}{1-\theta}(P_{2,2} - \theta\pi_2) & \cdots & \frac{1}{1-\theta}(P_{2,N} - \theta\pi_N) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{1-\theta}(P_{N,1} - \theta\pi_1) & \frac{1}{1-\theta}(P_{N,2} - \theta\pi_2) & \cdots & \frac{1}{1-\theta}(P_{N,N} - \theta\pi_N) \end{bmatrix} \end{aligned}$$

Clearly, all entries of Q_θ are nonnegative due to the condition

$$P_{i,j} \geq \theta\pi_j, \qquad i, j = 1, 2, \ldots, N.$$

In addition, for all $i = 1, 2, \ldots, N$ we have

$$\begin{aligned} \sum_{j=1}^{N} [Q_\theta]_{i,j} &= \frac{1}{1-\theta} \sum_{j=1}^{N} (P_{i,j} - \theta\varPi_{i,j}) \\ &= \frac{1}{1-\theta} \sum_{j=1}^{N} (P_{i,j} - \theta\pi_j) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{1-\theta}\sum_{j=1}^{N} P_{i,j} - \frac{\theta}{1-\theta}\sum_{j=1}^{N}\pi_j \\
&= \frac{1}{1-\theta} - \frac{\theta}{1-\theta} \\
&= 1, \qquad 0 < \theta < 1,
\end{aligned}
$$

and we conclude that Q_θ is a Markov transition matrix.

g) Clearly, the property holds for $n = 1$ by the definition of Q_θ. Next, assume that

$$P^n = \Pi + (1-\theta)^n \left(Q_\theta^n - \Pi\right)$$

for some $n \geq 1$. Noting that the condition $\pi P = \pi$ implies $\Pi P = \Pi$, we have

$$
\begin{aligned}
P^{n+1} &= \big(\Pi + (1-\theta)^n (Q_\theta^n - \Pi)\big) P \\
&= \Pi P + (1-\theta)^n Q_\theta^n P - (1-\theta)^n \Pi P \\
&= \Pi + (1-\theta)^n Q_\theta^n P - (1-\theta)^n \Pi \\
&= \Pi + (1-\theta)^n Q_\theta^n \big(\Pi + (1-\theta)(Q_\theta - \Pi)\big) - (1-\theta)^n \Pi \\
&= \Pi + \theta(1-\theta)^n Q_\theta^n \Pi + (1-\theta)^{n+1} Q_\theta^{n+1} - (1-\theta)^n \Pi.
\end{aligned}
$$

Next, we note that since Q_θ is a Markov transition matrix by Question (f) we have $Q_\theta \Pi = \Pi$, in other words we have $P\Pi = \Pi^2 = \Pi$, and

$$Q_\theta \Pi = \frac{1}{1-\theta}\left(P - \theta\Pi\right)\Pi = \frac{1}{1-\theta}\big(P\Pi - \theta\Pi^2\big) = \frac{1}{1-\theta}\left(\Pi - \theta\Pi\right) = \Pi,$$

and more generally $Q_\theta^n \Pi = \Pi$, $n \geq 1$, hence

$$
\begin{aligned}
P^{n+1} &= \Pi + \theta(1-\theta)^n Q_\theta^n \Pi + (1-\theta)^{n+1} Q_\theta^{n+1} - (1-\theta)^n \Pi \\
&= \Pi + \theta(1-\theta)^n \Pi + (1-\theta)^{n+1} Q_\theta^{n+1} - (1-\theta)^n \Pi \\
&= \Pi + (1-\theta)^{n+1} Q_\theta^{n+1} - (1-\theta)^{n+1} \Pi \\
&= \Pi + (1-\theta)^{n+1}\big(Q_\theta^{n+1} - \Pi\big).
\end{aligned}
$$

h) Let $k \in \{1, 2, \ldots, N\}$. By Question (g) we have

$$
\begin{aligned}
\|[P^n]_{k,\cdot} - \pi\|_{\mathrm{TV}} &= \|[P^n]_{k,\cdot} - \Pi_{k,\cdot}\|_{\mathrm{TV}} \\
&= \frac{1}{2}\sum_{j=1}^{N} \left|[P^n]_{k,j} - \pi_j\right| \\
&= \frac{1}{2}\sum_{j=1}^{N} \left|(1-\theta)^n [Q_\theta^n]_{k,j} - (1-\theta)^n \pi_j\right|
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(1-\theta)^n}{2} \sum_{j=1}^{N} |[Q_\theta^n]_{k,j} - \pi_j| \\
&= (1-\theta)^n \|[Q_\theta^n]_{k,\cdot} - \pi\|_{\mathrm{TV}} \\
&\le (1-\theta)^n, \qquad n \ge 0,
\end{aligned}
$$

where we applied the result of Question (a), since $\Pi_{k,\cdot} = \pi$ is a probability distribution and the same holds for $[Q_\theta^n]_{k,\cdot}$ for all $k = 1, 2, \ldots, N$ by Question (f).

The relation

$$
\|[P^n]_{k,\cdot} - \pi\|_{\mathrm{TV}} = (1-\theta)^n \|[Q_\theta^n]_{k,\cdot} - \pi\|_{\mathrm{TV}}, \qquad n \ge 0,
$$

also shows that, in total variation distance, at each time step the chain associated to P converges faster (by a factor $1-\theta$) to π than the chain associated to Q_θ.

Finally, we find

$$
d(n) = \max_{k=1,2,\ldots,N} \|[P^n]_{k,\cdot} - \pi\|_{\mathrm{TV}} \le (1-\theta)^n, \qquad n \ge 0.
$$

i) If $t_{\mathrm{mix}} = 0$ the inequality is clearly satisfied, so that we can suppose that $t_{\mathrm{mix}} \ge 1$. By the definition of t_{mix} and the result of Question (h) we have

$$
\frac{1}{4} < d(t_{\mathrm{mix}} - 1) \le (1-\theta)^{t_{\mathrm{mix}}-1},
$$

hence

$$
\log \frac{1}{4} < \log d(t_{\mathrm{mix}} - 1) \le \log\left((1-\theta)^{t_{\mathrm{mix}}-1}\right) = (t_{\mathrm{mix}} - 1)\log(1-\theta),
$$

and

$$
t_{\mathrm{mix}} - 1 \le \frac{\log d(t_{\mathrm{mix}} - 1)}{\log(1-\theta)} < \frac{\log 1/4}{\log(1-\theta)}.
$$

Hence we have

$$
t_{\mathrm{mix}} < 1 + \frac{\log 1/4}{\log(1-\theta)},
$$

which yields

$$
t_{\mathrm{mix}} < 1 + \left\lceil \frac{\log 1/4}{\log(1-\theta)} \right\rceil,
$$

and finally

$$
t_{\mathrm{mix}} \le \left\lceil \frac{\log 1/4}{\log(1-\theta)} \right\rceil.
$$

j) Given the transition matrix

$$P = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/3 & 1/2 & 1/6 \\ 1/6 & 2/3 & 1/6 \end{bmatrix}$$

and its stationary distribution

$$\pi = [\pi_1, \pi_2, \pi_3] = [11/24, 9/24, 4/24],$$

we check that in order to satisfy all nine conditions $P_{i,j} \geq \theta\pi_j$, $i, j = 1, 2, 3$, the value of θ should be in the range $[0, 4/11]$. The optimal value of θ is the one that minimizes the bound $\left\lceil \frac{\log 1/4}{\log(1-\theta)} \right\rceil$, *i.e.* $\theta = 4/11$, and

$$t_{\text{mix}} \leq \left\lceil \frac{\log 1/4}{\log 7/11} \right\rceil = \lceil 3.067 \rceil = 4.$$

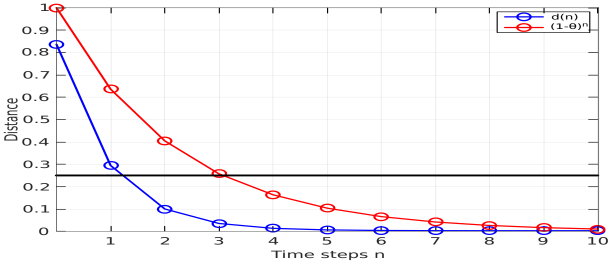


Fig. S.5: Graphs of distance to stationarity $d(n)$ and upper bound $(1-\theta)^n$.

We check from the above graph that the actual value of the mixing time is $t_{\text{mix}} = 2$. The value of $d(0)$ is the maximum distance between π and all deterministic initial distributions starting from states $k = 1, 2, \ldots, N$.

Remark. We have shown that the conditions $\pi P = \pi$ and $P_{i,j} \geq \theta\pi_j$, $i, j = 1, 2, \ldots, N$, for some $\theta \in (0, 1)$, define a unique (stationary) distribution π which is also a limiting distribution independent of the initial state. This is the case in particular when $P_{i,j} > 0$, $i, j = 1, 2, \ldots, N$, in which case the chain is irreducible and aperiodic, and admits a unique limiting and stationary distribution. More generally, the result holds when P is *regular*, *i.e.* when there exists $n \geq 1$ such that $[P^n]_{i,j} > 0$ for all $i, j = 1, 2, \ldots, N$, cf. § 4.3-4.5 of Levin et al. (2009).

