ECE59500RL Exam

Robert (Cars) Chandler — chandl71@purdue.edu

Problem 1

1.1

The set of trajectories described is the set of all trajectories such that we start at $X_0 = 1$ and all following states are either 5, 6, or 7. Because the transitions between 5, 6, and 7 are periodic, there is no way to return back to 1 once 5 is reached. As a result, we only have to analyze the first step in the trajectory. If the first step takes us to 5, then the trajectory must belong to the set described. The probability of transitioning to state 5 from the initial state, 1, is labeled as 0.5. If we let $\mathcal{T}$ be the set of trajectories described in the problem statement, then:

$$\mathbb{P}(X_1 = 5 | X_0 = 1) = \mathbb{P}(\tau \in \mathcal{T} | X_0 = 1) = 0.5$$

Where τ is any trajectory resulting from $X_0 = 1$.

Final answer: 0.5

1.2

We can solve this by finding the left eigenvectors of the $\mathbf{P}$ matrix:

```
import numpy as np

p = np.array(
    [
        [0, 0.1, 0, 0, 0.5, 0, 0, 0.4, 0, 0],
        [0, 0, 1, 0, 0, 0, 0, 0, 0, 0],
        [0, 0, 0, 1, 0, 0, 0, 0, 0, 0],
        [0, 0, 0.6, 0.4, 0, 0, 0, 0, 0, 0],
        [0, 0, 0, 0, 0.75, 0.25, 0, 0, 0, 0],
```

```python
        [0, 0, 0, 0, 0, 0, 1, 0, 0, 0],
        [0, 0, 0, 0, 0.2, 0.8, 0, 0, 0, 0],
        [0, 0, 0, 0, 0, 0, 0, 0, 1, 0],
        [0, 0, 0, 0, 0, 0, 0, 0.9, 0, 0.1],
        [0, 0, 0, 0, 0, 0, 0, 0, 0, 1],
    ]
)

# Using p.T gives us left eigenvectors instead of right
eigvals, eigvecs = np.linalg.eig(p.T)

# Get the indices where the eigenvalues are 1 (accounting for floating point
# math error)
i_ones = np.argwhere(np.abs(eigvals - 1) < 1e-10).squeeze()

# Get the eigenvectors with eigenvalues of 1 (each column of vecs1 is an eigenvector)
vecs1 = eigvecs[:, i_ones]

# The vectors are normalized by default to have a magnitude of 1, but we want
# them to sum to 1
vecs1 /= np.sum(vecs1, axis=0)

# Make each row an eigenvector
# vecs1 = vecs1.T

print("Stationary distributions:")

# Ensure theese eigenvectors display the desired properties
with np.printoptions(precision=4):
    for vec in vecs1.T:
        # Should sum to 1
        assert np.allclose(np.sum(vec), 1)

        # Should yield the same state after applying P
        assert np.allclose(vec @ p, vec)

        print(vec, "\n")
```

```
Stationary distributions:
[0. 0. 0. 0. 0. 0. 0. 0. 0. 1.]

[0.    0.    0.375 0.625 0.    0.    0.    0.    0.    0.   ]
```

```
[0.     0.     0.     0.     0.2857 0.3571 0.3571 0.     0.     0.    ]
```

The vectors above represeent the stationary distributions of the Markov chain. We have proven that they are indeed stationary distributions by proving that they all sum to 1 and that $\bar{\boldsymbol{\mu}}\mathbf{P} = \bar{\boldsymbol{\mu}}$ for each vector.

