Section 3 (STAT 171)

- TF: Jiaze Qiu
  - Email: jiazeqiug.harvard.edu
  - Section/OH: Wed 9-12 pm (ET)
- TF: Fernando Vicente
  - Email: fernando_vicentefas.harvard.edu
  - Section/OH: Mon 7-9 pm and Thu 9-11 am (ET)

- Acknolegement: This handout is partially based on notes created by Lisa Ruan and Christy Huo.

1. Review

- States i and j communicate if i is accessible from j and j is accessible from i.
- A Markov chain is irreducible if it has one communication class.
- Recurrent state: if the Markov chain started in j eventually revisits j: 1) If T is the time it takes to return to j, $P(T < \infty) = 1$; 2) However, it is possible for $\mathbb{E}[T] = \infty$. This is known as null recurrence.
- Transient state: if the Markov chain started in j has a positive probability of never returning to j.
- Periodicity: The period of a state i is the greatest common denominator (gcd) of all integers $n > 0$, for which $P_{ii}(n) > 0$
- An Ergodic Markov chain is irreducible, aperiodic, and all states have finite expected return times. Ergodic Markov chains have unique and positive limiting distributions!
- Lazy Chain: $\tilde{P} = \epsilon I + (1 - \epsilon)P$
- Time Reversibility: Let P be the transition matrix of a Markov chain. If x is a probability distribution which satisfies

  $$x_i P_{ij} = x_j P_{ji}, \forall i, j$$

  then x is the stationary distribution, and the Markov chain is reversible.

2. Excercises

2.1 Problem 1

Consider a Markov chain given by the following transition matrix,

$$\begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.6 & 0.1 & 0.3 \\ 0.2 & 0.6 & 0.2 \end{pmatrix}$$

- Use first-step analysis to find $E\left[T_1 \mid X_0 = 1\right]$, where $T_1 = \min\left\{n > 0 : X_n = 1\right\}$.
- What is the latent assumption(s) for the first-step analysis approach to work?

Solution.

- Let $e_1 = E\left[T_1 \mid X_0 = 1\right], e_2 = E\left[T_1 \mid X_0 = 2\right], e_3 = E\left[T_1 \mid X_0 = 3\right]$. First-step analysis yields

$$\begin{cases} e_1 = 0.4 + 0.3\left(1 + e_2\right) + 0.3\left(1 + e_3\right) \\ e_2 = 0.6 + 0.1\left(1 + e_2\right) + 0.3\left(1 + e_3\right) \\ e_3 = 0.2 + 0.6\left(1 + e_2\right) + 0.2\left(1 + e_3\right) \end{cases}$$

  The solution is

$$\begin{cases} e_1 = \dfrac{22}{9} \\ e_2 = \dfrac{55}{27} \\ e_3 = \dfrac{25}{9} \end{cases}$$

- Latent assumptions:

$$T_1 < \infty, \text{a.s.}$$

  and

$$E[T_1 | X_0 = 1] < \infty$$

2.2 Problem 2

In this problem we will investigate how reducibility of Markov chains interact with sta- tionary distributions and limiting distributions.

Problem: Suppose we are given 2 ergodic markov chains with transition matrices P and Q. We want to combine them into a new Markov chain with transition matrix R :

$$\left[\begin{array}{c|c} (1-\epsilon)*\mathbf{P} & \epsilon*\mathbf{I} \\ \hline 0 & \mathbf{Q} \end{array}\right]$$

- Draw this Markov chain
- How many communication classes are there?
- Does this have a limiting distribution? If so, what would be a reasonable guess in terms of the limiting distributions π_1 and π_2 of Q and P, respectively?

Hint: Try $\pi = (a\pi_1, b\pi_2)$ and solve for a and b.

- How does this relate to absorption states?

In this situation we can think of Q as an absorbing Markov Chain (as opposed to a single absorbing state). In this way we can adopt a macro view of the markov chain having sub-markov chains that interact in similar ways to what we have previously studied for states.

- Supposed I combined more Markov Chains $P1, P2$, and Q as follows:

$$\left[\begin{array}{c|c|c} (1-\epsilon)*\mathbf{P}1 & \epsilon*\mathbf{I} & 0 \\ \hline 0 & (1-\epsilon)*\mathbf{P}2 & \epsilon*\mathbf{I} \\ \hline 0 & 0 & \mathbf{Q} \end{array}\right]$$

What can you say about the limiting distribution? What if I chained N distributions?

Both P1 and P2 are transitive so Q will end up absorbing all the probability mass, similar to above. A chain of N distributions will follow the same properties as long as Q is the only absorbing chain.