Below is the Matlab/Octave code used to generate Figure S.5.

```
P = [2/3,1/6,1/6;
       1/3,1/2,1/6;
       1/6,2/3,1/6;]
pi = [11/24,9/24,4/25]
theta = 4/11
for n = 1:11
y(n)=n-1;
u(n)=0.25;
z(n)=(1-theta)^(n-1);
distance(n) = 0;
for k = 1:3
d = mpower(P,n-1)(k,1:3) - pi;
dist=0;
for i = 1:3
dist = dist + 0.5*abs(d(i));
end
distance(n) = max(distance(n) ,dist);
end
end
graphics_toolkit("gnuplot");
plot(y,distance,'-bo','LineWidth',8,y,z,'-ro','LineWidth',8,y,u,'-k',
       'LineWidth',8)
legend('d(n)','(1-\theta)^n')
set (gca, 'xtick', 1:10)
set (gca, 'ytick', 0:0.1:1)
grid on
xlabel('time steps n')
ylabel('distance')
pause
```

Problem 6.14 (cf. Lezaud (1998))

a) By the Perron-Frobenius theorem applied to the nonnegative matrix P, the largest eigenvalue λ_0 of P has a single multiplicity and satisfies

$$1 = \min_{1 \le i \le d} \sum_{j=1}^{d} P_{i,j} \le \lambda_0 \le \max_{1 \le i \le d} \sum_{j=1}^{d} P_{i,j} = 1.$$

Moreover, the eigenvector with eigenvalue $\lambda_0 = 1$ is clearly $\vec{e} = (1, \ldots, 1)$, as $P\vec{e} = \vec{e}$.

b) The projection operator Π onto $\vec{e}$ is the linear mapping given by

$$u \mapsto \Pi(u) = \frac{\langle u, \vec{e} \rangle}{\langle \vec{e}, \vec{e} \rangle} \vec{e} = \langle u, \vec{e} \rangle \vec{e} = \sum_{i=1}^{d} \langle u, \vec{e} \rangle \vec{e}_i,$$

where $\{\vec{e}_1, \ldots, \vec{e}_d\}$ is in the orthogonal basis

$$e_k := (0, \ldots, 0, \underset{\substack{\uparrow \\ k}}{1}, 0, \ldots, 0), \qquad k = 1, 2, \ldots, d,$$

of $\mathbb{R}^d$. Its matrix in $\{\vec{e}_1, \ldots, \vec{e}_d\}$ is given by

$$\Pi = (\Pi_{i,j})_{1\le i,j\le d} = (\langle \vec{e}_j, \vec{e}\rangle)_{1\le i,j\le d} = (\pi_j)_{1\le i,j\le d},$$

i.e.

$$\Pi := \begin{bmatrix} \pi \\ \\ \pi \\ \\ \pi \\ \\ \vdots \\ \\ \pi \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_d \\ \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_d \\ \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_d \\ \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_d \end{bmatrix}.$$

We also note that Π is self-adjoint with respect to $\langle\cdot,\cdot\rangle$, as

$$\langle \Pi u, v\rangle = \sum_{i,j=1}^{d} \pi_i\pi_j u_i v_j = \langle u, \Pi v\rangle,$$

and its highest eigenvalue is 1.

c) The equality clearly holds for $n = 0$, due to the convention $\sum_{l=1}^{0} = 0$. Assuming that it holds at the rank $n \ge 0$, we have

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\alpha\sum_{l=1}^{n+1} f(X_l)\right) \,\Big|\, X_0 = k\right] = \mathbb{E}\left[e^{\alpha f(X_1)} \exp\left(\alpha\sum_{l=2}^{n+1} f(X_l)\right) \,\Big|\, X_0 = k\right] \\
&= \sum_{r=1}^{d} \mathbb{E}\left[\mathbf{1}_{\{X_1=r\}} e^{\alpha f(X_1)} \exp\left(\alpha\sum_{l=2}^{n+1} f(X_l)\right) \,\Big|\, X_0 = k\right] \\
&= \frac{1}{\mathbb{P}(X_0 = k)} \sum_{r=1}^{d} e^{\alpha f(r)}\, \mathbb{E}\left[\mathbf{1}_{\{X_0=k, X_1=r\}} \exp\left(\alpha\sum_{l=2}^{n+1} f(X_l)\right)\right] \\
&= \sum_{r=1}^{d} e^{\alpha f(r)} \frac{\mathbb{P}(X_0 = k, X_1 = r)}{\mathbb{P}(X_0 = k)}\, \mathbb{E}\left[\exp\left(\alpha\sum_{l=2}^{n+1} f(X_l)\right) \,\Big|\, X_0 = k,\ X_1 = r\right] \\
&= \sum_{r=1}^{d} e^{\alpha f(r)} \mathbb{P}(X_1 = r \mid X_0 = k)\, \mathbb{E}\left[\exp\left(\alpha\sum_{l=2}^{n+1} f(X_l)\right) \,\Big|\, X_0 = k,\ X_1 = r\right] \\
&= \sum_{r=1}^{d} e^{\alpha f(r)} P_{k,r}\, \mathbb{E}\left[\exp\left(\alpha\sum_{l=2}^{n+1} f(X_l)\right) \,\Big|\, X_1 = r\right] \\
&= \sum_{r=1}^{d} e^{\alpha f(r)} P_{k,r}\, \mathbb{E}\left[\exp\left(\alpha\sum_{l=1}^{n} f(X_l)\right) \,\Big|\, X_= r\right] \\
&= \sum_{r=1}^{d} P_{k,r} e^{\alpha f(r)} \sum_{l=1}^{d} \left[\left(P e^{\alpha D_f}\right)^n\right]_{r,l}
\end{aligned}$$

$$= \sum_{l=1}^{d} \left[\left(P e^{\alpha D_f} \right)^{n+1} \right]_{k,l}.$$

d) We have

$$\begin{aligned} e^{\alpha\gamma n} \mathbb{P}\left(\sum_{l=1}^{n} f(X_l) \geq n\gamma \,\Big|\, X_0 = k \right) &= e^{\alpha\gamma n}\, \mathbb{E}\left[\mathbf{1}_{\left\{ \sum_{l=1}^{n} f(X_l) \geq n\gamma \right\}} \,\Big|\, X_0 = k \right] \\ &\leq \mathbb{E}\left[\exp\left(\alpha \sum_{l=1}^{n} f(X_l) \right) \,\Big|\, X_0 = k \right] \\ &= e^{-\alpha\gamma n} \sum_{l=1}^{d} \left[\left(P e^{\alpha D_f} \right)^{n} \right]_{k,l}, \quad n \geq 0. \end{aligned}$$

e) We have

$$\begin{aligned} \sum_{k,l=1}^{d} \pi_k \left[\left(P e^{\alpha D_f} \right)^{n} \right]_{k,l} &= \left\langle \vec{e}, \left(P e^{\alpha D_f} \right)^{n} \vec{e} \right\rangle \\ &= \left\langle \vec{e}, e^{-\alpha D_f/2} \left(e^{\alpha D_f/2} P e^{\alpha D_f/2} \right)^{n} e^{\alpha D_f/2} \vec{e} \right\rangle \\ &= \left\langle e^{-\alpha D_f/2} \vec{e}, \left(e^{\alpha D_f/2} P e^{\alpha D_f/2} \right)^{n} e^{\alpha D_f/2} \vec{e} \right\rangle \\ &\leq \| e^{-\alpha D_f/2} \vec{e} \| \cdot \left\| \left(e^{\alpha D_f/2} P e^{\alpha D_f/2} \right)^{n} e^{\alpha D_f/2} \vec{e} \right\| \\ &\leq \| e^{-\alpha D_f/2} \vec{e} \| \cdot \| e^{\alpha D_f/2} \vec{e} \| \cdot \left\| \left(e^{\alpha D_f/2} P e^{\alpha D_f/2} \right)^{n} \right\| \\ &\leq e^{\alpha} (\lambda_0(\alpha))^n. \end{aligned}$$

f) By Questions (d) and (e) we have

$$\mathbb{P}\left(\sum_{l=1}^{n} f(X_l) \geq n\gamma \,\Big|\, X_0 = k \right) \leq e^{-\alpha\gamma n} e^{\alpha} (\lambda_0(\alpha))^n = e^{\alpha - n(\alpha\gamma - \log \lambda_0(\alpha))},$$

$n \geq 0$.

