Probability Theory and Random Processes (MA 225)

Class Notes
September – November, 2020

Instructor
Ayon Ganguly
Department of Mathematics
IIT Guwahati

Contents

1 Probability **3**
- 1.1 Probability . . . 3
  - 1.1.1 Classical Probability . . . 3
  - 1.1.2 Countable and Uncountable Sets . . . 4
  - 1.1.3 Axiomatic Probability . . . 5
  - 1.1.4 Continuity of Probability . . . 10
- 1.2 Conditional Probability . . . 12
- 1.3 Independence . . . 14

2 Random Variable **18**
- 2.1 Random Variable . . . 18
- 2.2 Discrete Random Variable . . . 22
- 2.3 Continuous Random Variable . . . 26
- 2.4 Expectation of Random Variable . . . 32
- 2.5 Transformation of Random Variable . . . 36
  - 2.5.1 Technique 1 . . . 36
  - 2.5.2 Technique 2 . . . 38
  - 2.5.3 Expectation of Function of RV . . . 41
  - 2.5.4 Technique 3 . . . 45
- 2.6 Moment Inequality . . . 46
- 2.A Gamma Integral . . . 47
- 2.B Beta Integral . . . 49

3 Jointly Distributed Random Variables **51**
- 3.1 Random Vector . . . 51
- 3.2 Discrete Random Vector . . . 54
- 3.3 Continuous Random Vector . . . 57
- 3.4 Expectation of Function of Random Vector . . . 59
- 3.5 Some Useful Remarks . . . 60
- 3.6 Independent Random Variables . . . 62
- 3.7 Covariance and Correlation Coefficient . . . 63
- 3.8 Transformation Techniques . . . 65
  - 3.8.1 Technique 1 . . . 65
  - 3.8.2 Technique 2 . . . 68
  - 3.8.3 Technique 3 . . . 71
- 3.9 Conditional Distribution . . . 73
  - 3.9.1 For Discrete Random Vector . . . 73
  - 3.9.2 For Continuous Random Vector . . . 75

3.9.3 Computing Expectation by Conditioning 77
3.10 Bivariate Normal Distribution 83
3.11 Some Results on Independent and Identically Distributed Normal RVs 85

4 Limit Theorems **88**
4.1 Modes of Convergence 88
4.2 Limit Theorems 94

5 Markov Chain **98**
5.1 Introduction 98
5.2 Chapman-Kolmogrov Equations 101
5.3 Classification of States 104

Chapter 5

Markov Chain

5.1 Introduction

Definition 5.1 (Stochastic Processes)**.** *Let T be a countable set of time index. A function $X : \mathcal{S} \times T \to \mathbb{R}$ is called a stochastic process. As T is countable, we can take $T = \{0, 1, 2, 3, \ldots\}$, and we shall denote a stochastic process by $\{X_n : n \geq 0\}$.*

Example 5.1. Example of stochastic processes includes number of students obtaining AA grade in MA225 in each year, price of gold on different days, minimum temperature at Guwahati of each day. ||

Definition 5.2. *The set of all possible values taken by a stochastic process $\{X_n : n \geq 0\}$ is called state space.*

We will only consider atmost countable state spaces. In general, we will assume that the state space is $\{0, 1, 2, \ldots\}$, if the state space is infinite. If the state space is finite, we will use a finite subset of $\{0, 1, 2, \ldots\}$ as the state space. If $X_n = i$, then the process in said to be in state i at time n.

Definition 5.3 (Markov chain)**.** *A stochastic process $\{X_n : n \geq 0\}$ is said to be a Markov chain (MC) if*

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

for all states $i_0, i_1, \ldots, i_{n-1}, i$ and j and all $n \geq 0$.

Thus, for a MC, the conditional distribution of future X_{n+1} given the past $X_0, X_1, \ldots, X_{n-1}$ and present X_n depends only on the present X_n and not on past $X_0, X_1, \ldots, X_{n-1}$.

Definition 5.4 (Time-homogeneous MC)**.** *A MC $\{X_n : n \geq 0\}$ is said to be a time-homogeneous MC if*

$$P(X_{n+1} = j | X_n = i) = P(X_n = j | X_{n-1} = i)$$

for all $n \geq 1$ and for all states i and j.

Thus, for time-homogeneous MC, $P(X_{n+1} = j | X_n = i)$ does not depend on n and we will denote this probability by $p_{i,j}$. We will consider only time-homogeneous MC in this course. Hence, for the course MC will mean time-homogeneous MC, if not otherwise mentioned.

Definition 5.5. *The $p_{i,j}$ is called one-step transition probability from state i to state j. The matrix*

$$P = ((p_{i,j}))_{i,j}$$

is called one-step transition probability matrix or transition probability matrix (TPM).

As $p_{i,j}$'s are a conditional probabilities, $p_{i,j} \geq 0$ for all i and j. Also

$$\sum_{j=0}^{\infty} p_{i,j} = 1 \text{ for all states } i.$$

Thus, row sum of the matrix P is one for all its' rows.

