Stochastic Processes 1

STAT 333

Fall 2021 (1219)[1]

Cameron Roopnarine[2] Steve Drekic[3] Mirabelle Huynh[4]

24th October 2021

[1]Online Course
[2]LaTeXer
[3]Main Instructor
[4]Assisting Instructor

Contents

Week 6
20th to 27th October

0.1 Limiting Behaviour of DTMCs

The concepts of periodicity and transience/recurrence play an important role in characterizing the limiting behaviour of a DTMC. To demonstrate their influence, let us consider three examples with varying forms of limiting behaviour.

Example 3.10. Consider a DTMC with TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}.$$

Determine if $\lim_{n\to\infty} P^{(n)}$ exists.

Solution:

Example 3.11. Consider a DTMC with TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \end{array}.$$

Determine if $\lim_{n\to\infty} P^{(n)}$ exists.

Solution:

Example 3.12. Consider a DTMC with TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Examine $\lim_{n\to\infty} P^{(n)}$, and explain why the limiting probability of being in a state can depend on the initial state of this DTMC.

Solution:

In the above example, note that the second column of the limiting matrix contains all zeros. Not surprisingly, this is indicative of transient behaviour, implying that one will never end up in state 1 in the long run. This property can be proven more formally in the next theorem.

Theorem 3.6. For any state i and transient state j of a DTMC, $\lim_{n\to\infty} P_{i,j}^{(n)} = 0$.

Proof:

Mean Recurrent Time

As the previous three examples show, there is variation in the limiting behaviour of a DTMC. In particular, it is worthwhile to determine a set of conditions which ensure the "nice" limiting behaviour witnessed in Example 3.11. To ascertain when such conditions exist, we need to distinguish between two kinds of recurrence. Let

$$N_i = \min\{n \in \mathbb{Z}^+ : X_n = i\},$$

where state i is assumed to be recurrent. Clearly, the conditional rv $N_i \mid (X_0 = i)$ takes on values in $\mathbb{Z}^+$. Moreover, its conditional pmf is given by

$$\mathbb{P}(N_i = n \mid X_0 = i) = f_{i,i}^{(n)},\ n = 1, 2, 3, \ldots.$$

We observe that this is indeed a pmf since $\sum_{n=1}^{\infty} f_{i,i}^{(n)} = f_{i,i} = 1$, as state i is recurrent. This leads to the introduction of the following important quantity.

Definition: If state i is recurrent, then its *mean recurrent time* is given by

$$m_i = \mathbb{E}[N_i \mid X_0 = i] = \sum_{n=1}^{\infty} n f_{i,i}^{(n)}.$$

Positive and Null Recurrence

In words, m_i represents the average time it takes the DTMC to make successive visits to state i. Two notions of recurrence can now be defined based on the value of m_i.

Definition: Suppose that state i is recurrent. State i is said to be *positive recurrent* if $m_i < \infty$. On the other hand, state i is said to be *null recurrent* if $m_i = \infty$.

Remark: A fair question to ask is whether it is even possible for a discrete probability distribution on $\mathbb{Z}^+$ to have an **undefined** mean (i.e., a mean of ∞). To show that this is indeed possible, consider a rv X with pmf

$$\mathbb{P}(X = x) = \frac{1}{x(x+1)},\ x = 1, 2, 3, \ldots.$$

Let us first confirm that this is indeed a pmf:

$$\begin{aligned}\sum_{x=1}^{\infty}\frac{1}{x(x+1)} &= \lim_{n\to\infty}\sum_{x=1}^{n}\frac{1}{x(x+1)} \\ &= \lim_{n\to\infty}\sum_{x=1}^{n}\left(\frac{1}{x}-\frac{1}{x+1}\right) \\ &= \lim_{n\to\infty}\left\{\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\cdots+\left(\frac{1}{n}-\frac{1}{n+1}\right)\right\} \\ &= \lim_{n\to\infty}\left(1-\frac{1}{n+1}\right) \\ &= 1.\end{aligned}$$

However,

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} x \cdot \frac{1}{x(x+1)} = \sum_{x=1}^{\infty}\frac{1}{x+1} = \infty,$$

since the above harmonic series is known to diverge. In other words, a finite mean does not exist!