g) The first equality follows from the fact that $\Pi P = P$. Next, letting $M = (M_{i,j})_{1 \leq i,j \leq d}$, we have

$$\Pi D_f^n M D_f^m = \left(\sum_{l=1}^{d} \pi_l e^{n f(l)} M_{l,j} e^{m f(j)} \right)_{1 \leq i,j \leq d},$$

hence

$$\mathrm{tr}(\Pi D_f^n M D_f^m) = \sum_{j=1}^{d} \sum_{l=1}^{d} \pi_l e^{n f(l)} M_{l,j} e^{m f(j)} = \langle f^n, M f^m \rangle.$$

h) We apply II-(2.31) in Kato (1995) by matching the expansion

$$Pe^{\alpha D_f} = \sum_{n\geq 0} \alpha^n P \frac{(D_f)^n}{n!}$$

to II-(2.1) in Kato (1995) and by taking $m = 1$, see page 74 line -1 therein, since by Question (a) the multiplicity of the eigenvalue $\lambda_0(0) = 1$ of P is 1. We have

$$\begin{aligned} c_1 &= -\operatorname{tr}(PD_f S^{(0)}) \\ &= \operatorname{tr}(PD_f \mathit{\Pi}) \\ &= \operatorname{tr}(\mathit{\Pi} PD_f) \\ &= \operatorname{tr}(\mathit{\Pi} D_f) \\ &= \sum_{k=1}^{d} \pi_k f(k) \\ &= \mathbb{E}[f(X_1)] \\ &= 0, \end{aligned}$$

and

$$c_2 = -\frac{1}{2}\|f\|^2 + \frac{1}{2}\langle f, Sf\rangle \leq \frac{1}{2}\langle f, Sf\rangle \leq (1-\lambda_1)^{-1}.$$

where we used $S^{(0)} = -\mathit{\Pi}$ and $S^{(1)} = S$. Next, for $n \geq 2$ we have

$$\begin{aligned} c_n &= \sum_{p=1}^{n} \frac{(-1)^p}{p} \sum_{\substack{\nu_1+\cdots+\nu_p=n \\ k_1+\cdots+k_p=p-1 \\ \nu_1\geq 1,\ldots,\nu_p\geq 1 \\ k_1\geq 0,\ldots,k_p\geq 0}} \operatorname{tr}\left(P\frac{(D_f)^{\nu_1}}{\nu_1!} S^{(k_1)} \cdots P\frac{(D_f)^{\nu_p}}{\nu_p!} S^{(k_p)} \right) \\ &= \sum_{p=1}^{n} \frac{(-1)^{p+1}}{p} \sum_{\substack{\nu_1+\cdots+\nu_p=n \\ k_1+\cdots+k_p=p-1 \\ \nu_1\geq 1,\ldots,\nu_p\geq 1 \\ k_1\geq 0,\ldots,k_p\geq 0}} \frac{1}{\nu_1!\cdots\nu_p!} \operatorname{tr}\big(\mathit{\Pi} P(D_f)^{\nu_1} S^{(k_1')} \cdots S^{(k_{p-1}')} P(D_f)^{\nu_p}\big) \\ &= \sum_{p=1}^{n} \frac{(-1)^{p+1}}{p} \\ &\quad \sum_{\substack{\nu_1+\cdots+\nu_p=n \\ k_1+\cdots+k_p=p-1 \\ \nu_1\geq 1,\ldots,\nu_p\geq 1 \\ k_1\geq 0,\ldots,k_p\geq 0}} \frac{1}{\nu_1!\cdots\nu_p!} \langle f^{\nu_1}, S^{k_1'} P(D_f)^{\nu_2} \cdots S^{k_{p-2}'} P(D_f)^{\nu_{p-1}} S^{k_{p-1}'} P f^{\nu_p}\rangle, \end{aligned}$$

where we used $S^{(0)} = -\mathit{\Pi}$, $S^{(n)} = S^n$, Question (g), and the relation $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

i) We have

$$\sum_{\substack{k_1+\cdots+k_p=p-1\\ k_1\geq 0,\ldots,k_p\geq 0}} \mathbf{1} = \sum_{\substack{\nu_1+\cdots+\nu_p-p=p-1\\ \nu_1\geq 1,\ldots,\nu_p\geq 1}} \mathbf{1} = \binom{2p-2}{p-1}.$$

j) Since $|\lambda_1| \leq 1$ by the Perron-Frobenius theorem, we have $0 \leq 1 - \lambda_1 \leq 2$, hence

$$\begin{aligned}
c_n &= \sum_{p=1}^{n} \frac{(-1)^{p+1}}{p} \\
&\sum_{\substack{\nu_1+\cdots+\nu_p=n\\ k_1+\cdots+k_p=p-1\\ \nu_1\geq 1,\ldots,\nu_p\geq 1\\ k_1\geq 0,\ldots,k_p\geq 0}} \frac{1}{\nu_1!\cdots\nu_p!} \langle f^{\nu_1}, S^{k'_1}P(D_f)^{\nu_2}\cdots S^{k'_{p-2}}P(D_f)^{\nu_{p-1}}S^{k'_{p-1}}Pf^{\nu_p}\rangle \\
&\leq \sum_{p=1}^{n} \frac{1}{p} \\
&\sum_{\substack{\nu_1+\cdots+\nu_p=n\\ k_1+\cdots+k_p=p-1\\ \nu_1\geq 1,\ldots,\nu_p\geq 1\\ k_1\geq 0,\ldots,k_p\geq 0}} \frac{1}{\nu_1!\cdots\nu_p!} \|f^{\nu_1}\| \cdot \left\|S^{k'_1}P(D_f)^{\nu_2}\cdots S^{k'_{p-2}}P(D_f)^{\nu_{p-1}}S^{k'_{p-1}}P\right\| \cdot \|f^{\nu_p}\| \\
&\leq \sum_{p=1}^{n} \frac{1}{p} \sum_{\substack{\nu_1+\cdots+\nu_p=n\\ k_1+\cdots+k_p=p-1\\ \nu_1\geq 1,\ldots,\nu_p\geq 1\\ k_1\geq 0,\ldots,k_p\geq 0}} \frac{1}{\nu_1!\cdots\nu_p!} \left\|S^{k'_1}\cdots S^{k'_{p-1}}\right\| \\
&\leq \sum_{p=1}^{n} \frac{1}{p} \sum_{\substack{\nu_1+\cdots+\nu_p=n\\ k_1+\cdots+k_p=p-1\\ \nu_1\geq 1,\ldots,\nu_p\geq 1\\ k_1\geq 0,\ldots,k_p\geq 0}} \frac{1}{2^{\nu_1-1}\cdots 2^{\nu_p-1}} \left\|S^{k'_1}\cdots S^{k'_{p-1}}\right\| \\
&\leq \sum_{p=1}^{n} \frac{(1-\lambda_1)^{-(p-1)}}{p2^{n-p}} \sum_{\substack{\nu_1+\cdots+\nu_p=n\\ k_1+\cdots+k_p=p-1\\ \nu_1\geq 1,\ldots,\nu_p\geq 1\\ k_1\geq 0,\ldots,k_p\geq 0}} \mathbf{1} \\
&\leq \sum_{p=1}^{n} \frac{((1-\lambda_1)/2)^{-(p-1)}}{p2^{n-1}} \binom{n-1}{p-1}\binom{2p-2}{p-1} \\
&\leq \sum_{p=1}^{n} \frac{((1-\lambda_1)/2)^{-(n-1)}}{p2^{n-1}} \binom{n-1}{p-1}\binom{2p-2}{p-1} \\
&= (1-\lambda_1)^{-(n-1)} \sum_{p=1}^{n} \frac{1}{p}\binom{n-1}{p-1}\binom{2p-2}{p-1} \\
&\leq (1-\lambda_1)^{-(n-1)} \left(1 + \sum_{p=2}^{n} \frac{1}{p}\binom{n-1}{p-1}\frac{2^{2p-2}}{\sqrt{\pi p}}\right)
\end{aligned}$$

$$\leq (1-\lambda_1)^{-(n-1)} \sum_{p=0}^{n-1} \frac{1}{p+1} \binom{n-1}{p} 4^p, \qquad n \geq 2.$$

Next, we note that for $x > 0$ we have

$$\begin{aligned}
\sum_{p=0}^{n-1} \binom{n-1}{p} \frac{x^p}{p+1} &= \frac{1}{x} \int_0^x \sum_{p=0}^{n-1} \binom{n-1}{p} y^p dy \\
&= \frac{1}{x} \int_0^x (1+y)^{n-1} dy \\
&= \frac{(1+x)^n - 1}{nx} \\
&\leq \frac{(1+x)^n}{nx},
\end{aligned}$$

hence, taking $x := 4$ we obtain

$$c_n \leq (1-\lambda_1)^{-(n-1)} \frac{5^n}{4n} \leq (1-\lambda_1)^{-(n-1)} \frac{5^n}{25}, \qquad n \geq 7.$$

and we check by hand calculation that the bound

$$1 + \sum_{p=2}^{n} \frac{1}{p} \binom{n-1}{p-1} \frac{4^{p-1}}{\sqrt{\pi p}} \leq \frac{5^n}{25}$$

is also valid for $n = 3, 4, 5, 6$, hence we have

$$c_n \leq (1-\lambda_1)^{-(n-1)} \frac{5^n}{25}, \qquad n \geq 2.$$

k) Noting that $c_1 = 0$, we have

$$\begin{aligned}
\lambda_0(\alpha) &= 1 + \sum_{n\geq 2} c_n \alpha^n \\
&\leq 1 + \sum_{n\geq 2} \frac{5^{n-2}\alpha^n}{(1-\lambda_1)^{n-1}} \\
&\leq 1 + \sum_{n\geq 2} \frac{5^{n-2}\alpha^n}{(1-\lambda_1)^{n-1}} \\
&\leq 1 + \frac{\alpha^2}{1-\lambda_1} \frac{1}{1-5\alpha/(1-\lambda_1)} \\
&= 1 + \frac{\alpha^2}{1-\lambda_1-5\alpha}, \qquad \alpha \in [0, (1-\lambda_1)/5),
\end{aligned}$$

hence

$$\mathbb{P}\left(\frac{1}{n}\sum_{l=1}^{n} f(X_l) \geq \gamma\right) \leq \exp\left(\alpha - n\left(\alpha\gamma - \log\left(1+\frac{\alpha^2}{1-\lambda_1-5\alpha}\right)\right)\right)$$
$$\leq \exp\left(\frac{1-\lambda_1}{5} - n\gamma\alpha + \frac{n\alpha^2}{1-\lambda_1-5\alpha}\right),$$

$\alpha \in [0,(1-\lambda_1)/5)$.