Example 5.2. Suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today and not on past weather condition. Suppose that if it is raining today, then it will rain tomorrow with probability α. If it is not raining today, then it will rain tomorrow with probability β. Define

$$X_n = \begin{cases} 0 & \text{if it rains on day } n \\ 1 & \text{if it does not rain on day } n. \end{cases}$$

Then $\{X_n : n \geq 0\}$ is a two-state MC with state space $\{0, 1\}$. The one-step transition probability matrix of the MC is given by

$$\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{bmatrix} \end{array}.$$

||

Example 5.3. Suppose that the chance of rain tomorrow depends on previous weather conditions through last two days. Assume that

$$\begin{aligned}
&P\,(\text{Rain tomorrow}|\text{rain for past two days}) = 0.7, \\
&P\,(\text{Rain tomorrow}|\text{rain today, no rain yesterday}) = 0.5, \\
&P\,(\text{Rain tomorrow}|\text{rain yesterday, no rain today}) = 0.4, \\
&P\,(\text{Rain tomorrow}|\text{no rain for past two days}) = 0.2.
\end{aligned}$$

If we define a process like the previous example, then it will not be a MC, as for example $P\,(X_{n+1} = 0|X_n = 0,\, X_{n-1} = 0)$ may not be equal to $P\,(X_{n+1} = 0|X_n = 0)$. However, we can transform this model into a MC by saying that the state at any time is determined by the weather conditions during both that day and the previous day. Thus, define

$$Y_n = \begin{cases} 0 & \text{if it rained both on days } n \text{ and } n-1 \\ 1 & \text{if it rained on day } n \text{ but not on day } n-1 \\ 2 & \text{if it rained on day } n-1 \text{ but not on day } n \\ 3 & \text{if it did not rain on days } n \text{ and } n-1. \end{cases}$$

Clearly, $\{Y_n : n \geq 1\}$ is a MC with TPM

$$\begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & 0.7 & 0 & 0.3 & 0 \\ 1 & 0.5 & 0 & 0.5 & 0 \\ 2 & 0 & 0.4 & 0 & 0.6 \\ 3 & 0 & 0.2 & 0 & 0.8 \end{array}.$$

||

Example 5.4. Consider a communication system that transmit the digits 0 and 1. Each digit, transmitted must pass several stages, at each stage of which there is a probability p that the digit entered will remain unchanged when it leaves. Letting X_n denote the digit entering the nth stage, then $\{X_n : n \geq 0\}$ is a two-state MC with TPM

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & p & 1-p \\ 1 & 1-p & p \end{array}.$$

||

Example 5.5. A MC whose state space is given by set of integers is said to be simple random walk (SRW) if for some $0 < p < 1$, $p_{i,i+1} = p = 1 - p_{i,i-1}$ for $i = 0, \pm 1, \pm 2, \ldots$. The preceding MC is called a random walk for we may think of it as being a model for an individual walking on a straight line who at each point of time either takes one step to the right with probability p or one step to the left with probability $1 - p$. It is said to be simple symmetric random walk (SSRW) if $p = 1/2$. ||

Example 5.6. Let $\{S_n : n \geq 0\}$ be a SSRW, which starts from state 0, *i.e.*, $P(S_0 = 0) = 1$. Define $M_n = \max_{0 \leq i \leq n} S_i$ for $n \geq 0$. Then $\{M_n : n \geq 0\}$ is not a MC. To see it, we can take two different paths of the process M_n for $0 \leq n \leq 4$ and show that the conditional probabilities of the final step are not same. Let the first path be $M_0 = 0$, $M_1 = 1$, $M_2 = 1$, $M_3 = 1$, and $M_4 = 2$ and the second path be $M_0 = 0$, $M_1 = 0$, $M_2 = 0$, $M_3 = 1$, and $M_4 = 2$. Note that both the paths have same values for the last two steps (1 and 2 here). Now, for the first path

$$P(M_4 = 2, M_3 = 1, M_2 = 1, M_1 = 1, M_0 = 0) = P(S_0 = 0, S_1 = 1, S_2 = 0, S_3 = 1, S_4 = 2) = \left(\frac{1}{2}\right)^4.$$

$$P(M_3 = 1, M_2 = 1, M_1 = 1, M_0 = 0) = P(S_0 = 0, S_1 = 1, S_2 = 0, S_3 = 1 \text{ or } -1) = \left(\frac{1}{2}\right)^2.$$

Thus,

$$P(M_4 = 2 | M_3 = 1, M_2 = 1, M_1 = 1, M_0 = 0) = \frac{1}{4}.$$

For the second path,

$$P(M_4 = 2, M_3 = 1, M_2 = 0, M_1 = 0, M_0 = 0) = P(S_0 = 0, S_1 = -1, S_2 = 0, S_3 = 1, S_4 = 2) = \left(\frac{1}{2}\right)^4.$$

$$P(M_3 = 1, M_2 = 0, M_1 = 0, M_0 = 0) = P(S_0 = 0, S_1 = -1, S_2 = 0, S_3 = 1) = \left(\frac{1}{2}\right)^3.$$

Thus,

$$P(M_4 = 2 | M_3 = 1, M_2 = 1, M_1 = 1, M_0 = 0) = \frac{1}{2}.$$

To summarize, this specific example shows that the conditional probability of the step $M_3 = 1 \to M_4 = 2$ depends not only on the fact that $M_3 = 1$ but on M_k for $0 \le k \le 2$ as well. Thus, $\{M_n : n \ge 0\}$ is not a MC. ||

Example 5.7 (A Gambling Model)**.** Consider a gambler who at each play of the game either wins Re. 1 with probability p or losses Re. 1 with probability $1 - p$. Suppose that the gambler quits plays either when he goes broke or he attains a fortune of Rs. N. Let X_n denote the gambler's fortune after nth play. Then $\{X_n : n \ge 0\}$ is a MC with transition probabilities

$$p_{i,i+1} = p = 1 - p_{i,i-1} \quad \text{for } i = 1,\, 2,\, 3,\, \ldots,\, N-1 \text{ and } p_{0,0} = p_{N,N} = 1.$$

State 0 and N is called absorbing states since once entered they are never left. Note that this is a finite state random walk with absorbing barriers (state 0 and N). ||

5.2 Chapman-Kolmogrov Equations

Theorem 5.1 (Chapman-Kolmogorov Equations)**.** *Consider a Markov Chain having state space $\{0,\, 1,\, 2,\, \ldots\}$ and one-step transition probabilities $p_{i,j}$ for $i,\, j = 0,\, 1,\, 2,\, \ldots$. Let us define the n-step transition probabilities*

$$p_{i,j}^{(n)} = P\left(X_n = j | X_0 = i\right) = P\left(X_{n+m} = j | X_m = i\right).$$

The Chapman-Kolmogorov equations are given by

$$p_{i,j}^{(m+n)} = \sum_{k=0}^{\infty} p_{i,k}^{(m)} p_{k,j}^{(n)}$$

for all $m, n \ge 0$ and all $i,\, j = 0,\, 1,\, 2,\, \ldots$.