Some Facts About Positive and Null Recurrence:

1. If $i \leftrightarrow j$ and state i is positive recurrent, then state j is also positive recurrent. This means that positive recurrence is also a class property. An obvious by-product of this result is that null recurrence is a class property too.
2. In a finite-state DTMC, there can **never** be any null recurrent states.

Remarks:

(1) The above facts are provided without formal justification, as their proofs are rather lengthy and depend on material beyond the scope of STAT 333.

(2) Positive recurrent, aperiodic states are referred to as *ergodic* states.

Stationary Distribution

Before stating the main result governing the "nice" limiting behaviour demonstrated in Example 3.11, we introduce a special type of probability distribution.

Definition: A probability distribution $\{p_i\}_{i=0}^{\infty}$ is called a *stationary distribution* of a DTMC if $\{p_i\}_{i=0}^{\infty}$ satisfies the conditions $\sum_{i=0}^{\infty} p_i = 1$ and $p_j = \sum_{i=0}^{\infty} p_i P_{i,j}$ $\forall j \in \mathbb{N}$.

Remark: If we define the row vector

$$\underline{p} = (p_0, p_1, \ldots, p_j, \ldots),$$

then the above conditions can be represented in matrix form as

$$\underline{p}\,\underline{e}^{\top} = 1 \text{ and } \underline{p} = \underline{p}P,$$

where $\underline{e}^{\top} = (1, 1, \ldots, 1, \ldots)^{\top}$ denotes a column vector of ones (in general, the $^{\top}$ notation will be used to represent column vectors).

A logical question to ask is "*Why is such a distribution called stationary*?"
To answer this question, suppose that the initial conditions of the DTMC are given by $\underline{\alpha}_0 = \underline{p}$. As a result, we have that $\alpha_{0,j} = \mathbb{P}(X_0 = j) = p_j \ \forall j \in \mathbb{N}$. Now, for any $j \in \mathbb{N}$, note that

$$\alpha_{1,j} = \mathbb{P}(X_1 = j) = \sum_{i=1}^{\infty} \alpha_{0,i} P_{i,j} = \sum_{i=0}^{\infty} p_i P_{i,j} = p_j = \alpha_{0,j}.$$

The above equation indicates that X_1 has the same probability distribution as X_0 when $\underline{\alpha}_0 = \underline{p}$. More generally, it is straightforward to show (using mathematical induction) that each X_i, $i \in \mathbb{Z}^+$, is *identically distributed* to X_0, provided that $\underline{\alpha}_0 = \underline{p}$.
In other words, if a DTMC is started according to a stationary distribution, then the probability of being in a given state remains *unchanged* (i.e., stationary) over time.

Remarks:

(1) In some texts, the stationary probability distribution is sometimes called the *invariant probability distribution* or *steady-state probability distribution*.

(2) A known fact (which again we do not prove formally) is that a stationary distribution will not exist if all the states of the DTMC are either null recurrent or transient. On the other hand, an irreducible DTMC is positive recurrent iff a stationary distribution exists.

(3) Stationary distributions are not necessarily unique. This happens when a DTMC has more than one positive recurrent communication class. For instance, it is not difficult to verify that the DTMC in Example 3.10 has an **infinite** number of stationary distributions (left as an upcoming exercise).

The Basic Limit Theorem

We are now in position to state the fundamental limiting theorem for DTMCs, generally referred to as the Basic Limit Theorem (BLT).

