6. Stochastic processes and its classes

The theory of stochastic processes turns to be a useful tool in solving problems in various fields such as engineering, genetics, statistics, economics, finance, etc. The word *stochastic* means *random* or *chance*. Stochastic processes can be thought of collection of random variables indexes by some parameter.

Definition 6.1 (**Stochastic process).** A real stochastic process is a collection of random variables $\{X_t : t \in \mathbf{T}\}$ defined on a common probability space $(\Omega, \mathbb{P}, \mathcal{F})$ with values in $\mathbb{R}$.

- $\mathbf{T}$ is called the index set of the process which is usually a subset of $\mathbb{R}$.
- The set of values that the random variable X_t can take is called state space, denoted by $\mathcal{S}$.

Example 6.1. *Consider a random event occurring in time such as number of telephone calls received at a switchboard. Suppose that X_t is the random variable which represents the number of incoming calls in an interval $(0,t)$ of duration t units. The number of calls within a fixed interval of specified duration, say one unit of time, is a random variable X_1 and the family $\{X_t; t \in \mathbf{T}\}$ constitutes a stochastic process.*

Stochastic processes can be classified, in general, into the following four types of processes:
a) **Discrete time, discrete state space**: the number of individuals in a population at the end of year t can be modeled as a stochastic process $\{X_t; t \in \mathbf{T}\}$, where $\mathbf{T} = \{0, 1, 2, \ldots\}$ and the state space $\mathcal{S} = \{0, 1, 2, \ldots\}$.

b) **Discrete time, continuous state space**: the share price for an asset at the close of trading on day t with $\mathbf{T} = \{0, 1, 2, \ldots\}$ and $\mathcal{S} = \{x : 0 \le x < \infty\}$.

c) **Continuous time, discrete state space**: the number of incoming calls X_t in an interval $[0,t]$. The stochastic process $\{X_t; t \in \mathbf{T}\}$ has $\mathbf{T} = \{t : 0 \le t < \infty\}$ and the state space $\mathcal{S} = \{0, 1, 2, \ldots\}$.

d) **Continuous time, continuous state space**: the value of the Dow-Jones index (stock market index) at time t. Then The stochastic process $\{X_t; t \in \mathbf{T}\}$ has $\mathbf{T} = \{t : 0 \le t < \infty\}$ and the state space $\mathcal{S} = \{x : 0 \le x < \infty\}$.

Definition 6.2 (**Independent Increments).** Let $\{X_t : t \in \mathbf{T}\}$ be a stochastic process. We say that it has independent increments, if for all $t_0 < t_1 < t_2 < \ldots < t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.

Definition 6.3 (**Stationary Increments).** A stochastic process $\{X_t : t \in \mathbf{T}\}$ is said to have stationary increments if $X_{t_2+\tau} - X_{t_1+\tau}$ has the same distribution as $X_{t_2} - X_{t_1}$ for all choices of t_1, t_2 and $\tau > 0$.

Definition 6.4 (**Second order process).** A stochastic process $\{X_t : t \in \mathbf{T}\}$ is called a second order process if $\mathbb{E}[X_t^2] < \infty$ for all $t \in \mathbf{T}$.

Example 6.2. *Let Z_1, Z_2 be independent normally distributed random variables with mean 0 and variance σ^2. For $\lambda \in \mathbb{R}$, define*

$$X_t = Z_1 \cos(\lambda t) + Z_2 \sin(\lambda t), \quad t \in \mathbb{R}^+.$$

Then X_t is a second-order stochastic process. Indeed, by using independent property,

$$\begin{aligned}\mathbb{E}[|X_t|^2] &= \mathbb{E}\big[Z_1^2 \cos^2(\lambda t) + Z_2^2 \sin^2(\lambda t) + 2 Z_1 Z_2 \cos(\lambda t) \sin(\lambda t)\big] \\ &= \mathbb{E}[Z_1^2] \cos^2(\lambda t) + \mathbb{E}[Z_2^2] \sin^2(\lambda t) + \sin(2\lambda t) \mathbb{E}[Z_1 Z_2]\end{aligned}$$

$$= \sigma^2(\cos^2(\lambda t) + \sin^2(\lambda t)) = \sigma^2 \quad \forall t \in \mathbb{R}^+.$$

Definition 6.5 (**Covariance Stationary Process).** A second order stochastic process $\{X_t : t \in \mathbf{T}\}$ is called *covariance stationary process* if

i) $m(t) = \mathbb{E}[X_t]$ is independent of t.
ii) $\mathrm{Cov}(X_t, X_s)$ depends only on the difference $|t-s|$ for all $s, t \in \mathbf{T}$.