Proof:

$$\begin{aligned}
p_{i,j}^{(m+n)} &= P\left(X_{m+n} = j | X_0 = i\right) \\
&= \sum_{k=0}^{\infty} P\left(X_{m+n} = j,\, X_m = k | X_0 = i\right), \quad \text{using the Theorem of Total Probability} \\
&= \sum_{k=0}^{\infty} P\left(X_{m+n} = j | X_m = k,\, X_0 = i\right) P\left(X_m = k | X_0 = i\right) \\
&= \sum_{k=0}^{\infty} P\left(X_{m+n} = j | X_m = k\right) P\left(X_m = k | X_0 = i\right) \\
&= \sum_{k=0}^{\infty} p_{i,k}^{(m)} p_{k,j}^{(n)}.
\end{aligned}$$

This theorem is most easily understood by noting that $p_{i,k}^{(n)} p_{k,j}^{(m)}$ represents the probability that starting from i the process will go to state j in $n+m$ steps through a path which takes it into state k at the nth transition. Hence, summing over all intermediate states k yields the probability that the process will be in state j after $n+m$ transition. □

Corollary 5.1. *If we denote n-step transition probability matrix by* $P^{(n)} = \left(\left(p_{i,j}^{(n)}\right)\right)_{i,j}$, *then*

$$P^{(n+m)} = P^{(n)} P^{(m)}.$$

Moreover,

$$P^{(n)} = P^n.$$

Proof: Straight forward from the previous theorem. $\square$

Example 5.8 (Continuision of Example 5.2)**.** Suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today and not on past weather condition. Suppose that if it is raining today, then it will rain tomorrow with probability $\alpha = 0.7$. If it is not raining today, then it will rain tomorrow with probability $\beta = 0.4$. Then the probability that it will rain four days from today given that it is raining today, i.e., $P\left(X_4 = 0 | X_0 = 0\right)$ can be calculated from P^4, where P is the TPM. Here,

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \implies P^{(4)} = P^4 = \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix}.$$

Thus, $P\left(X_4 = 0 | X_0 = 0\right) = 0.5749$. ||

Example 5.9. Suppose that balls are successively distributed among 8 urns, with each ball being equally likely to be put in any of these urns. Suppose that we are interested to find the probability that there will be exactly 3 occupied urns after 9 balls have been distributed. Let X_n denote the number of occupied urn after n balls have been distributed. Then $\{X_n : n \geq 0\}$ is a MC with state space $\{0, 1, 2, \ldots, 8\}$ and TPM

$$\begin{array}{c c} & \begin{array}{ccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{8} & \frac{7}{8} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{8} & \frac{6}{8} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{8} & \frac{5}{8} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{4}{8} & \frac{4}{8} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{5}{8} & \frac{3}{8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{6}{8} & \frac{2}{8} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{7}{8} & \frac{1}{8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

The desired probability is $p_{0,3}^{(9)}$. This probability can be found by computing 9th power of the TPM and collecting the (1, 3)th element. However, notice that $p_{0,3}^{(9)} = \sum_{k=0}^{8} p_{0,k} p_{k,3}^{(8)} = p_{1,3}^{(8)}$ as $p_{0,1} = 1$ and $p_{0,k} = 0$ for $k \neq 1$. Thus, we need the probability that there are exactly 3 occupied urns after 8 balls have been distributed starting with a single occupied urn. To do so, we can use the fact that the state of the MC cannot decrease and collapse all states 4, 5, 6, 7, and 8 into a single state, call it state 4. Thus, we define a new MC as follows:

$$Y_n = \begin{cases} 1 & \text{if exactly 1 urn is occupied after } n \text{ balls had been distributed} \\ 2 & \text{if exactly 2 urns are occupied after } n \text{ balls had been distributed} \\ 3 & \text{if exactly 3 urns are occupied after } n \text{ balls had been distributed} \\ 4 & \text{if more than 3 urns are occupied after } n \text{ balls had been distributed.} \end{cases}$$

The TPM of the MC $\{Y_n\}$ is given by

$$P = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} \frac{1}{8} & \frac{7}{8} & 0 & 0 \\ 0 & \frac{2}{8} & \frac{6}{8} & 0 \\ 0 & 0 & \frac{3}{8} & \frac{5}{8} \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Now, we need to find 8 step transition probability $p_{1,3}^{(8)}$ for the MC $\{Y_n\}$. We can compute $P^8 = P^4 P^4$ and collect the (1, 3)th element of P^8. Note that

$$P^{(4)} = P^4 = \begin{bmatrix} 0.0002 & 0.0256 & 0.2563 & 0.7178 \\ 0 & 0.0039 & 0.0952 & 0.9009 \\ 0 & 0 & 0.0198 & 0.9802 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus, $p_{1,3}^{(8)} = 0.0002 \times 0.2563 + 0.0256 \times 0.0952 + 0.2563 \times 0.0198 + 0.7178 \times 0 = 0.00756.$ ||