- How does this compare to the Fundamental Theorem of Ergodic Markov Chains?

The Fundamental theorem does not apply here since the chain is not irreducible. However, we can still have a limiting distribution, but it doesn't have to have positive probability for every state.

2.3 Problem 3

M balls are initially distributed among m urns. At each stage one of the balls is selected at random, taken from whichever urn it is in, and placed, at random, in one of the other $m-1$ urns. Consider the Markov chain whose state at any time is the vector $(n_1, \ldots, n_m)$ where n_i denotes the number of balls in urn i. Guess at the limiting probabilities for this Markov chain and then verify your guess and show at the same time that the Markov chain is time reversible.

Solution. Recall that if we can find a probability distribution x which satisfies

$$x_i P_{ij} = x_j P_{ji}, \forall i, j$$

then x is the stationary distribution, and the Markov chain is reversible. Thus, we would first guess that x and then verify the equation. That will prove the chain is reversible and also x is the stationary distribution. It's not hard to guess that the stationary distribution would be symmetric in some sense, because all the balls and urns are equivalent. Hence, multinomial distribution would be a very natural guess. Next we are going to verify it. Note that we can't jump from a state to any other state; we can only go to a state where the vector representing numbers of balls only differ from the current one at two coordinates, with one increasing by 1 and the other decreasing by 1. We only need to consider these pairs of states, because for other pairs, the above equation is just $0 = 0$. Without loss of generality (due to symmetry), let's consider the following two states,

$$(n_1, n_2, \ldots, n_m) \text{ and } (n_1 + 1, n_2 - 1, \ldots, n_m)$$

Let's call them state A and B. To go from A to B, we would need to pick a ball from urn 2, with probability $\frac{n_2}{M}$, and then put it in urn 1, with probability $\frac{1}{m-1}$. That gives

$$P_{AB} = \frac{n_2}{M(m-1)}$$

Similarly,

$$P_{BA} = \frac{n_1 + 1}{M(m-1)}$$

For a multinomial distribution, $x_A = \frac{M!}{n_1! n_2! \cdots n_m!}$ and $x_B = \frac{M!}{(n_1+1)!(n_2-1)! \cdots n_m!}$. It is clear now

$$x_A P_{AB} = x_B P_{BA}$$

Because A and B can be any two neighboring states, we have verified the detailed-balance equation.

2.4 Problem 4 (Symmetric Simple Random Walk on $\mathbb{Z}^2$)

Consider a symmetric simple random walk on $\mathbb{Z}^2$, i.e., it moves to one of its four neighbors with equal probability. Show that it's recurrent.

Proof. WLOG, we assume the chain starts from the origin. Recall that we just need to show that $\sum_{n=0}^{\infty} P_{00}^{n} = \infty$. Clearly, $P_{00}^{n} \neq 0$ if and only if n is even, so we show

$$\sum_{n=0}^{\infty} P_{00}^{2n} = \infty$$

To find P_{00}^{2n}, we need to count how many ways there are for the chain to return in exactly $2n$ steps. We need to first choose k steps from $2n$ to go right, and then choose k steps from the remaining $(2n-k)$ to go left, which is $\left(\begin{array}{c} 2n \\ k \end{array}\right)\left(\begin{array}{c} 2n-k \\ k \end{array}\right)$, and then choose $(n-k)$ steps from the remaining $(2n-2k)$ steps to go up, and the remaining $(n-k)$ steps we have to go down. k can be any integer from 0 to n. That means, the number of possible ways is

$$\begin{aligned} &\sum_{k=0}^{n}\left(\begin{array}{c} 2n \\ k \end{array}\right)\left(\begin{array}{c} 2n-k \\ k \end{array}\right)\left(\begin{array}{c} 2n-2k \\ n-k \end{array}\right) \\ =&\frac{(2n)!}{k!k!(n-k)!(n-k)!} \\ =&\sum_{k=0}^{n}\left(\begin{array}{c} 2n \\ n \end{array}\right)\left(\begin{array}{c} n \\ k \end{array}\right)\left(\begin{array}{c} n \\ n-k \end{array}\right) \\ =&\left(\begin{array}{c} 2n \\ n \end{array}\right)^{2} \end{aligned}$$

Thus, $P_{00}^{2n} = \frac{1}{4^{2n}}\left(\begin{array}{c} 2n \\ n \end{array}\right)^{2}$. Plug in Stirling's approximation $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}$, we get $P_{00}^{2n} \sim \frac{C}{n}$ where C is a constant. The divergence of the harmonic series gives us the result.