Example 6.3. *Let $\{X_n : n \geq 1\}$ be uncorrelated random variables with mean 0 and variance 1. Then*

$$\mathrm{Cov}(X_m, X_n) = \mathbb{E}[X_m X_n] = \begin{cases} 0, & m \neq n \\ 1\,, & m = n\,. \end{cases}$$

Thus, $\{X_n : n \geq 1\}$ is a covariance stationary process.

Example 6.4. *Let A be a random variable with mean m and variance σ^2. For fixed $\eta \in \mathbb{R}$, consider the stochastic process*

$$X_t := A\sin(\eta t + \Phi), \quad t \geq 0$$

where Φ is uniformly distributed over $(0, 2\pi)$ and independent of A. Then $\{X_t : t \geq 0\}$ is a covariance stationary process. To check this, we first compute the mean.

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[A\sin(\eta t + \Phi)] = \mathbb{E}[A]\mathbb{E}[\sin(\eta t + \Phi)] = m\mathbb{E}[\sin(\eta t + \Phi)] \\ &= \frac{m}{2\pi}\int_0^{2\pi} \sin(\eta t + \theta)\, d\theta = 0. \end{aligned}$$

Next we calculate the covariance of X_t and X_s.

$$\begin{aligned} &\mathrm{Cov}(X_t, X_s) = \mathbb{E}[X_t X_s] = \frac{1}{2}\mathbb{E}[A^2 2\sin(\eta t + \Phi)\sin(\eta s + \Phi)] \\ &= \frac{1}{2}\mathbb{E}[A^2]\mathbb{E}[\cos(\eta(t-s)) - \cos(\eta(t+s) + 2\Phi)] \quad (\because 2\sin(a)\sin(b) = \cos(a-b) - \cos(a+b)) \\ &= \frac{1}{2}\mathbb{E}[A^2]\Big(\cos(\eta(t-s)) - \frac{1}{2\pi}\int_0^{2\pi}\cos(\eta(t+s) + 2\theta)\, d\theta\Big) \\ &= \frac{1}{2}\mathbb{E}[A^2]\cos(\eta(t-s)) = \frac{1}{2}(\sigma^2 + m^2)\cos(\eta(t-s)). \end{aligned}$$

Thus, $\mathrm{Cov}(X_t, X_s)$ depends only on the difference $|t-s|$. Next we show that it is a second ordet process. Indeed,

$$\begin{aligned} \mathbb{E}[X_t^2] &= \mathbb{E}[A^2]\mathbb{E}[\sin^2(\eta t + \Phi)] = (\sigma^2 + m^2)\frac{1}{2\pi}\int_0^{2\pi}\sin^2(\eta t + \theta)\, d\theta \\ &= (\sigma^2 + m^2)\frac{1}{4\pi}\int_0^{2\pi}\big(1 - \cos(2\eta t + 2\theta)\big)\, d\theta \quad (\because 2\sin^2(a) = 1 - \cos(2a)) \\ &= \frac{\sigma^2 + m^2}{2} < +\infty \quad \forall\, t. \end{aligned}$$

Thus, $\{X_t : t \geq 0\}$ is a covariance stationary process.

Definition 6.6 (**Markov Process).** Let $\{X_t : t \geq 0\}$ be a stochastic process defined over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with state space $(\mathbb{R}, \mathcal{B})$. We say that $\{X_t : t \geq 0\}$ is a Markov process if for any $0 \leq t_1 < t_2 < \ldots < t_n$ and for any $B \in \mathcal{B}$

$$\mathbb{P}\Big(X_{t_n} \in B | X_{t_1}, \ldots, X_{t_{n-1}}\Big) = \mathbb{P}\Big(X_{t_n} \in B | X_{t_{n-1}}\Big)\,.$$

Roughly, a Markov process is a process such that given the value X_s, the distribution of X_t for $t > s$, does not depend on the values of $X_u, u < s$.

Remark 6.1. Any stochastic process which has independent increments is a Markov process.