Basic Limit Theorem: For an irreducible, recurrent, and aperiodic DTMC, $\lim_{n\to\infty} P_{i,j}^{(n)}$ exists and is independent of state i, satisfying

$$\lim_{n\to\infty} P_{i,j}^{(n)} = \pi_j = \frac{1}{m_j} \ \forall i,j \in \mathbb{N}.$$

If the DTMC also happens to be positive recurrent, then $\{\pi_j\}_{j=0}^{\infty}$ is the unique, positive solution to the system of linear equations defined by

$$\begin{cases} \pi_j = \sum_{i=0}^{\infty} \pi_i P_{i,j} \ \forall j \in \mathbb{N}, \\ \sum_{j=0}^{\infty} \pi_j = 1. \end{cases}$$

Remarks:

(1) A formal proof of the BLT is beyond the scope of STAT 333. However, it is not difficult to understand why $\{\pi_j\}_{j=0}^{\infty}$ (if they exist) satisfies the above system of linear equations. Specifically, recall the Chapman-Kolmogorov equations with $m = n - 1$, namely

$$P_{i,j}^{(n)} = \sum_{k=0}^{\infty} P_{i,k}^{(n-1)} P_{k,j} \ \forall i,j \in \mathbb{N}$$

Taking the limit as $n \to \infty$ of both sides of this equation and assuming that it is permissible to pass the limit through the summation sign, we obtain

$$\lim_{n\to\infty} P_{i,j}^{(n)} = \lim_{n\to\infty} \sum_{k=0}^{\infty} P_{i,k}^{(n-1)} P_{k,j}.$$

$$\pi_j = \sum_{k=0}^{\infty} \lim_{n\to\infty} P_{i,k}^{(n-1)} P_{k,j} = \sum_{k=0}^{\infty} \pi_k P_{k,j} \forall j \in \mathbb{N},$$

which is precisely the above system of equations.

(2) If we define the row vector of limiting probabilities

(3)

$$\underline{\pi} = (\pi_0, \pi_1, \ldots, \pi_j, \ldots),$$

then the above system of linear equations can be written succinctly in matrix form as:

$$\left\{ \underline{\pi} = \underline{\pi} P, \underline{\pi}\underline{e}^{\top} = 1. \right.$$

Therefore, if a DTMC is irreducible and ergodic, then the BLT states that the limiting probability distribution is the unique stationary distribution.

(4) When a DTMC has a finite number of states (i.e., suppose that the state space is $\{0, 1, \ldots, N\}$ where $N < \infty$), the BLT states that there are $N + 1$ linear equations to consider of the form

$$\pi_j = \sum_{i=0}^{N} \pi_i P_{i,j}, \; j = 0, 1, \ldots, N. \tag{3.8}$$

Along with the condition $\sum_{j=0}^{N} \pi_j = 1$, this leads to $N + 2$ equations in $N + 1$ unknowns, of which a unique solution must exist. In fact, the first $N + 1$ equations given by (3.8) are linearly dependent (implying that there is a redundancy), and so we can drop any one of the equations given by (3.8) and solve the remaining $N + 1$ equations to obtain a unique solution.

(5) If the conditions of the BLT are satisfied and state j happens to be null recurrent, then $\pi_j = 0$ which interestingly is similar to the limiting behaviour of a transient state.

Example 3.11. (*continued*) Recall that we previously considered a DTMC with TPM

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \end{array}.$$

Find the limiting probabilities for this DTMC.

Solution:

Doubly Stochastic TPM

Recall that the TPM of a DTMC is stochastic, with all row sums of P being equal to 1. However, a TPM is said to be *doubly stochastic* if all column sums of P are also equal to 1 (i.e., $\sum_{i=0}^{\infty} P_{i,j} = 1 \; \forall j \in \mathbb{N}$). The following theorem provides an interesting result concerning the limiting behaviour of a class of such DTMCs.