l) We minimize

$$\alpha \mapsto -\gamma\alpha + \frac{\alpha^2}{1-\lambda_1-5\alpha}$$

over $\alpha \in [0,(1-\lambda_1)/5)$ by noting that the vanishing of its derivative

$$5\left(\frac{\alpha}{1-\lambda_1-5\alpha}\right)^2 + 2\frac{\alpha}{1-\lambda_1-5\alpha} - \gamma = 0$$

occurs at

$$\frac{\alpha_*}{1-\lambda_1-5\alpha_*} = \frac{-1+\sqrt{1+5\gamma}}{5},$$

i.e.

$$\alpha_* = (1-\lambda_1)\frac{-1+\sqrt{1+5\gamma}}{5\sqrt{1+5\gamma}} = \frac{(1-\lambda_1)\gamma}{1+5\gamma+\sqrt{1+5\gamma}} < \frac{1-\lambda_1}{5},$$

hence

$$\begin{aligned}
-\gamma\alpha_* + \frac{\alpha_*^2}{1-\lambda_1-5\alpha_*} &= \alpha_*\left(-\gamma + \frac{\alpha_*}{1-\lambda_1-5\alpha_*}\right) \\
&= \alpha_*\frac{-1-5\gamma+\sqrt{1+5\gamma}}{5} \\
&= (1-\lambda_1)\gamma\frac{-1-5\gamma+\sqrt{1+5\gamma}}{5(1+5\gamma+\sqrt{1+5\gamma})} \\
&= (1-\lambda_1)\gamma\frac{1+5\gamma-(1+5\gamma)^2}{5(1+5\gamma+\sqrt{1+5\gamma})^2} \\
&= -\frac{(1-\lambda_1)\gamma^2(1+5\gamma)}{(1+5\gamma+\sqrt{1+5\gamma})^2} \\
&= -\frac{(1-\lambda_1)\gamma^2}{(1+\sqrt{1+5\gamma})^2} \\
&\leq -\frac{(1-\lambda_1)\gamma^2}{(1+\sqrt{6})^2} \\
&\leq -\frac{(1-\lambda_1)\gamma^2}{7+2\sqrt{6}} \\
&< -(1-\lambda_1)\frac{\gamma^2}{12}.
\end{aligned}$$

Chapter 7 - Ising Model

Exercise 7.1 (See also **here**). By first step analysis, we have

$$\begin{cases} h(3) = 1 + h(2), \\ h(2) = 1 + \dfrac{2}{3}h(1) + \dfrac{1}{3}h(3), \\ h(1) = 1 + \dfrac{1}{3} \times 0 + \dfrac{2}{3}h(2), \\ h(0) = 0, \end{cases}$$

which yields

$$\begin{aligned} h(2) &= 1 + \frac{2}{3}\left(1 + \frac{2}{3}h(2)\right) + \frac{1}{3}(1 + h(2)) \\ &= 1 + \frac{2}{3} + \frac{4}{9}h(2) + \frac{1}{3} + \frac{1}{3}h(2) \\ &= 2 + \frac{7}{9}h(2), \end{aligned}$$

hence

$$\begin{cases} h(3) = 10, \\ h(2) = 9, \\ h(1) = 7 \\ h(0) = 0. \end{cases}$$

Problem 7.2 (See also **here**).

a) We have $h(d) = 0$.
b) We have $h(0) = 1 + h(1)$.
c) We have

$$h(r) = 1 + \frac{r}{d}h(r-1) + \frac{d-r}{d}h(r+1), \qquad r = 1, 2, \ldots, d-1.$$

d) We have

$$h(r) = 1 + \frac{r}{d}h(r-1) + \frac{d-r}{d}h(r+1), \qquad r = 1, 2, \ldots, d-1,$$

hence

$$\frac{r}{d}h(r) + \frac{d-r}{d}h(r) = 1 + \frac{r}{d}h(r-1) + \frac{d-r}{d}h(r+1),$$

hence

$$\frac{r}{d}f(r-1) = 1 + \frac{d-r}{d}f(r), \qquad r = 1,2,\ldots,d-1.$$

e) We have $f(0) = h(1) - h(0) = -1$, and

$$f(r) = -\frac{d}{d-r} + \frac{r}{d-r}f(r-1), \qquad r = 1,2,\ldots,d,$$

hence

$$f(r) = -\frac{1}{\binom{d-1}{r}}\sum_{l=0}^{r}\binom{d}{l}, \qquad r = 0,1,\ldots,r.$$

f) We have

$$\begin{aligned} h(r) &= h(d) + \sum_{k=r}^{d-1}(h(k) - h(k+1)) \\ &= h(d) - \sum_{k=r}^{d-1} f(k) \\ &= \sum_{k=r}^{d-1}\frac{1}{\binom{d-1}{k}}\sum_{l=0}^{k}\binom{d}{l}, \qquad r = 0,1,\ldots,d. \end{aligned}$$

g) We have

$$h(0) = \sum_{k=0}^{d-1}\frac{1}{\binom{d-1}{k}}\sum_{l=0}^{k}\binom{d}{l}$$

and

$$h(1) = \sum_{k=1}^{d-1}\frac{1}{\binom{d-1}{k}}\sum_{l=0}^{k}\binom{d}{l}, \quad \text{and} \quad h(2) = \sum_{k=2}^{d-1}\frac{1}{\binom{d-1}{k}}\sum_{l=0}^{k}\binom{d}{l}.$$

h) i) When $d = 1$ we find $h(0) = 1$, $h(1) = 0$.
ii) When $d = 2$ we find $h(0) = 4$, $h(1) = 3$, h(2) = 0.
iii) When $d = 3$ we have $h(0) = 10$, $h(1) = 9$, $h(2) = 7$, $h(3) = 0$.

Remark. This random walk is the same as the one in Exercises 6.7 and 7.3 in Privault (2018) on the Ehrenfest chain.

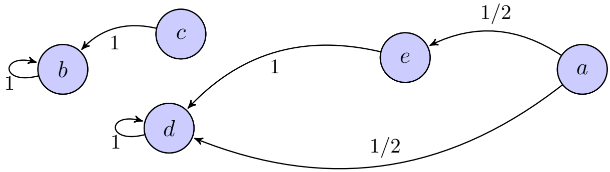
Chapter 8 - Search Engines

Problem 8.1

a) The transition matrix of the chain $(X_n)_{n\geq 0}$ is given as follows:

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

b) The chain $(X_n)_{n\geq 0}$ admits the following graph, and is clearly reducible:



c) Starting from state ⓐ, ⓓ or ⓔ, the limiting distribution is $(0, 0, 0, 1, 0)$, starting from state ⓑ or ⓒ, the limiting distribution is $(0, 1, 0, 0, 0)$, so that although the chain admits limiting distributions, it does *not* admit a limiting distribution independent of the initial state. More precisely, it can be checked that the powers P^n of the transition matrix P take the form

$$P^n = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{for all } n \geq 2, \text{ hence} \quad \lim_{n\to\infty} P^n = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

d) The equation $\pi = \pi P$ is satisfied by any probability distribution of the form

$$\pi = [\pi_a, \pi_b, \pi_c, \pi_d, \pi_e] = [0, p, 0, 1-p, 0],$$

with $p \in [0, 1]$. The stationary distribution is not unique here because the chain is reducible.

e) All rows in the matrix $\widetilde{P}$ clearly add up to 1, so $\widetilde{P}$ is a Markov transition matrix. On the other hand, all states become accessible from each other so that the new chain is irreducible and all states have period 1.

f) Since the chain is irreducible, aperiodic and has a finite state space, we know by Corollary 6.7 that it admits a unique stationary distribution $\tilde{\pi}$. The equation $\tilde{\pi} = \tilde{\pi}\widetilde{P}$ reads

$$\tilde{\pi} = \tilde{\pi}\widetilde{P}$$

$$= \frac{\varepsilon}{n}\tilde{\pi}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} + (1-\varepsilon)\tilde{\pi}P$$

$$= \left[\frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}\right] + (1-\varepsilon)\tilde{\pi}P.$$

g) The equation

$$\tilde{\pi} = \left[\frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}\right] + (1-\varepsilon)\tilde{\pi}P$$

reads

$$[\pi_a, \pi_b, \pi_c, \pi_d, \pi_e] = \left[\frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}\right] + (1-\varepsilon)\tilde{\pi}\begin{bmatrix} 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

which admits the solution

$$\begin{cases} \pi_a = \dfrac{\varepsilon}{5}, \\[2ex] \pi_b = \dfrac{2-\varepsilon}{5}, \\[2ex] \pi_c = \dfrac{\varepsilon}{5}, \\[2ex] \pi_d = \dfrac{(2-\varepsilon)(3-\varepsilon)}{10}, \\[2ex] \pi_e = \dfrac{(3-\varepsilon)\varepsilon}{10}. \end{cases} \tag{S.8.6}$$

h) We note that

$$\pi_a = \pi_c < \pi_e < \pi_b < \pi_d,$$

hence we will rank the states as

Rank	State
1	d
2	b
3	e
4	$a \simeq c$

based on the idea that the most visited states should rank higher. In the graph of Figure S.6 the stationary distribution is plotted as a function of $\varepsilon \in [0, 1]$.