Example 5.10. In a sequence of independent flips of a fair coin, let N denote the number of flips until there is a run of three consecutive heads. We can find $P(N \leq 8)$ by constructing an appropriate MC. Let us define a MC with states 0, 1, 2, and 3, where, for $i = 0, 1, 2$, state i means we are currently on a run of i heads flips and run of 3 heads has not occurred yet, and state 3 means a run of 3 heads has already occurred. Thus, the TPM of the MC is given by

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

We can find the TPM using the following arguments. If currently we are on a run of 0 heads, we will move to either state 0 (if a tail occurs in the next flip) or to state 1 (if a head occurs on the next flip). Hence, $p_{0,0} = \frac{1}{2} = p_{0,1}$. Similarly, the 2nd and 3rd row of the TPM can be found. If we enter the state 3, we will never leave it. Hence, $p_{3,3} = 1$. Notice that $P(N \leq 8) = p_{0,3}^{(8)}$ as $N \leq 8$ occurs if and only if we are in state 3 after 8 flips. Hence, we need to find the P^8, which is given by

$$P^8 = \begin{bmatrix} \frac{81}{256} & \frac{44}{256} & \frac{24}{256} & \frac{107}{256} \\ \frac{68}{256} & \frac{37}{256} & \frac{20}{256} & \frac{131}{256} \\ \frac{44}{256} & \frac{24}{256} & \frac{13}{256} & \frac{175}{256} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Thus, $P(N \leq 8) = \frac{107}{265}$. We can also find

$$P(N = 8) = P(N \leq 8) - P(N \leq 7) = p_{0,3}^{(8)} - p_{0,3}^{(7)}.$$

Of course, we need to perform some calculations, but once we can construct appropriate MC finding out such probabilities become quite easy. ||

Remark 5.1. Like a random variable is probabilistically specified by its distribution, a stochastic process is specified by its finite dimensional distributions, *i.e.*, by

$$P\left(X_0=i_0,\, X_1=i_1,\, \ldots,\, X_n=i_n\right) \text{ for all } n \geq 0 \text{ and all } i_0,\, i_1,\, \ldots,\, i_n \in S,$$

where S is the state space. Now, assume that $\{X_n : n \geq 0\}$ is a MC and $P(X_0 = i) = \mu_i$ for $i \in S$. Then

$$\begin{aligned}
&P\left(X_0=i_0,\, X_1=i_1,\, \ldots,\, X_n=i_n\right)\\
&= P\left(X_n=i_n | X_{n-1}=i_{n-1},\, \ldots,\, X_0=i_0\right) P\left(X_{n-1}=i_{n-1},\, \ldots,\, X_0=i_0\right)\\
&= P\left(X_n=i_n | X_{n-1}=i_{n-1}\right) P\left(X_{n-1}=i_{n-1},\, \ldots,\, X_0=i_0\right)\\
&= p_{i_{n-1},i_n} P\left(X_{n-1}=i_{n-1},\, \ldots,\, X_0=i_0\right)\\
&= \left(\prod_{k=0}^{n-1} p_{i_k i_{k+1}}\right) \mu_{i_0}.
\end{aligned}$$

Thus, a MC is probabilistically specified by its initial distribution $\{\mu_i\}$ and one step TPM. †

5.3 Classification of States

Definition 5.6 (Accessibility)**.** *State j is said to be accessible from state i if there exists an $n \geq 0$ such that $p_{i,j}^{(n)} > 0$, where*

$$p_{i,j}^{(0)} = \delta_{i,j} = \begin{cases} 1 & \textit{if } i = j \\ 0 & \textit{otherwise.} \end{cases}$$

We will use the notation $i \to j$ to denote the fact j is accessible from i. State j is said to be not accessible from state $i\ (\neq j)$ if $p_{i,j}^{(n)} = 0$ for all $n > 1$.

Remark 5.2. If state j is not accessible from state i, then

$$P\left(\text{Ever be in } j | \text{starting from } i\right) = P\left(\cup_{n=1}^{\infty}\{X_n=j\} | X_0=i\right) \leq \sum_{n=1}^{\infty} P\left(X_n=j | X_0=i\right) = 0.$$

Thus, starting from state i, the chain will not visit state j if state j is not accessible from state i. †

Definition 5.7 (Communication)**.** *Two states i and j are said to communicate if i and j are accessible from each other, i.e., if there exist $m \geq 0$ and $n \geq 0$ such that $p_{i,j}^{(n)} > 0$ and $p_{j,i}^{(m)} > 0$. We will use the notation $i \leftrightarrow j$ to denote i and j communicate.*

Theorem 5.2. *Communication is an equivalence relation, i.e.,*

1. *(Reflexivity) $i \leftrightarrow i$.*
2. *(Symmetry) $i \leftrightarrow j \iff j \leftrightarrow i$.*
3. *(Transitivity) $i \leftrightarrow k$ and $k \leftrightarrow j \implies i \leftrightarrow j$.*

Hence, it partitions the state space into equivalence classes. This equivalence classes are also called communicating classes.

Proof: 1. $p_{i,i}^{(0)} = 1 > 0$. Thus, $i \leftrightarrow i$.

2. Trival.

3. As $i \rightarrow k$ and $k \rightarrow j$, there exists $m \geq 0$ and $n \geq 0$ such that $p_{i,k}^{(m)} > 0$ and $p_{k,j}^{(n)} > 0$. Now, using Chapman-Kolmogrov equations

$$p_{i,j}^{(m+n)} = \sum_{l \in S} p_{i,l}^{(m)} p_{l,k}^{(n)} \geq p_{i,k}^{(m)} p_{k,j}^{(n)} > 0.$$

Thus, $i \rightarrow j$. Similarly, we can prove that $j \rightarrow i$. Hence, $i \leftrightarrow j$.