1.3

If $\alpha\bar{\boldsymbol{\mu}}^1 + (1-\alpha)\bar{\boldsymbol{\mu}}^2$ is a stationary distribution, then it will satisfy the following properties:

$$\alpha\bar{\mu}^1(s) + (1-\alpha)\bar{\mu}^2(s) \geq 0, \quad \forall s \in \mathcal{S}$$
$$\sum_{s\in\mathcal{S}} \alpha\bar{\mu}^1(s) + (1-\alpha)\bar{\mu}^2(s) = 1$$
$$\alpha\bar{\boldsymbol{\mu}}^1 + (1-\alpha)\bar{\boldsymbol{\mu}}^2 = \left(\alpha\bar{\boldsymbol{\mu}}^1 + (1-\alpha)\bar{\boldsymbol{\mu}}^2\right)\mathbf{P}$$

Note that if the quantity above is a convex combination, then both α and $(1-\alpha)$ must be nonnegative, and therefore:

$$0 \leq \alpha \leq 1$$

First we show that for every element $\bar{\mu}_i$ of any stationary distribution $\bar{\boldsymbol{\mu}}$:

$$0 \leq \bar{\mu}_i \leq 1$$

The left side of the inequality holds by the definition of a stationary distribution, and the right side holds because $\sum_{s\in\mathcal{S}} \bar{\mu}_i = 1$ and since each $\bar{\mu}_i$ is positive, the maximum value for any one $\bar{\mu}_i$ is 1.

With this in mind, it becomes apparent that the first of the three properties above holds, because if each element of μ^1, μ^2 is between 0 and 1 and α and $(1-\alpha)$ are also between 0 and 1, then we just have a sum of two products of positive values, the result of which will always be positive.

Next:

$$
\begin{aligned}
&\sum_{s\in\mathcal{S}} \alpha\bar{\mu}^1(s) + (1-\alpha)\bar{\mu}^2(s) \\
&= \alpha \sum_{s\in\mathcal{S}} \bar{\mu}^1(s) + (1-\alpha)\sum_{s\in\mathcal{S}} \bar{\mu}^2(s) && \text{by linearity of summation} \\
&= \alpha \cdot 1 + (1-\alpha)\cdot 1 && \text{since } \sum_{s\in\mathcal{S}} \bar{\mu} = 1 \text{ by definition} \\
&= 1
\end{aligned}
$$

And finally:

$$
\begin{aligned}
\alpha\bar{\boldsymbol{\mu}}^1 + (1-\alpha)\bar{\boldsymbol{\mu}}^2 &= \left(\alpha\bar{\boldsymbol{\mu}}^1 + (1-\alpha)\bar{\boldsymbol{\mu}}^2\right)\mathbf{P} \\
\alpha\bar{\boldsymbol{\mu}}^1 + (1-\alpha)\bar{\boldsymbol{\mu}}^2 &= \alpha\bar{\boldsymbol{\mu}}^1\mathbf{P} + (1-\alpha)\bar{\boldsymbol{\mu}}^2\mathbf{P} \\
\alpha\bar{\boldsymbol{\mu}}^1 - \alpha\bar{\boldsymbol{\mu}}^1\mathbf{P} &= (1-\alpha)\bar{\boldsymbol{\mu}}^2\mathbf{P} - (1-\alpha)\bar{\boldsymbol{\mu}}^2 \\
\alpha\left(\bar{\boldsymbol{\mu}}^1 - \bar{\boldsymbol{\mu}}^1\mathbf{P}\right) &= (1-\alpha)\left(\bar{\boldsymbol{\mu}}^2\mathbf{P} - \bar{\boldsymbol{\mu}}^2\right) \\
\alpha\left(\bar{\boldsymbol{\mu}}^1 - \bar{\boldsymbol{\mu}}^1\right) &= (1-\alpha)\left(\bar{\boldsymbol{\mu}}^2 - \bar{\boldsymbol{\mu}}^2\right) && \text{since } \bar{\boldsymbol{\mu}}\mathbf{P} = \bar{\boldsymbol{\mu}} \text{ by definition} \\
\alpha \cdot 0 &= (1-\alpha)\cdot 0 \\
0 &= 0 && \blacksquare
\end{aligned}
$$

1.4

Initializing our chain at $\alpha\mu_0^1 + (1-\alpha)\mu_0^2$ will cause μ_t to converge to

$$\alpha\bar{\mu}^1 + (1-\alpha)\bar{\mu}^2$$

as $t \to \infty$. Intuitively, the pieces of the initial state distribution that were carried to their respective stationary distribution will still be carried to the same distribution, but in proportion based on the fractions α and $(1-\alpha)$.