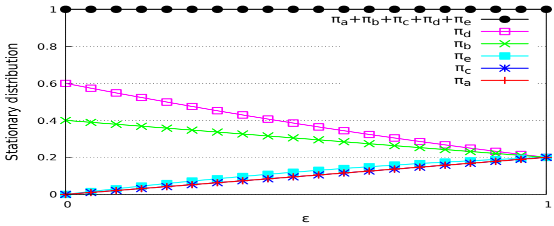


Fig. S.6: Stationary distribution as a function of $\varepsilon \in [0, 1]$.

We note again that the ranking of states is clearer for smaller values of ε. On the other hand, ε cannot be be chosen too large, for example taking $\varepsilon = 1$ makes all mean return times equal and corresponds to a uniform stationary distribution. This can be illustrated using the following R code.

```
library("igraph"); library(markovchain)
P<-matrix(c(0,0,0,0.5,0.5,0,1,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,1,0),nrow=5,
       byrow=TRUE)
MC <-new("markovchain",transitionMatrix=P,states=c("a","b","c","d","e"))
graph <- as(MC, "igraph")
plot(graph,vertex.size=50,edge.label.cex=2, edge.label=E(graph)$prob,
       edge.color='black', vertex.color='dodgerblue', vertex.label.cex=3)
page_rank(graph,damping=0.97)
$vector
       a b c d e
0.00600 0.39400 0.00600 0.58509 0.00891
```

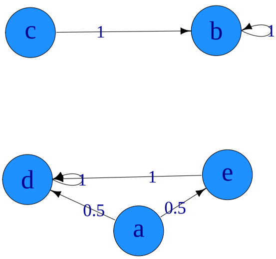


Fig. S.7: Markovchain package output.

i) By Corollary 6.7, we find

$$
\begin{cases}
\mu_a(a) = \dfrac{5}{\varepsilon} \\[2ex]
\mu_b(b) = \dfrac{5}{2-\varepsilon} \\[2ex]
\mu_c(c) = \dfrac{5}{\varepsilon} \\[2ex]
\mu_d(d) = \dfrac{10}{(2-\varepsilon)(3-\varepsilon)} \\[2ex]
\mu_e(e) = \dfrac{10}{\varepsilon(3-\varepsilon)}.
\end{cases}
$$

In the graph of Figure S.8 the mean return times are plotted as a function of $\varepsilon \in [0,1]$. A commonly used value in the literature is $\varepsilon = 1/7$.

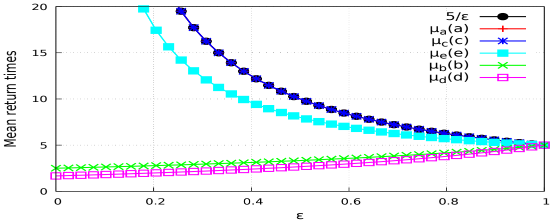


Fig. S.8: Mean return times as functions of $\varepsilon \in [0,1]$.

For small values of ε the mean return times can be higher, and therefore the simulations may take a longer time.

Chapter 9 - Hidden Markov Model

Exercise 9.1

a) By summing over $o_1, \ldots, o_t$ we have

$$\mathbb{P}(X_t = i_t, \ldots, X_0 = i_0)$$

$$= \mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1}) \cdots \mathbb{P}(X_1 = i_1 \mid X_0 = i_0)\mathbb{P}(X_0 = i_0)$$
$$= \mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1})\mathbb{P}(X_{t-1} = i_{t-1}, \ldots, X_0 = i_0),$$

which recovers (1.1) as

$$\mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1}), \quad t \geq 1.$$

b) We have

$$\begin{aligned}
&\mathbb{P}(X_t = i_t, \ldots, X_0 = i_0,\ O_t = o_t, \ldots, O_1 = o_1) \\
&= \mathbb{P}(O_t = o_t \mid X_t = i_t)\mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1}) \\
&\quad \mathbb{P}(X_{t-1} = i_{t-1}, \ldots, X_0 = i_0, O_{t-1} = o_{t-1}, \ldots, O_1 = o_1),
\end{aligned}$$

hence by summing over $i_0, i_1, \ldots, i_{t-2}$ and o_{t-1}, we have

$$\begin{aligned}
&\mathbb{P}(X_t = i_t, X_{t-1} = i_{t-1},\ O_{t-1} = o_{t-1}, \ldots, O_1 = o_1) \\
&= \mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1})\mathbb{P}(X_{t-1} = i_{t-1}, O_{t-1} = o_{t-1}, \ldots, O_1 = o_1),
\end{aligned}$$

which implies

$$\begin{aligned}
&\mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1},\ O_{t-1} = o_{t-1}, \ldots, O_1 = o_1) \\
&\qquad = \mathbb{P}(X_t = i_t \mid X_{t-1} = i_{t-1}), \qquad t \geq 1. \qquad \text{(S.9.7)}
\end{aligned}$$

Exercise 9.2

a) We have

$$\begin{aligned}
&\mathbb{P}(O_{t+1} = v,\ O_t = u) = \sum_{x \in \mathbb{S}} \mathbb{P}(O_{t+1} = v,\ O_t = u,\ X_t = x) \\
&= \sum_{x \in \mathbb{S}} \mathbb{P}(O_{t+1} = v \mid X_t = x)\mathbb{P}(X_t = x, O_t = u) \\
&= \sum_{x \in \mathbb{S}} \mathbb{P}(O_{t+1} = v,\ X_t = x)\mathbb{P}(O_t = u \mid X_t = x) \\
&= \sum_{x,y \in \mathbb{S}} \mathbb{P}(O_{t+1} = v,\ X_{t+1} = y,\ X_t = x) M_{x,u} \\
&= \sum_{x,y \in \mathbb{S}} \mathbb{P}(O_{t+1} = v \mid X_{t+1} = y,\ X_t = x)\mathbb{P}(X_{t+1} = y,\ X_t = x) M_{x,u} \\
&= \sum_{x,y \in \mathbb{S}} \mathbb{P}(O_{t+1} = v \mid X_{t+1} = y,\ X_t = x)\mathbb{P}(X_{t+1} = y \mid X_t = x)\mathbb{P}(X_t = x) M_{x,u} \\
&= \sum_{x,y \in \mathbb{S}} \pi_x P_{x,y} M_{x,u} \mathbb{P}(O_{t+1} = v \mid X_{t+1} = y) \\
&= \sum_{x,y \in \mathbb{S}} \pi_x P_{x,y} M_{x,u} M_{y,v}, \quad u, v \in \mathcal{O}.
\end{aligned}$$

b) We have

$$\begin{aligned}\mathbb{P}(O_{t+1} \in \mathcal{B},\ O_t \in \mathcal{A}) &= \sum_{u\in\mathcal{A}} \sum_{v\in\mathcal{B}} \mathbb{P}(O_{t+1} = v,\ O_t = u)\\ &= \sum_{x,y\in\mathbb{S}} \pi_x P_{x,y} \sum_{v\in\mathcal{B}} M_{y,v} \sum_{u\in\mathcal{A}} M_{x,u}.\end{aligned}$$

c) We find

$$\begin{aligned}\mathbb{P}(O_t \in \mathcal{A}) &= \sum_{u\in\mathcal{A}} \sum_{v\in\mathcal{O}} \mathbb{P}(O_{t+1} = v,\ O_t = u)\\ &= \sum_{u\in\mathcal{A}} \sum_{v\in\mathcal{O}} \sum_{x,y\in\mathbb{S}} M_{y,v} \pi_x P_{x,y} M_{x,u}\\ &= \sum_{x\in\mathbb{S}} \pi_x \sum_{u\in\mathcal{A}} M_{x,u},\end{aligned}$$

and

$$\mathbb{P}(O_{t+1} \in \mathcal{B} \mid O_t \in \mathcal{A}) = \frac{\mathbb{P}(O_{t+1} \in \mathcal{B} \mid O_t \in \mathcal{A})}{\mathbb{P}(O_t \in \mathcal{A})}.$$

d) If

$$\begin{bmatrix} \sum_{u\in\mathcal{A}} M_{0,u} & \sum_{v\in\mathcal{B}} M_{0,v} \\ \sum_{u\in\mathcal{A}} M_{1,u} & \sum_{v\in\mathcal{B}} M_{1,v} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

then

$$\mathbb{P}(O_{t+1} \in \mathcal{A},\ O_t \in \mathcal{A}) = \sum_{u,v\in\mathcal{A}} \sum_{x,y\in\mathbb{S}} M_{y,v} \pi_x P_{x,y} M_{x,u} = \pi_0 P_{0,0},$$

and similarly

$$\mathbb{P}(O_t \in \mathcal{A}) = \sum_{u\in\mathcal{A}} \sum_{v\in\mathcal{O}} \sum_{x,y\in\mathbb{S}} M_{y,v} \pi_x P_{x,y} M_{x,u} = \pi_0,$$

hence $\mathbb{P}(O_{t+1} \in \mathcal{A},\ O_t \in \mathcal{A}) = P_{0,0}$, and more generally,

$$\begin{bmatrix} \mathbb{P}(O_{t+1} \in \mathcal{A} \mid O_t \in \mathcal{A}) & \mathbb{P}(O_{t+1} \in \mathcal{A} \mid O_t \in \mathcal{B}) \\ \mathbb{P}(O_{t+1} \in \mathcal{B} \mid O_t \in \mathcal{A}) & \mathbb{P}(O_{t+1} \in \mathcal{B} \mid O_t \in \mathcal{B}) \end{bmatrix} = \begin{bmatrix} P_{0,0} & P_{0,1} \\ P_{1,0} & P_{1,1} \end{bmatrix}.$$