□

Definition 5.8 (Irreducibility)**.** *A MC is said to be irreducible if all states communicate with each other, i.e., there is a single communicating class.*

Example 5.11. Let a MC with state space $\{1, 2, 3\}$ has the following TPM

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

We want to check if the MC is irreducible or not. In this case it is easy to see that the MC is irreducible. For example, we can go form state 1 to state 3 in 2 steps. However, normally it is very helpful to draw a picture which depicts the communications through states. In the

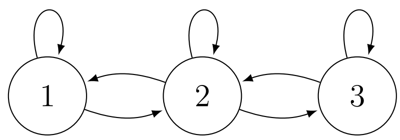
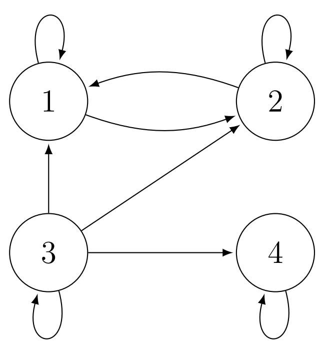


Figure 5.1: Figure for Example 5.11

Figure 5.1, there is a directed edge from state i to state j if $p_{i,j} > 0$. Now, this graph shows that we can move from one state to other states of the MC. This movement may not be in one step. For example, state 3 can not be reached form state 1 in one step. However, there is a positive probability that one can go to state 3 from state 1 in two steps. ||

Example 5.12. Let a MC with state space $\{1, 2, 3, 4\}$ has the following TPM

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



From the figure, it is clear that there are three communicating classes, *viz.*, $\{1, 2\}$, $\{3\}$, and

$\{4\}$. Note that state 1 or 2 are accessible form state 3, but the reverse is not true. On the other hand, state 4 is absorbing state (i.e., $p_{4,4} = 1$) and hence on other state is accessible from state 4. ||

Definition 5.9 (Hitting Time). *For any $A \subset S$, the hitting time or first passage time T_A is defined by*

$$T_A = \inf\left\{n \geq 1 : X_n \in A\right\},$$

with the convention that $\inf \phi = \infty$*. Thus, T_A is the first time, after time 0, when the process enters A. For simplicity, we will use T_i instead of $T_{\{i\}}$ for $i \in S$.*

Definition 5.10 (Reccurent and Transient States). *A state i is called recurrent if $P(T_i < \infty|X_0 = i) = 1$. A state i is called transient if $P(T_i < \infty|X_0 = i) < 1$. Thus, i is recurrent if and only if*

$$f_{ii} = P\left(X_n = i \text{ for some } n \geq 1|X_0 = i\right) = 1.$$

As $P\left(T_i = \infty|X_0 = i\right) = 1 - P\left(T_i < \infty|X_0 = i\right)$, a state i is recurrent if and only if

$$P\left(T_i = \infty|X_0 = i\right) = 0$$

and transient if and only if

$$P\left(T_i = \infty|X_0 = i\right) > 0.$$

Loosely speaking, conditioning on $X_0 = i$, T_i is a proper RV if state i is recurrent state. However, T_i is not a proper RV (in the sense that T_i takes infinite value with positive probability) if state i is a transient state. Thus, $E\left(T_i|X_0 = i\right)$ is infinite when state i is transient. On the other hand, $E\left(T_i|X_0 = i\right)$ could be finite or infinite as T_i is non-negative proper RV. Thus, we have the following definitions of null and positive recurrent.

Definition 5.11 (Null and Positive Cecurrent States). *A recurrent state i is called null recurrent if $E(T_i|X_0 = i) = \infty$ and positive recurrent if $E(T_i|X_0 = i) < \infty$.*

Loosely speaking, for a positive RV X, $E\left(X\right)$ exists if X does not take large values with high probability and (X) does not exists if X takes large values with high probability. Thus, a state i is positive recurrent means that the chain will come back to state i quickly (with high probability). The returing times are large (with high probability) for a null recurrent states.

To discuss the next example, we need two results. One of them is from real analysis and another from probability. The proof of the result from real analysis is skipped here. However, the proof of the result from probability is provided.

Theorem 5.3. *If $0 \leq q_n < 1$, then $\prod_{n=1}^{\infty}(1 - q_n) \to l \neq 0$ if and only if $\sum_{n=1}^{\infty} q_n$ converges.*

Proof: The proof is skipped. □

Theorem 5.4. *Let X be a DRV with support $\{1, 2, 3, \ldots\}$. Then*

$$E\left(X\right) = 1 + \sum_{k=1}^{\infty} P\left(X > k\right).$$

Proof:

$$\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} x P(X=x) \\
&= \sum_{x=1}^{\infty} \left(\sum_{k=1}^{x} 1 \right) P(X=x), \quad \text{as } x = \sum_{k=1}^{x} 1 \\
&= \sum_{k=1}^{\infty} \sum_{x=k}^{\infty} P(X=x), \quad \text{interchanging the order of summation} \\
&= \sum_{k=1}^{\infty} P(X \geq k) \\
&= 1 + \sum_{k=1}^{\infty} P(X > k).
\end{aligned}$$

□

Example 5.13 (Frog in the Well)**.** Let the state space be $S = \{1, 2, \ldots\}$. For $i \geq 1$ and $0 < \alpha_i < 1$, the one-step transition probabilities are given by

$$p_{i,i+1} = \alpha_i \quad \text{and} \quad p_{i,1} = 1 - \alpha_i. \tag{5.1}$$

This can be viewed as a model for movement of a frog in a well. Suppose that the frog is trying to jump up. In the process, if the frog is successful in a jump, it reaches one step ahead in the upward direction. If it it is not successful, it slips to the bottom of the well. Denoting the event that the frog is at ith step of the well by the state i (bottom is state 1), and assuming that the probability of success is α_i when it jumps from the step i, the position of the frog is a MC with transition probabilities given by (5.1).

