EE5137 : Stochastic Processes (Spring 2022)
Some Additional Notes on Markov Chains and Classification of States

Vincent Y. F. Tan

February 28, 2022

In this document, we provide some supplementary material to Lecture 7. You need to know Sections 1, 2 and 3 here.

1 Classification of States

Recall that for a finite-state Markov chain, two distinct states i and j communicate (abbreviated $i \leftrightarrow j$) if i is accessible from j and j is accessible from i.

Proposition 1. *Communication is an equivalence relation. That is*

1. $i \leftrightarrow i$;
2. *if* $i \leftrightarrow j$, *then* $j \leftrightarrow i$;
3. *if* $i \leftrightarrow j$ *and* $j \leftrightarrow k$, *then* $i \leftrightarrow k$;

Proof. The first two parts follow directly from the definition. For part 3, suppose that $i \leftrightarrow j$ and $j \leftrightarrow k$; then there exists m and n such that $P_{ij}^m > 0$ and $P_{jk}^n > 0$. Hence,

$$P_{ik}^{m+n} = \sum_{r=0}^{\infty} P_{ir}^m P_{rk}^n \ge P_{ij}^m P_{jk}^n > 0. \tag{1}$$

Similarly, we can show that there exists an s such that $P_{ki}^s > 0$. □

For any states i and j define f_{ij}^n to be the probability that, starting in i, the first transition into j occurs at time n. Formally,

$$f_{ij}^0 = 0, \quad f_{ij}^n = \Pr(X_n = j, X_k \neq j, k = 1, \dots, n-1 | X_0 = i). \tag{2}$$

Let

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^n. \tag{3}$$

Then f_{ij} denotes the probability of ever making a transition into state j give that the process starts in i. Note that for $i \neq j$, $f_{ij} > 0$ if and only if j is accessible from i. State j is said to be *recurrent* if $f_{jj} = 1$, and *transient* otherwise. These definitions are consistent with that in the book.

Proposition 2. *State j is recurrent if, and only if,*

$$\sum_{n=1}^{\infty} P_{jj}^n = \infty. \tag{4}$$

Proof. State j is recurrent if, with probability 1, a process starting at state j with eventually return. However, by the Markovian property, it follows that the process probabilistically restarts itself upon returning to state j. Hence, with probability 1, it will return again to j. Repeating this argument, we see that, with probability 1, the number of visits to state j will be infinite and tus will have infinite expectation. On the other hand, suppose j is transient. Then each time the process returns to j there is a positive probability of $1 - f_{jj}$ that it will never again return; hence, the number of visits is geometric with finite mean $1/(1 - f_{jj})$.

By the above argument, we see that state j is recurrent if and only if

$$\mathbb{E}[\text{number of visits to } j | X_0 = j] = \infty \tag{5}$$

But letting

$$I_n = \begin{cases} 1 & \text{if } X_n = j \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

it follows that $\sum_{n=0}^{\infty} I_n$ denotes the number of visits to j. Since

$$\mathbb{E}\left[\sum_{n=0}^{\infty} I_n \,\middle|\, X_0 = j\right] = \sum_{n=0}^{\infty} \mathbb{E}\left[I_n \mid X_0 = j\right] = \sum_{n=0}^{\infty} P_{jj}^n \tag{7}$$

the result follows. □

Corollary 3. *If i is recurrent and $i \leftrightarrow j$, then j is recurrent.*

Proof. Let m and n be such that $P_{ij}^n > 0$ and $P_{ji}^m > 0$. Now for any $s \geq 0$,

$$P_{jj}^{m+n+s} \geq P_{ji}^m P_{ii}^s P_{ij}^n \tag{8}$$

and thus,

$$\sum_s P_{jj}^{m+n+s} \geq P_{ji}^m P_{ij}^n \sum_s P_{ii}^s = \infty \tag{9}$$

and the result follows from Proposition 2. □

2 The Simple Random Walk

The Markov chain whose state space is the set of all integers and has transition probabilities

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \quad i \in \mathbb{Z} \tag{10}$$

where $p \in (0, 1)$, is called the simple random walk. One interpretation of this process is that it represents the wanderings of a drunken man as he walks along a straight line. Another is that it represents the winnings of a gambler who on each play of the game either wins or loses one dollar.

Since all states communicate with one, it follows from Corollary 3 that the states are either all recurrent or all transient. Let's just focus on state 0 and attempt to determine whether $\sum_n P_{00}^n$ is finite or infinite.