e) We have

$$\begin{aligned}[\pi_0, \pi_1] &= \left[\frac{0.6842348}{0.8564253 + 0.6842348}, \frac{0.8564253}{0.8564253 + 0.6842348}\right]\\ &= [0.444117947, 0.555882053].\end{aligned}$$

f) We have

$$
\begin{aligned}
\widehat{\mathbb{P}}(O_{t+1} \in \mathcal{A},\ O_t \in \mathcal{A}) &= \sum_{x,y \in \mathbb{S}} \pi_x P_{x,y} \sum_{v \in \mathcal{A}} \widehat{M}_{y,v} \sum_{u \in \mathcal{A}} \widehat{M}_{x,u} \\
&= \pi_0 P_{0,0} \sum_{v \in \mathcal{A}} \widehat{M}_{0,v} \sum_{u \in \mathcal{A}} \widehat{M}_{0,u} + \pi_0 P_{0,1} \sum_{v \in \mathcal{A}} \widehat{M}_{1,v} \sum_{u \in \mathcal{A}} \widehat{M}_{0,u} \\
&\quad + \pi_1 P_{1,0} \sum_{v \in \mathcal{A}} \widehat{M}_{0,v} \sum_{u \in \mathcal{A}} \widehat{M}_{1,u} + \pi_1 P_{1,1} \sum_{v \in \mathcal{A}} \widehat{M}_{1,v} \sum_{u \in \mathcal{A}} \widehat{M}_{1,u} \\
&= 0.444117947 \times 0.1435747 \times 0.53605372 \times 0.53605372 \\
&\quad + 0.444117947 \times 0.8564253 \times 0.02345197 \times 0.53605372 \\
&\quad + 0.555882053 \times 0.6842348 \times 0.53605372 \times 0.02345197 \\
&\quad + 0.555882053 \times 0.3157652 \times 0.02345197 \times 0.02345197 \\
&= 0.027982632,
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{P}(O_t \in \mathcal{A}) &= \sum_{x \in \mathbb{S}} \pi_x \sum_{u \in \mathcal{A}} M_{x,u} \\
&= \pi_0 \sum_{u \in \mathcal{A}} M_{0,u} + \pi_1 \sum_{u \in \mathcal{A}} M_{1,u} \\
&= 0.444117947 \times 0.53605372 + 0.555882053 \times 0.02345197 \\
&= 0.251107607,
\end{aligned}
$$

hence

$$
\widehat{\mathbb{P}}(O_{t+1} \in \mathcal{A} \mid O_t \in \mathcal{A}) = \frac{0.027982632}{0.251107607} = 0.1114368,
$$

and more generally,

$$
\begin{bmatrix}
\widehat{\mathbb{P}}(O_{t+1} \in \mathcal{A} \mid O_t \in \mathcal{A}) & \widehat{\mathbb{P}}(O_{t+1} \in \mathcal{A} \mid O_t \in \mathcal{B}) \\
\widehat{\mathbb{P}}(O_{t+1} \in \mathcal{B} \mid O_t \in \mathcal{A}) & \widehat{\mathbb{P}}(O_{t+1} \in \mathcal{B} \mid O_t \in \mathcal{B})
\end{bmatrix}
= \begin{bmatrix}
0.1114368 & 0.8885632 \\
0.2957185 & 0.7042815
\end{bmatrix}. \tag{S.9.8}
$$

g) We find that (S.9.8) is a close approximation of (9.20).

Problem 9.3 (Wolfer and Kontorovich (2021))

a) For all $i = 1, \ldots, d$ we have

$$
\mathbb{E}\left[\sum_{j=1}^{d} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j} \right| \right] = \frac{1}{n} \sum_{j=1}^{d} \mathbb{E}\left[\left| \sum_{k=1}^{n} \mathbf{1}_{\{Z_i(k)=j\}} - nP_{i,j} \right| \right]
$$

$$
\begin{aligned}
&\leq \frac{1}{n} \sum_{j=1}^{d} \sqrt{\mathbb{E}\left[\left|\sum_{k=1}^{n} \left(\mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j}\right)\right|^2\right]} \\
&= \frac{1}{n} \sum_{j=1}^{d} \sqrt{\operatorname{Var}\left[\sum_{k=1}^{n} \mathbf{1}_{\{Z_i(k)=j\}}\right]} \\
&= \frac{1}{n} \sum_{j=1}^{d} \sqrt{n(1-P_{i,j})P_{i,j}} \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^{d} \sqrt{P_{i,j}} \\
&\leq \sqrt{\frac{d}{n}} \sqrt{\sum_{j=1}^{d} P_{i,j}} \\
&= \sqrt{\frac{d}{n}}, \qquad n \geq 1,
\end{aligned}
$$

where we used the Cauchy-Schwarz inequality.

b) Using the inequality $||u| - |v|| \leq |u - v|$, $u, v \in \mathbb{R}$, we have

$$
\begin{aligned}
&\left| \sum_{j=1}^{d} \left| \frac{1}{n} \mathbf{1}_{\{x=j\}} + \frac{1}{n} \sum_{k=1,\ k\neq i}^{n} \mathbf{1}_{\{z(k)=j\}} - P_{i,j} \right| \right. \\
&\qquad \left. - \sum_{j=1}^{d} \left| \frac{1}{n} \mathbf{1}_{\{y=j\}} + \frac{1}{n} \sum_{k=1,\ k\neq i}^{n} \mathbf{1}_{\{z(k)=j\}} - P_{i,j} \right| \right| \\
&\leq \frac{1}{n} \sum_{j=1}^{d} \left| \mathbf{1}_{\{x=j\}} + \sum_{k=1,\ k\neq i}^{n} \mathbf{1}_{\{z(k)=j\}} - P_{i,j} \right. \\
&\qquad \left. - \left(\sum_{j=1}^{d} \mathbf{1}_{\{y=j\}} + \sum_{k=1,\ k\neq i}^{n} \mathbf{1}_{\{z(k)=j\}} - P_{i,j} \right) \right| \\
&= \frac{1}{n} \sum_{j=1}^{d} \left| \mathbf{1}_{\{x=j\}} - \mathbf{1}_{\{y=j\}} \right| \\
&\leq \frac{2}{n} := c_i, \qquad i = 1, \ldots, n.
\end{aligned}
$$

c) Using McDiarmid's inequality, for all $i = 1, \ldots, d$ we have

$$
\begin{aligned}
&\mathbb{P}\left(\sum_{j=1}^d \left|\frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j}\right| > \varepsilon\right) \\
&= \mathbb{P}\left(\sum_{j=1}^d \left|\frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j}\right| - \mathbb{E}\left[\sum_{j=1}^d \left|\frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j}\right|\right]\right. \\
&\qquad\qquad\qquad\qquad \left. > \varepsilon - \mathbb{E}\left[\sum_{j=1}^d \left|\frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j}\right|\right]\right) \\
&\leq \mathbb{P}\left(\sum_{j=1}^d \left|\frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j}\right| - \mathbb{E}\left[\sum_{j=1}^d \left|\frac{1}{n}\sum_{k=1}^n \mathbf{1}_{\{Z_i(k)=j\}} - P_{i,j}\right|\right] > \varepsilon - \sqrt{\frac{d}{n}}\right) \\
&\leq \exp\left(-\frac{2}{\sum_{i=1}^d c_i^2} \max\left(0, \varepsilon - \sqrt{\frac{d}{n}}\right)^2\right) \\
&= \exp\left(-\frac{n}{2} \max\left(0, \varepsilon - \sqrt{\frac{d}{n}}\right)^2\right).
\end{aligned}
$$

d) When $\widetilde{N}_i(m) = n \geq 1$, we have

$$
\begin{aligned}
\widetilde{P}_{i,j}(m) &:= \frac{1}{\widetilde{N}_i(m)} \sum_{k=1}^{m-1} \mathbf{1}_{\{\widetilde{X}_k = i, \widetilde{X}_{k+1} = j\}} \\
&= \frac{1}{n} \sum_{k=1}^{m-1} \mathbf{1}_{\{\widetilde{X}_k = i, Z_{\widetilde{X}_k}(1+\widetilde{N}_{\widetilde{X}_k}(k)) = j\}} \\
&= \frac{1}{n} \sum_{k=1}^{m-1} \mathbf{1}_{\{\widetilde{X}_k = i, Z_i(1+\widetilde{N}_i(k)) = j\}} \\
&= \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{Z_i(k) = j\}} \qquad i, j = 1, \ldots, d.
\end{aligned}
$$

e) This follows from the fact that $\widetilde{X}_{k+1}$ has the same distribution as Z_i given that $\widetilde{X}_k = i$.