Here, we want to study the state 1 and shall try to find out the conditions under which the state 1 is transient, null recurrent or positive recurrent. Now,

$$P(T_1 > k | X_0 = 1) = \alpha_1 \alpha_2 \ldots \alpha_k.$$

To see it, notice that starting from floor, $T_1 > k$ means that the first k steps are taken in upward direction. Therefore, as $A_k = \{T_1 > k\}$ is an decreasing sequence of events,

$$P(T_1 = \infty | X_0 = 1) = \lim_{k \to \infty} P(T_1 > k | X_0 = 1) = \lim_{k \to \infty} \alpha_1 \alpha_2 \ldots \alpha_k.$$

As $0 < \alpha_k < 1$ for all $k \geq 1$, $\{x_n = \alpha_1 \alpha_2 \ldots \alpha_n : n \geq 1\}$ is a decreasing sequence and $0 < x_n < 1$ for all $n \geq 1$. Thus, x_n converges and $0 \leq l = \lim_{n \to \infty} x_n \leq 1$. Now, using Theorem 5.3, $l = 0$ if and only if $\sum_{n=1}^{\infty} (1 - \alpha_n)$ diverges, *i.e.*, $\sum_{n=1}^{\infty} (1 - \alpha_n) = \infty$. Thus, the state 1 is recurrent if and only if $\sum_{n=1}^{\infty} (1 - \alpha_n) = \infty$. The state 1 is transient if and only if $\sum_{n=1}^{\infty} (1 - \alpha_n)$ converges.

Using Theorem 5.4,

$$E(T_i | X_0 = i) = 1 + \sum_{k=1}^{\infty} \alpha_1 \alpha_2 \ldots \alpha_k.$$

Thus, the state 1 is positive recurrent if and only if $\sum_{n=1}^{\infty} (1 - \alpha_n) = \infty$ and $\sum_{k=1}^{\infty} \alpha_1 \alpha_2 \ldots \alpha_k < \infty$. The state 1 is a null recurrent if and only if $\sum_{n=1}^{\infty} (1 - \alpha_n) = \infty$ and $\sum_{k=1}^{\infty} \alpha_1 \alpha_2 \ldots \alpha_k = \infty$.

Now, let us consider some specific choices of α_i's. If $\alpha_i = 1 - \frac{1}{2i^2}$ for all $i = 1, 2, 3, \ldots$, then the state 1 is transient. To see it notice that

$$\sum_{n=1}^{\infty}(1-\alpha_n) = \sum_{n=1}^{\infty} \frac{1}{2n^2} < \infty.$$

If $\alpha_i = \alpha$ for all $i = 1, 2, 3, \ldots$, then the state 1 is positive recurrent as

$$\sum_{k=1}^{\infty}(1-\alpha_k) = \sum_{k=1}^{\infty}(1-\alpha) = \infty$$

and

$$\sum_{k=1}^{\infty} \alpha_1\alpha_2\ldots\alpha_k = \sum_{k=1}^{\infty} \alpha^k < \infty.$$

If $\alpha_i = 1 - \frac{1}{2i}$ for all $i = 1, 2, 3, \ldots$, then the state 1 is null recurrent. To see it, note that

$$\sum_{n=1}^{\infty}(1-\alpha_n) = \sum_{n=1}^{\infty} \frac{1}{2n} = \infty.$$

Now, we shall show that $\sum_{k=1}^{\infty} \alpha_1\alpha_2\ldots\alpha_k$ diverges as

$$\alpha_1\alpha_2\ldots\alpha_k > \frac{1}{2k} \quad \text{for all } k > 1$$

We shall prove the claim by induction. Note that $\alpha_1\alpha_2 = \frac{3}{8} > \frac{1}{4}$. Thus, the claim is true for $k = 2$. Now, now assuming that the claim is true for $k = n$, we need to show that the claim is true for $k = n + 1$.

$$\alpha_1\alpha_2\ldots\alpha_n\alpha_{n+1} > \frac{\alpha_{n+1}}{2n} = \frac{2n+1}{4n(n+1)} > \frac{1}{2(n+1)}.$$

Thus, our claim is true for all values of $k > 1$. ||

Theorem 5.5 (Strong Markov Property). *For any state i, any initial distribution $\mu = \{\mu_i\}$, any $k < \infty$, and any states $i_1, i_2, \ldots, i_k$,*

$$P_\mu\left(X_{T_i+j} = i_j,\, j = 1, 2, \ldots, k,\, T_i < \infty\right) = P_\mu\left(T_i < \infty\right) P\left(X_j = i_j,\, j = 1, 2, \ldots, k | X_0 = i\right).$$

Proof: For $n \in \mathbb{N}$,

$$\begin{aligned}
&P_\mu\left(X_{T_i+j} = i_j,\, 1 \le j \le k,\, T_i = n\right) \\
&= P_\mu\left(X_{n+j} = i_j,\, 1 \le j \le k, X_n = i,\, X_r \ne i,\, 1 \le r \le n-1\right) \\
&= P_\mu\left(X_{n+j} = i_j,\, 1 \le j \le k | X_n = i,\, X_r \ne i,\, 1 \le r \le n-1\right) \\
&\qquad \times P_\mu\left(X_n = i,\, X_r \ne i,\, 1 \le r \le n-1\right) \\
&= P\left(X_j = i_j,\, 1 \le j \le k | X_0 = i\right) P_\mu\left(T_i = n\right).
\end{aligned}$$

Now, adding over n, we get

$$\begin{aligned}
&\sum_{n=1}^{\infty} P_\mu\left(X_{T_i+j} = i_j,\, 1 \le j \le k,\, T_i = n\right) = P\left(X_j = i_j,\, 1 \le j \le k | X_0 = i\right) \sum_{n=1}^{\infty} P_\mu\left(T_i = n\right) \\
&\implies P_\mu\left(X_{T_i+j} = i_j,\, 1 \le j \le k,\, T_i < \infty\right) = P\left(X_j = i_j,\, 1 \le j \le k | X_0 = i\right) P_\mu\left(T_i < \infty\right).
\end{aligned}$$

□

The above theroem says that after each return to state i, the MC starts afresh.