Since it is impossible to be even (win 0 dollars) after an odd number of steps,

$$P_{00}^{2n+1} = 0, \quad n \in \mathbb{N}. \tag{11}$$

On the other hand, the gambler would be even after $2n$ trials if and only if he won n of those trials and lost n. This probability is

$$P_{00}^{2n} = \binom{2n}{n} p^n (1-p)^n = \frac{(2n)!}{(n!)^2} (p(1-p))^n, \quad n \in \mathbb{N}. \tag{12}$$

By using the Stirling approximation,[1]

$$n! \sim n^{n+1/2} e^{-n} \sqrt{2\pi}, \tag{13}$$

[1]We write $a_n \sim b_n$ if $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$

we obtain

$$P_{00}^{2n} \sim \frac{(4p(1-p))^n}{\sqrt{\pi n}}. \tag{14}$$

Hence, $\sum_n P_{00}^n < \infty$ if and only if $p \neq 1/2$ (note that $\sum_n \frac{1}{\sqrt{n}} = \infty$). Thus the chain is recurrent if $p = 1/2$ and transient if $p \neq 1/2$.

When $p = 1/2$, the above process is called a *one-dimensional symmetric random walk*. We could also look at symmetry random walks in more than one dimension. For instance, in the two-dimensional symmetric random walk, the process would, at each transition, either take a step to the left, right, up or down, each with probability $1/4$. Similarly in three dimensions, the process would, with probability $1/6$, make a transition to any of the six adjacent points. By using the same method as the one-dimensional random walk, it can be shown that the two-dimensional symmetric random walk is recurrent, but all higher-dimensional symmetric random walks are transient.

3 Clarification about the Proof of Theorem 4.2.8

We know that

$$d(i) \mid t$$

where t is any number in the set $T := \{t : P_{jj}^t > 0\}$. Thus $d(i)$ is a common divisor of the elements of T. By definition, $d(j)$ is the *greatest* common divisor of T. It is known that every common divisor of a set of numbers T divides the greatest common divisor, i.e., it holds that

$$d(i) \mid d(j).$$

For a proof of this non-trivial fact, see `https://proofwiki.org/wiki/Common_Divisor_Divides_GCD` or `https://www.cut-the-knot.org/Generalization/gcd.shtml`. Also see Lemma 4 below for a self-contained proof. Note that $d(j)$ need not be in set T.

Consider the example in which $T = \{t : P_{jj}^t > 0\} = \{4, 8, 10, \ldots\}$ for any $j \in \{1, \ldots, 9\}$ in Fig. 4.2(b) in Gallager's book. Note that $d(j) = \gcd(T) = 2$. Note that $d(j) = 2$ need not be in T, i.e., $P_{jj}^2 = P_{jj}^{d(j)}$ could be (and in fact is) 0. However, $d(i) = 2$ divides every element in the set T. It also divides $d(j)$, i.e.,

$$d(i) \mid d(j).$$

In fact, both $d(i)$ and $d(j)$ are 2 in this case.

Lemma 4. *Say we have two natural numbers m and n, whose greatest common divisor is $d = \gcd(m, n)$. Let a be any common divisor of m and n. It holds that $a \mid d$.*

Proof. Assume, to the contrary, that a cannot exactly divide d, then by definition of exact division, there exist x, y and z such that $a = xy$, and $d = xz$, but $y > 1$ and $z > 1$ are relatively prime. Since a and d divide m, $xy \mid m$ and $xz \mid m$ where y, z are relatively prime. Thus, there exists $a_1, d_1 \in \mathbb{N}$ such that $m = xya_1 = xzd_1$. This implies that $ya_1 = zd_1$. This implies that $z \mid ya_1$ and since y and z are relatively prime, $z \mid a_1$. This implies that $xyz \mid xya_1$ and so $xyz \mid m$.

Similarly, there exists $a_2, d_2 \in \mathbb{N}$ such that $n = xya_2 = xzd_2$. This implies that $ya_2 = zd_2$. This implies that $z \mid ya_2$ and since y and z are relatively prime, $z \mid a_2$. This implies that $xyz \mid xya_2$ and so $xyz \mid n$.

Hence, xyz is a common divisor of m and n. But $xyz > xz = d$. This contradicts the fact that d is the *greatest* common divisor of m and n. $\square$