f) Letting $n_i := \lceil m\pi_i/2 \rceil$, $i = 1, \ldots, d$, letting $c_1 := (1 - 1/\sqrt{2})^2$ we have

$$
0 \leq \varepsilon - \sqrt{\frac{d}{n}} \leq \varepsilon\sqrt{c_1}, \qquad n \geq n_i \geq 2d/\varepsilon^2,
$$

hence

$$
\begin{aligned}
&\sum_{n=n_i}^{3n_i} \mathbb{P}\left(\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m) - P_{i,j}\right| > \varepsilon \text{ and } N_i(m) = n\right) \\
&\quad = \sum_{n=n_i}^{3n_i} \mathbb{P}\left(\sum_{j=1}^{d}\left|\widetilde{P}_{i,j}(m) - P_{i,j}\right| > \varepsilon \text{ and } \widetilde{N}_i(m) = n\right) \\
&\quad = \sum_{n=n_i}^{3n_i} \exp\left(-\frac{n}{2}\max\left(0, \varepsilon - \sqrt{\frac{d}{n}}\right)^2\right) \\
&\quad \le \sum_{n=n_i}^{3n_i} \mathrm{e}^{-2nc_1\varepsilon^2} \\
&\quad \le (2n_i + 1)\mathrm{e}^{-2n_i c_1 \varepsilon^2} \\
&\quad \le (2n_i + 1)\mathrm{e}^{-m\pi_i c_1 \varepsilon^2},
\end{aligned}
$$

provided that $n_i \ge 2d/\varepsilon^2$, or $m \ge 4d/(\varepsilon^2 \pi_i)$.

g) We have

$$
\begin{aligned}
&\sum_{i=1}^{d}\sum_{n=n_i}^{3n_i} \mathbb{P}\left(\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m) - P_{i,j}\right| > \varepsilon \text{ and } N_i(m) = n\right) \\
&\quad \le \sum_{i=1}^{d}(2n_i + 1)e^{-c_1 m\pi_i\varepsilon^2} \\
&\quad \le \sum_{i=1}^{d}\frac{2n_i + 1}{c_1 m\pi_i\varepsilon^2}e^{-c_1 m\pi_i\varepsilon^2/2} \\
&\quad \le \sum_{i=1}^{d}\frac{2\lceil m\pi_i/2\rceil + 1}{c_1 m\pi_i\varepsilon^2}e^{-c_1 m\pi_i\varepsilon^2/2} \\
&\quad \le \sum_{i=1}^{d}\frac{m + 3/\pi_i}{c_1 m\varepsilon^2}e^{-c_1 m\pi_i\varepsilon^2/2} \\
&\quad \le \frac{1}{c_1\varepsilon^2}\sum_{i=1}^{d}\left(1 + \frac{3}{m\pi_*}\right)e^{-c_1 m\pi_i\varepsilon^2/2} \\
&\quad = \frac{d}{c_1\varepsilon^2}\left(1 + \frac{3}{m\pi_*}\right)e^{-c_1 m\pi_i\varepsilon^2/2} \\
&\quad \le \frac{d}{c_1\varepsilon^2}\left(1 + \frac{3\varepsilon^2}{4d}\right)e^{-c_1 m\pi_i\varepsilon^2/2} \\
&\quad \le \frac{2d}{c_1\varepsilon^2}e^{-c_1 m\pi_*\varepsilon^2/2},
\end{aligned}
$$

provided that $m \geq 4d/(\varepsilon^2\pi_*)$ and $\varepsilon \in (0,1)$.

h) For all $\varepsilon > 0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i=1,\ldots,d}\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m)-P_{i,j}\right| > \varepsilon\right)\\
&= \mathbb{P}\left(\max_{i=1,\ldots,d}\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m)-P_{i,j}\right| > \varepsilon \text{ and } \bigcap_{j=1}^{d}\{N_i(m)\in[n_i,3n_i]\}\right)\\
&\quad+\mathbb{P}\left(\max_{i=1,\ldots,d}\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m)-P_{i,j}\right| > \varepsilon \text{ and } \bigcup_{j=1}^{d}\{N_i(m)\notin[n_i,3n_i]\}\right)\\
&\leq \mathbb{P}\left(\bigcup_{i=1,\ldots,d}\left\{\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m)-P_{i,j}\right| > \varepsilon \text{ and } N_i(m)\in[n_i,3n_i]\right\}\right)\\
&\quad+\mathbb{P}\left(\bigcup_{j=1}^{d}\{N_i(m)\notin[n_i,3n_i]\}\right)\\
&\leq \sum_{i=1}^{d}\mathbb{P}\left(\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m)-P_{i,j}\right| > \varepsilon \text{ and } N_i(m)\in[n_i,3n_i]\right)\\
&\quad+\mathbb{P}\left(\bigcup_{j=1}^{d}\{N_i(m)\notin[n_i,3n_i]\}\right)\\
&= \sum_{i=1}^{d}\sum_{n=n_i}^{3n_i}\mathbb{P}\left(\sum_{j=1}^{d}\left|\widehat{P}_{i,j}(m)-P_{i,j}\right| > \varepsilon \text{ and } N_i(m)=n\right)\\
&\quad+\mathbb{P}(\exists i\in\{1,\ldots,d\}\ :\ N_i(m)\notin[n_i,3n_i]).
\end{aligned}$$

i) Letting $f_i(x) := \mathbf{1}_{\{x=i\}} - \pi_i$, $i = 1,\ldots,d$, we have

$$N_i(m)-(m-1)\pi_i = \sum_{k=1}^{m-1} f_i(X_k)$$

and

$$\mathbb{E}[f_i(X_k)] = \mathbb{E}[N_i(m)-(m-1)\pi_i] = \mathbb{E}\left[\sum_{k=1}^{m-1} f_i(X_k)\right] = (m-1)\pi_i = 0,$$

hence by the bound in Question (l) of Problem 6.14, we have

$\mathbb{P}(\exists i\in\{1,\ldots,d\}\ :\ N_i(m)\notin[n_i,3n_i])$

$$
\begin{aligned}
&= \mathbb{P}(\exists i \in \{1,\dots,d\} \ : \ N_i(m) > 3n_i) + \mathbb{P}(\exists i \in \{1,\dots,d\} \ : \ N_i(m) < n_i) \\
&\leq \mathbb{P}(\exists i \in \{1,\dots,d\} \ : \ N_i(m) > 3(m-1)\pi_i/2) \\
&\quad + \mathbb{P}(\exists i \in \{1,\dots,d\} \ : \ N_i(m) < 2 + (m-1)\pi_i/2) \\
&= \mathbb{P}\left(\exists i \in \{1,\dots,d\} \ : \ \frac{1}{m-1}\sum_{k=1}^{m-1} f_i(X_k) > \frac{\pi_i}{2}\right) \\
&\quad + \mathbb{P}\left(\exists i \in \{1,\dots,d\} \ : \ \frac{1}{m-1}\sum_{k=1}^{m-1} f_i(X_k) < -\frac{\pi_i}{2} + \frac{2}{m-1}\right) \\
&\leq \mathbb{P}\left(\max_{i=1,\dots,d} \frac{1}{m-1}\sum_{k=1}^{m-1} f_i(X_k) > \frac{\pi_i}{2}\right) \\
&\quad + \mathbb{P}\left(\max_{i=1,\dots,d} \frac{1}{m-1}\sum_{k=1}^{m-1} (-f_i(X_k)) > \frac{\pi_i}{2} - \frac{2}{m-1}\right) \\
&\leq e^{(1-\lambda_1)/5} e^{-(1-\lambda_1)m\pi_i^2/48} + e^{(1-\lambda_1)/5} e^{-(1-\lambda_1)m(\pi_i/2-2/(m-1))^2/12} \\
&\leq c_2 d e^{-c_3 m(1-\lambda_1)\pi_*^2}, \qquad m \geq 2,
\end{aligned}
$$

where $c_2 = 2e^{(1-\lambda_1)/5}$ and

$$
c_3 = \max\left(\frac{1}{48}, \frac{1}{12}\left(1 - \frac{4}{\pi_*(m-1)}\right)\right) \leq \frac{5}{12},
$$

provided that $m \geq 1 + 4/\pi_*$.

j) We upper bound

$$
\begin{aligned}
&\sum_{i=1}^{d}\sum_{n=n_i}^{3n_i} \mathbb{P}\left(\sum_{j=1}^{d} \left|\widehat{P}_{i,j}(m) - P_{i,j}\right| > \varepsilon \text{ and } N_i(m) = n\right) \\
&\qquad \leq \frac{2d}{c_1\varepsilon^2} e^{-c_1 m\pi_*\varepsilon^2/2} \\
&\qquad < \frac{\delta}{2}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{P}(\exists i \in \{1,\dots,d\} \ : \ N_i(m) \notin [n_i, 3n_i]) &\leq c_2 d e^{-c_3 m(1-\lambda_1)\pi_*^2} \\
&< \frac{\delta}{2},
\end{aligned}
$$

which yields

$$
m > \frac{2}{c_1\pi_*\varepsilon^2} \log \frac{4d}{\delta c_1 \varepsilon^2}
$$

and

$$m > \frac{1}{c_3(1-\lambda_1)\pi_*^2} \log \frac{2c_2 d}{\delta},$$

hence, using the facts that $d \geq 2$ and $y + \log x < 2 \log x$, $x > e^y$, we find that there is a constant $c > 0$ such that for all

$$m \geq c \max\left(\frac{1}{\varepsilon^2 \pi_*} \max\left(d, \log \frac{d}{\delta\varepsilon}\right), \frac{1}{(1-\lambda_1)\pi_*^2} \log \frac{d}{\delta}\right),$$

we have

$$\mathbb{P}\left(\max_{i=1,\ldots,d} \sum_{j=1}^{d} \left| \widehat{P}_{i,j}(m) - P_{i,j} \right| \leq \varepsilon \right) \geq 1 - \delta.$$

For example, taking $\varepsilon = \delta = 5\%$ and $\pi_* = 1/d$ with $d = 26$ we find $m \gtrapprox 62300$.

