Math 525: Assignment 8 Solutions

1.

(a) From class, $Pe = e$. Therefore, $\mu^\intercal P e = \mu^\intercal e$.

(b) First, note that $\mu^\intercal P^k = ((P^k)^\intercal \mu)^\intercal = ((P^\intercal)^k \mu)^\intercal$. Moreover,

$$\begin{aligned}
(P^\intercal)^k \mu &= (P^\intercal)^{k-1} P^\intercal \mu \\
&= (P^\intercal)^{k-1} P^\intercal (c_1 v_1 + \cdots + c_m v_m) \\
&= (P^\intercal)^{k-1} (c_1 \lambda_1 v_1 + \cdots + c_m \lambda_m v_m) \\
&= \cdots \\
&= c_1 \lambda_1^k v_1 + \cdots + c_m \lambda_m^k v_m.
\end{aligned}$$

Therefore, $\mu^\intercal P^k = c_1 \lambda_1^k v_1^\intercal + \cdots + c_m \lambda_m^k v_m^\intercal$.

(c) We know $\rho(P) \leq 1$ and $\lambda_1 = 1$. Therefore, $|\lambda_j| < 1$ whenever $j \neq 1$ and hence

$$\lim_{n\to\infty} \lambda_1^n = 1 \qquad \text{and} \qquad \lim_{n\to\infty} \lambda_j^n = 0 \text{ if } j \neq 1.$$

It follows that

$$\begin{aligned}
\lim_{n\to\infty} \mu^\intercal P^n &= \lim_{n\to\infty} \{c_1 \lambda_1^n v_1^\intercal + \cdots + c_m \lambda_m^n v_m^\intercal\} \\
&= c_1 v_1^\intercal \lim_{n\to\infty} \lambda_1^n + \cdots + c_m v_m^\intercal \lim_{n\to\infty} \lambda_m^n = c_1 v_1^\intercal.
\end{aligned}$$

(d) This follows from the fact that products are continuous:

$$(c_1 v_1^\intercal) P = \left(\lim_{n\to\infty} \mu^\intercal P^n \right) P = \lim_{n\to\infty} \mu^\intercal P^{n+1} = \lim_{n\to\infty} \mu^\intercal P^n = c_1 v_1^\intercal.$$

(e) By part (a), we know that $(\mu^\intercal P^n) e = 1$ for any n. Therefore,

$$1 = \lim_{n\to\infty} ((\mu^\intercal P^n) e) = \left(\lim_{n\to\infty} \mu^\intercal P^n \right) e = (c_1 v_1^\intercal) e.$$

Since v_1 is a positive vector, c_1 must be positive, since otherwise $(c_1 v_1^\intercal) e \leq 0$.

(f) Part (c) is akin to saying $c_1 v_1^\intercal$ is the limiting distribution of the Markov chain and part (d) is akin to saying that the limiting distribution is an equilibrium/stationary distribution of the Markov chain.

2.

(a) Simplifying completely, the answer is

$$\mathbb{P}(\tau \leq n) = \begin{cases} \dfrac{(1-p)^{n+1}}{2p-1} - \dfrac{p^{n+1}}{2p-1} + 1 & \text{if } p \neq 1/2 \\ 2^{-n}\left(1 + 2^n - n\right) & \text{if } p = 1/2. \end{cases} \tag{1}$$

Below are two methods to obtain this solution.

i. The first method uses four states.

- Let $S = \{HH, HT, TH, TT\}$. The interpretation is as follows:
  - HH: the most recent coin flip was H, immediately preceded by H.
  - HT: the most recent coin flip was T, immediately preceded by H.
  - TH: the most recent coin flip was H, immediately preceded by T.
  - TT: the most recent coin flip was T, immediately preceded by T.
- The transitions between these states is given by the matrix

$$\tilde{P} = \begin{pmatrix} p & 1-p & & \\ & & p & 1-p \\ p & 1-p & & \\ & & p & 1-p \end{pmatrix}.$$

- In the context of our problem, this isn't quite the transition matrix we want! That's because as soon as we encounter HT, the "game" should terminate. Therefore, we modify it to

$$P = \begin{pmatrix} p & 1-p & & \\ & 1 & & \\ p & 1-p & & \\ & & p & 1-p \end{pmatrix}.$$

- The initial distribution (i.e., the distribution of X_0) is given by

$$\mu^\intercal = \begin{pmatrix} p^2 & p\,(1-p) & p\,(1-p) & (1-p)^2 \end{pmatrix}.$$

- The terminal distribution we are interested in is

$$\nu^\intercal = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}.$$

- Putting this all together, we get $\mathbb{P}(\tau \leq n) = \mu^\intercal P^{n-2} \nu$. The reason we are using $n-2$ instead of n here is that in the Markov chain, X_0 corresponds to after we have already seen the first two coin flips. You can (optionally) simplify this to get (1).

ii. The second method uses only three states. The reason we can do this is that we are actually maintaining some redundant information in method (i).

- Let $S = \{\epsilon, H, HT\}$. The interpretation is as follows:

- ϵ: we have not flipped any coins yet.
- H: the most recent coin flip was H (it is understood that if we are in this state, we have not yet seen HT).
- HT: the most recent coin flip was T, immediately preceded by H.

- The transition matrix is

$$P = \begin{pmatrix} 1-p & p & \\ & p & 1-p \\ & & 1 \end{pmatrix}$$

(draw a picture of the graph to get a better sense of what is going on).

- The initial distribution (i.e., the distribution of X_0) is given by

$$\mu^{\mathsf{T}} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

- The terminal distribution we are interested in is

$$\nu^{\mathsf{T}} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}.$$

- Putting this all together, $\mathbb{P}(\tau \leq n) = \mu^{\mathsf{T}} P^n \nu$. You can (optionally) simplify this to get (1).

(b) Once we have an expression for $\mathbb{P}(\tau \leq n)$, we can obtain an expression for $\mathbb{P}(\tau = n)$ as follows. Since $\mathbb{P}(\tau = 1) = 0$, we can assume $n > 1$. Then,

$$\mathbb{P}(\tau = n) = \mathbb{P}(\tau \leq n) - \mathbb{P}(\tau < n) = \mathbb{P}(\tau \leq n) - \mathbb{P}(\tau \leq n-1)$$

(you can simplify the above further if you want to obtain an expression like (1)).

(c) Plugging in $n = 10$ and $p = 1/3$ into (1), we get

$$\mathbb{P}\left(\tau \leq 10 \middle| p = \frac{1}{3}\right) = \frac{57002}{59049} \approx 0.96533.$$