Theorem 3.7. Suppose that a finite-state DTMC with state space $\mathcal{S} = \{0, 1, \ldots, N-1\}$ is irreducible and aperiodic. If the associated TPM is doubly stochastic, then the limiting probabilities $\{\pi_j\}_{j=0}^{N-1}$ exist and are given by

$$\pi_j = \frac{1}{N},\ j = 0, 1, \ldots, N-1.$$

Proof:

Alternative Interpretation

The primary interpretation of the limiting distribution of a DTMC is that after the process has been in operation for a "long" period of time, the probability of finding the process in state j is π_j (assuming the conditions of the BLT are met). In such situations, however, another interpretation exists for π_j. Specifically, π_j also represents the "long-run" mean fraction of time that the process spends in state j.
To see that this interpretation is valid, define the sequence of indicator random variables $\{A_k\}_{k=1}^{\infty}$ as follows:

$$A_k = \begin{cases} 0, & \text{if } X_k \neq j, \\ 1, & \text{if } X_k = j. \end{cases}$$

The fraction of time the DTMC visits state j during the time interval from 1 to n inclusive is therefore given by

$$\frac{1}{n}\sum_{k=1}^{n} A_k.$$

Looking at the quantity

$$\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n} A_k \,\middle|\, X_0 = i\right],$$

which is interpreted as the mean fraction of time spent in state j during the time interval from 1 to n inclusive, given that the process starts in state i, note that

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n} A_k \,\middle|\, X_0 = i\right] &= \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}[A_k \mid X_0 = i] \\
&= \frac{1}{n}\sum_{k=1}^{n} \big(0 \cdot \mathbb{P}(A_k = 0 \mid X_0 = i) + 1 \cdot \mathbb{P}(A_k = 1 \mid X_0 = i)\big) \\
&= \frac{1}{n}\sum_{k=1}^{n} \mathbb{P}(X_k = j \mid X_0 = i) \\
&= \frac{1}{n}\sum_{k=1}^{n} P_{i,j}^{(k)}.
\end{aligned}$$

We have: $\mathbb{E}\big[\frac{1}{n}\sum_{k=1}^{n} A_k \,\big|\, X_0 = i\big] = \frac{1}{n}\sum_{k=1}^{n} P_{i,j}^{(k)}$.
Recall: If $\{a_n\}_{n=1}^{\infty}$ is a real sequence such that $a_n \to a$ as $n \to \infty$, then $\frac{1}{n}\sum_{k=1}^{n} a_k \to a$ as $n \to \infty$.
Thus, if the conditions of the BLT are satisfied, then $P_{i,j}^{(n)} \to \pi_j$ as $n \to \infty$. Therefore, applying the above result with $a_n = P_{i,j}^{(n)}$ and $a = \pi_j$, we obtain

$$\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n} A_k \,\middle|\, X_0 = i\right] \to \pi_j \text{ as } n \to \infty,$$

implying that the long-run mean fraction of time spent in state j is also equal to π_j.

<u>Remark</u>: If one begins in recurrent state j, we realize that the process spends one unit of time in state j every N_j time units. On average, this amounts to one unit of time in state j every $\mathbb{E}[N_j \mid X_0 = j] = m_j$ time units. If the conditions of the BLT are satisfied, then it makes sense intuitively that $\pi_j = 1/m_j$, as the BLT specifies. For a more formal justification in the positive recurrent case, let $\{N_j^{(n)}\}_{n=1}^{\infty}$ be a sequence of rvs where $N_j^{(n)}$ represents the number of transitions between the $(n-1)^{\text{th}}$ and n^{th} visits into state j, as illustrated in the diagram below. By the Markov property and the stationary assumption of the DTMC, $\{N_j^{(n)}\}_{n=1}^{\infty}$ is actually an iid sequence of rvs with common mean $m_j < \infty$. Therefore, the long-run fraction of time spend in state j can be viewed as

$$\pi_j = \lim_{n\to\infty} \frac{n}{\sum_{i=1}^{n} N_j^{(i)}} = \lim_{n\to\infty} \frac{1}{\frac{1}{n}\sum_{i=1}^{n} N_j^{(i)}} = \frac{1}{m_j},$$

where the last equality follows from the SLLN.