Definition 5.12. *Let i be a state. Define $T_i^{(0)} = 0$ and for $k \geq 0$*

$$T_i^{(k+1)} = \begin{cases} \inf\left\{n : n > T_i^{(k)},\, X_n = i\right\} & \text{if } T_i^{(k)} < \infty \\ \infty & \text{otherwise.} \end{cases}$$

Note that $T_i^{(1)} = T_i$ is the first time when the chain visits the state i. $T_i^{(2)}$ is the second time when the chain visits the state i. In general, $T_i^{(k)}$ is the kth time when the chain visits the state i.

Theorem 5.6. *Let i be a recurrent state. Then for all $k \geq 0$,*

$$P\left(T_i^{(k)} < \infty | X_0 = i\right) = 1.$$

Proof: We shall use induction to prove the theorem. By definition of recurrence, the claim is true for $k = 1$. Now assume that the clain is true for $k = n$. Then

$$\begin{aligned} P\left(T_i^{(n+1)} < \infty | X_0 = i\right) &= \sum_{m=n}^{\infty} P\left(T_i^{(n+1)} < \infty,\, T_i^{(n)} = m | X_0 = i\right) \\ &= \sum_{m=n}^{\infty} P\left(T_i^{(n+1)} < \infty | T_i^{(n)} = m,\, X_0 = i\right) P\left(T_i^{(n)} = m | X_0 = i\right) \\ &= \sum_{m=n}^{\infty} P\left(T_i^{(1)} < \infty | X_0 = i\right) P\left(T_i^{(n)} = m | X_0 = i\right) \\ &= P\left(T_i < \infty | X_0 = i\right) P\left(T_i^{(n)} < \infty | X_0 = i\right) \\ &= 1. \end{aligned}$$

□

Definition 5.13 (Cycles)**.** *Let $\eta_r = \left\{X_j : T_i^{(r)} \leq j < T_i^{(r+1)},\, T_i^{(r+1)} - T_i^{(r)}\right\}$ for $r = 0,\, 1,\, \ldots$. The η_r's are called cycles or excursions.*

Theorem 5.7. *Let i be a recurrent state. Under $X_0 = i$, the sequence $\{\eta_r\}_{r=0}^{\infty}$ are i.i.d. as random vectors with a random number of components, i.e., for any $k \in \mathbb{N}$,*

$$\begin{aligned} &P\left(\eta_r = (x_{r0}, x_{r1}, \ldots, x_{rj_r}, j_r), r = 0, 1, \ldots, k | X_0 = i\right) \\ &\quad = \prod_{r=0}^{k} P\left(\eta_1 = (x_{r0}, x_{r1}, \ldots, x_{rj_r}, j_r) | X_0 = i\right), \end{aligned}$$

for any states $x_{r0}, x_{r1}, \ldots, x_{rj_r}$ and any time j_r, $r = 0, 1, \ldots, k$.

Proof: This theorem can be proved applying the strong Markov property repeatedly. □

Theorem 5.8 (Number of Visits)**.** *For any state i, let N_i be the number of visits to the state i after time 0. Then,*

1. *i recurrent implies $P(N_i = \infty | X_0 = i) = 1$.*

2. *i transient implies $P(N_i = n|X_0 = i) = f_{ii}^n(1 - f_{ii})$ for $n = 0, 1, 2, \ldots,$ where $f_{ii} = P(T_i < \infty|X_0 = i)$ is the probability of returning to state i starting from state i. Thus, $N_i|X_0 = i \sim Geo(1 - f_{ii})$.*

Proof: Follows from the strong Markov property. □

Theorem 5.9. *A state i is recurrent if and only if*

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty.$$

A state i is transient if and only if

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty.$$

Proof: Let us define

$$\delta_{X_n,i} = \begin{cases} 1 & \text{if } X_n = i \\ 0 & \text{otherwise,} \end{cases}$$

for all $n > 0$. Then $N_i = \sum_{n=1}^{\infty} \delta_{X_n,i}$ and

$$E\left(N_i|X_0 = i\right) = E\left(\sum_{n=1}^{\infty} \delta_{X_n,i}|X_0 = i\right) = \sum_{n=1}^{\infty} E\left(\delta_{X_n,i}|X_0 = i\right) = \sum_{n=1}^{\infty} p_{i,i}^{(n)}.$$

Now, if the state i is recurrent, then $P\left(N_i = \infty|X_0 = i\right) = 1$ and hence $E\left(N_i|X_0 = i\right) = \infty$. Thus, $\sum_{n=1}^{\infty} p_{i,i}^{(n)} = \infty$ if the state i is recurrent. On the other hand, if state i is tansient, then $E\left(N_i|X_0 = i\right) = \frac{f_{i,i}}{1-f_{i,i}}$. Thus, $\sum_{n=1}^{\infty} p_{i,i}^{(n)} < \infty$ if the state i is transient.