Chapter 10 - Markov Decision Processes

Exercise 10.1 By first step analysis, we have

$$\begin{cases} V_a(a) = 0 \\ V_a(b) = -1 + \frac{2}{3} V_a(a) + \frac{1}{3} V_a(c) \\ V_a(c) = 2 + V_a(b), \end{cases}$$

which has for solution $V_a(a) = 0$, $V_a(b) = -1/2$, $V_a(c) = 3/2$, as confirmed by the following R code.

```
library("igraph"); library(markovchain); statenames <- c("a", "b", "c")
P<-matrix(c(1,0,0,2/3,0,1/3,0,1,0),nrow=3,byrow=TRUE, dimnames =
      list(statenames,statenames));
MC <-new("markovchain",transitionMatrix=P); graph <- as(MC, "igraph")
plot(graph,vertex.size=50,edge.label.cex=2,
      edge.label=E(graph)$prob,edge.color='black', vertex.color='dodgerblue',
      vertex.label.cex=3)
expectedRewards(MC,100,c(0,-1,2))
0.0 -0.5 1.5
meanAbsorptionTime(object = MC)
  b c
a 2 3
```

Exercise 10.2 By first step analysis, we have

$$
\begin{cases}
V(1) = -2 + (1-p)\gamma V(1) + p\gamma V(2) \\
V(2) = 3 + (1-q)\gamma V(1) + q\gamma V(3) \\
V(3) = 1 + \gamma V(3)
\end{cases}
$$

hence

$$
\begin{cases}
V(1) = -2 + (1-p)\gamma V(1) + p\gamma V(2) \\
V(2) = 3 + (1-q)\gamma V(1) + \dfrac{q\gamma}{1-\gamma} \\
V(3) = \dfrac{1}{1-\gamma} = \displaystyle\sum_{n\geq 0} \gamma^n,
\end{cases}
$$

and

$$
\begin{cases}
V(1) = \dfrac{(3p\gamma - 2)(1-\gamma) + pq\gamma^2}{(1-(1-p)\gamma - (1-q)p\gamma^2)(1-\gamma)} \\
V(2) = 3 + \dfrac{q\gamma}{1-\gamma} + \dfrac{(1-q)((3p\gamma-2)(\gamma-\gamma^2) + pq\gamma^3)}{(1-(1-p)\gamma-(1-q)p\gamma^2)(1-\gamma)} \\
V(3) = \dfrac{1}{1-\gamma}.
\end{cases}
$$

In particular, when $p = q = 1$ we check that

$$
\begin{cases}
V(1) = -2 + 3\gamma + \dfrac{\gamma^2}{1-\gamma} \\
V(2) = 3 + \dfrac{\gamma}{1-\gamma}, \\
\\
V(3) = \dfrac{1}{1-\gamma} = \displaystyle\sum_{n\geq 0} \gamma^n.
\end{cases}
$$

Exercise 10.3

a) We have

$$
\begin{aligned}
h(k) &= \mathbb{E}\left[\sum_{i\geq 0} \beta^i c(X_i) \,\middle|\, X_0 = k \right] \\
&= \mathbb{E}[c(X_0) \mid X_0 = k] + \mathbb{E}\left[\sum_{i\geq 1} \beta^i c(X_i) \,\middle|\, X_0 = k \right] \\
&= c(k) + \sum_{j\in S} P_{k,j}\, \mathbb{E}\left[\sum_{i\geq 1} \beta^i c(X_i) \,\middle|\, X_1 = j \right]
\end{aligned}
$$

$$
\begin{aligned}
&= c(k) + \beta \sum_{j\in S} P_{k,j}\, \mathbb{E}\left[\sum_{i\geq 0} \beta^i c(X_i) \;\Big|\; X_0 = j \right] \\
&= c(k) + \beta \sum_{j\in S} P_{k,j} h(j), \qquad k \in S.
\end{aligned}
$$

This type of equation may be difficult to solve in full generality.

b) The chain has the following graph:

$$
\begin{tikzcd}
0 \arrow[r, bend left, "0.5"] \arrow[d, bend right, "0.5"'] & 1 \arrow[l, bend left, "0.5"] \arrow[loop right, "0.5"] \\
2 \arrow[loop right, "1"] &
\end{tikzcd}
$$

The average utility $h(k)$ solves the first step analysis equations

$$
\begin{cases}
h(0) = c(0) + \dfrac{1}{2} h(1) = 5 + \dfrac{1}{2} h(1) \\
\\
h(1) = c(1) + \dfrac{1}{2} h(0) + \dfrac{1}{2} h(1) = -2 + \dfrac{1}{2} h(0) + \dfrac{1}{2} h(1) \\
\\
h(2) = 0,
\end{cases}
$$

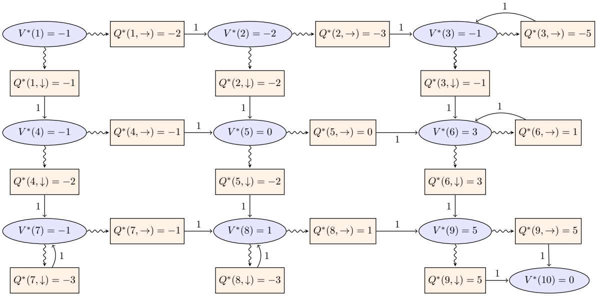
which yields

$$
h(0) = 6, \quad h(1) = 2, \quad h(2) = 0.
$$

See also Problem 5.22 in Privault (2018) for a related problem with explicit solution.

Exercise 10.4

a) The optimal action-value functional $Q^*(k, a)$ is obtained as follows:



b) The optimal value function $V^*(k)$, $k = 1, 2, \ldots, 9$, is given in the next table.

① $V^*(1) = -1$	② $V^*(2) = -2$	③ $V^*(3) = -1$
④ $V^*(4) = -1$	⑤ $V^*(5) = 0$	⑥ $V^*(6) = +3$
⑦ $V^*(7) = -1$	⑧ $V^*(8) = 1$	⑨ $V^*(9) = +5$

c) The optimal policy $\pi^*(k) \in \{\rightarrow, \downarrow\}$, $k = 1, 2, \ldots, 9$, is given as follows.

① $\pi^*(1) = \downarrow$	② $\pi^*(2) = \downarrow$	③ $\pi^*(3) = \downarrow$
④ $\pi^*(4) = \rightarrow$	⑤ $\pi^*(5) = \rightarrow$	⑥ $\pi^*(6) = \downarrow$
⑦ $\pi^*(7) = \rightarrow$	⑧ $\pi^*(8) = \rightarrow$	⑨ $\pi^*(9) =$ ↱

References

[1] V. Anantharam, P. Varaiya, and J. Walrand. Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays - Part II: Markovian rewards. *IEEE Transactions on Automatic Control*, 32(11): 977–982, 1987. (Cited on page 53.)

[2] T. Antal and S. Redner. The excited random walk in one dimension. *J. Phys. A*, 38(12):2555–2577, 2005. (Cited on pages 26, 28, and 38.)

[3] A.N. Borodin. *Stochastic processes.* Probability and its Applications. Birkhäuser/Springer, Cham, 2017. Original Russian edition published by LAN Publishing, St. Petersburg, 2013. (Cited on page 26.)

[4] S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and non-stochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012. (Cited on page 56.)

[5] S. Chewi. Wald's identity. https://inst.eecs.berkeley.edu/~ee126/fa17/wald.pdf, 2017. Accessed: 2022-08-25. (Cited on page 52.)

[6] S. Foucart. Linear algebra and matrix analysis, Lecture 6. https://www.math.drexel.edu/~foucart/TeachingFiles/F12/M504Lect6.pdf, 2010. Accessed: 2022-08-25. (Cited on page 49.)

[7] D. Freedman. *Markov chains.* Springer-Verlag, New York-Berlin, 1983. (Cited on page 52.)

[8] T. Kato. *Perturbation Theory for Linear Operators.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition. (Cited on pages 67 and 68.)

[9] C.A. Léon and F. Perron. Optimal Hoeffding bounds for discrete reversible Markov chains. *The Annals of Applied Probability*, pages 958–970, 2004. (Cited on page 50.)

[10] D.A. Levin, Y. Peres, and E.L. Wilmer. *Markov Chains and Mixing Times.* American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson. (Cited on pages 59 and 64.)

[11] P. Lezaud. Chernoff-type bound for finite Markov chains. *Ann. Appl. Probab.*, 8(3):849–867, 1998. (Cited on page 65.)

[12] N. Privault. *Understanding Markov Chains.* Springer Undergraduate Mathematics Series. Springer, second edition, 2018. (Cited on pages 2, 40, 73, and 88.)

[13] S. Redner. *A guide to first-passage processes.* Cambridge University Press, Cambridge, 2001. (Cited on pages 26 and 28.)

[14] R. Serfozo. *Basics of applied stochastic processes.* Probability and its Applications (New York). Springer-Verlag, Berlin, 2009. (Cited on pages 45 and 52.)

[15] J.M. Steele. *Stochastic Calculus and Financial Applications*, volume 45 of *Applications of Mathematics.* Springer-Verlag, New York, 2001. (Cited on page 2.)

[16] R. J. Williams. Asymptotic variance parameters for the boundary local times of reflected Brownian motion on a compact interval. *J. Appl. Probab.*, 29(4):996–1002, 1992. (Cited on pages 26 and 27.)

[17] G. Wolfer and A. Kontorovich. Statistical estimation of ergodic Markov chain kernel over discrete state space. *Bernoulli*, 27(1):532–553, 2021. (Cited on page 80.)