Now, assume that $\sum_{n=1}^{\infty} p_{i,i}^{(n)} = \infty$. If possible, assume that the state i is transient. Then $\sum_{n=1}^{\infty} p_{i,i}^{(n)} < \infty$, which is a contradiction. Hence, our assumption is wrong and the state i is rucurrent. Similarly, $\sum_{n=1}^{\infty} p_{i,i}^{(n)} < \infty$ implies that the state i is recurrent. □

Remark 5.3. $P(X_n = i \,\text{for infinitely many}\, n|X_0 = i) = 1 \,\text{or}\, 0$ if and only if recurrent or transient. †

Theorem 5.10. *If the state space S is finite, then at least one state must be recurrent.*

Proof: Let the state space $S = \{1,\ 2,\ \ldots,\ K\}$, where $K < \infty$. Since, $n = \sum_{i=1}^{K} \sum_{j=1}^{n} \delta_{X_j,i}$, there exists an i_0 such that as $n \to \infty$, $\sum_{j=1}^{n} \delta_{X_j,i_0} \to \infty$ with positive probability. This implies that i_0 must be recurrent. □

Theorem 5.11. *Let $i \leftrightarrow j$. If i is transient, then j is transient. If i is recurrent, then j is recurrent.*

Proof: As i and j communicate, there exist $m > 0$ and $n > 0$ such that $p_{i,j}^{(m)} > 0$ and $p_{j,i}^{(n)} > 0$. Then using Chapman-Kolmogorov equations

$$p_{i,i}^{(m+k+n)} > p_{i,j}^{(m)} p_{j,j}^{(k)} p_{j,i}^{(n)} \implies \sum_{k=1}^{\infty} p_{i,i}^{(m+k+n)} > p_{i,j}^{(m)} p_{j,i}^{(n)} \sum_{k=1}^{\infty} p_{j,j}^{(k)}.$$

Now, if the state i is transient, then $\sum_{k=1}^{\infty} p_{i,i}^{(m+k+n)} < \infty$, which implies $\sum_{k=1}^{\infty} p_{j,j}^{(k)} < \infty$. Hence, the state j is transient. Also,

$$\sum_{k=1}^{\infty} p_{j,j}^{(n+k+m)} > p_{j,i}^{(n)} p_{i,j}^{(m)} \sum_{k=1}^{\infty} p_{i,i}^{(k)}.$$

Now, if i is recurrent, then $\sum_{k=1}^{\infty} p_{i,i}^{(k)} = \infty$, which implies $\sum_{k=1}^{\infty} p_{j,j}^{(k)} = \infty$. Hence, the state j is recurrent. □

Theorem 5.12 (Solidarity property). *Let i be recurrent and $i \to j$. Then*

$$f_{j,i} = P\left(T_i < \infty | X_0 = j\right) = 1$$

and j is recurrent.

Proof: By the strong Markov property,

$$\begin{aligned}
1 - f_{i,i} &= P\left(T_i = \infty | X_0 = i\right) \\
&\geq P\left(T_j < \infty,\, T_i = \infty | X_0 = i\right) \\
&= P\left(T_j < T_i,\, T_i = \infty | X_0 = i\right) \\
&= P\left(T_i = \infty | T_j < T_i,\, X_0 = i\right) P\left(T_j < T_i | X_0 = i\right) \\
&= P\left(T_i = \infty | X_0 = j\right) P\left(T_j < T_i | X_0 = i\right) \\
&= \left(1 - f_{j,i}\right) f_{i,j}^{*},
\end{aligned}$$

where $f_{i,j}^{*} = P\left(T_j < T_i | X_0 = i\right) = P\left(\text{visiting } j \text{ before visiting } i | X_0 = i\right)$. Now, i recurrent and $i \to j$ yield $1 - f_{i,i} = 0$ and $f_{i,j}^{*} > 0$ and so $1 - f_{j,i} = 0$, *i.e.*, $f_{j,i} = 1$. Thus, starting from j, the chain visits i with probability 1. From i, it keeps returning to i infinitely often. In each of these excursions, $f_{i,j}^{*} > 0$ and since there are infinite number of such excursions and they are *i.i.d.*, the chain does visit j in one of these excursions with probability 1. That is j is recurrent. □

Theorem 5.13. *Suppose that $\{X_n\}$ is irreducible and recurrent. Then for all $i \in S$, $P_\mu\left(T_i < \infty\right) = 1$ for any initial distribution μ.*

Proof:

$$P_\mu\left(T_i < \infty\right) = \sum_{j \in S} P\left(T_i < \infty | X_0 = j\right) \mu_j = \sum_{j \in S} \mu_j = 1.$$

The second equality is true due to the previous theorem. □

Example 5.14. Let the MC consisting of the states 0, 1, 2, and 4 have the TPM

$$P = \begin{bmatrix} 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

It is clear that all states communicate, *i.e.*, it is a irreducible MC. As it has finite state space, at least one state is recurrent. As it is an irreducible MC, all the states are recurrent. ||

Example 5.15. Condiser the MC having states 0, 1, 2, 3, and 4 and TPM

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 0 & 0 & 1/2 \end{bmatrix}.$$

This MC consists of three communicating classes, *viz.*, $\{0,\, 1\}$, $\{2, 3\}$ and $\{4\}$. Note that $p_{i,i}^{(n)} = \frac{1}{2}$ for all $i = 0,\, 1,\, 2,\, 3$ and for all $n \geq 1$. This shows that $\sum_{n=1}^{\infty} p_{i,i}^{(n)} = \infty$ for $i = 0,\, 1,\, 2,\, 3$, and hence states 0, 1, 2, and 3 are recurrent states. Now, if possible, suppose that the state 4 is recurrent. Note that $4 \to 1$, but 4 is not accessible from 1, which contradicts the solidarity property. Thus, state 4 is transient. ||