7. Markov Chain

Definition 7.1 (Markov Chain). A sequence of random variables $(X_n)_{n\in\mathbb{N}}$ with discrete state space (countable or finite) is called a discrete-time Markov chain (**DTMC**) if it satisfies the following condition

$$\mathbb{P}\Big(X_{n+1} = j \Big| X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big) = \mathbb{P}\big(X_{n+1} = j \big| X_n = i\big) \tag{7.1}$$

for all $n \in \mathbb{N}$ and for all $i_0, i_1, \dots, i_{n-1}, i, j \in S$ (state space) with $\mathbb{P}\big(X_0 = i_0, \dots, X_n = i_n\big) > 0$. In other words, the condition (7.1) implies the following: if we know the present state "$X_n = i$", the knowledge of the past history "$X_{n-1}, \dots, X_0$" has no influence on the probability structure of the future state X_{n+1}.

Example 7.1. *Suppose that a coin is tossed repeatedly and let X_n denotes the number of heads in the first n-tossed. Then the number of heads obtained in the first $(n+1)$ tossed only depends on the knowledge of the number of heads obtained in the first n tosses. Thus, (7.1) is satisfied. Hence it is a Markov chain.*

Example 7.2 (A simple Queueing Model). *Let $0, 1, 2, \dots$ be the times at which an elevator starts. It is assume that only one person can be transported by the elevator at a time. Between the times n and $n+1$, let Y_n denotes the no. of people who wants to get into the elevator arrive. Assume that $Y_n's$ are independent. The Queue length X_n immediately before the start of the elevator at time n is equal to*

$$X_n = \max\{0, X_{n-1} - 1\} + Y_{n-1} \quad n \ge 1\,.$$

Suppose $X_0 = 0$. Since X_i with $i \le n$ can be expressed in terms of $Y_1, Y_2, \dots, Y_{n-1}$, we have that Y_n is independent of $(X_0, X_1, \dots, X_n)$ as well. Thus, if $i_n \ge 1$, then

$$\begin{aligned}
&\mathbb{P}\Big(X_{n+1} = i_{n+1} \Big| X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big) \\
&= \frac{\mathbb{P}\Big(X_{n+1} = i_{n+1}, X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big)}{\mathbb{P}\Big(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big)} \\
&= \frac{\mathbb{P}(Y_n = i_{n+1} - i_n + 1)\mathbb{P}\Big(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big)}{\mathbb{P}\Big(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big)} \\
&= \mathbb{P}\big(Y_n = i_{n+1} - i_n + 1\big) = \mathbb{P}\big(X_{n+1} = i_{n+1} \big| X_n = i_n\big)\,.
\end{aligned}$$

If $i_n = 0$, then $X_{n+1} = Y_n$, and in this case

$$\begin{aligned}
&\mathbb{P}\Big(X_{n+1} = i_{n+1} \Big| X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big) \\
&= \frac{\mathbb{P}\Big(Y_n = i_{n+1}, X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big)}{\mathbb{P}\Big(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big)} \\
&= \mathbb{P}\big(Y_n = i_{n+1}\big) = \mathbb{P}\big(X_{n+1} = i_{n+1} \big| X_n = 0\big)\,.
\end{aligned}$$

Hence $(X_n)_{n\in\mathbb{N}}$ is a Markov chain.

Lemma 7.1. *If $\{X_n : n \ge 0\}$ is a Markov chain, the for all n and all $i_0, i_1, \dots, i_n \in S$, we have*

$$\mathbb{P}\Big(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big) = \mathbb{P}(X_0 = i_0) \prod_{k=1}^{n} \mathbb{P}\Big(X_k = i_k \Big| X_{k-1} = i_{k-1}\Big)\,.$$

Proof. By using Markov property, we see that

$$\mathbb{P}(X_0 = i_0) \prod_{k=1}^{n} \mathbb{P}\Big(X_k = i_k \Big| X_{k-1} = i_{k-1}\Big) = \mathbb{P}(X_0 = i_0) \prod_{k=1}^{n} \frac{\mathbb{P}\Big(X_k = i_k, X_{k-1} = i_{k-1}, \dots, X_0 = i_0\Big)}{\mathbb{P}\Big(X_{k-1} = i_{k-1}, \dots, X_0 = i_0\Big)}$$

$$= \mathbb{P}\Big(X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\Big).$$

□

Definition 7.2. Let $\{X_n : n \ge 0\}$ be a Markov chain.

- The pmf of X_n is defined by $p_j(n) := \mathbb{P}(X_n = j)$. Moreover, we define
$$p_{jk}(m,n) := \mathbb{P}(X_n = k \big| X_m = j) \quad 0 \le m \le n, \ j,k \in \mathcal{S}.$$
- $\{X_n : n \ge 0\}$ is called homogeneous Markov chain if $p_{jk}(m,n)$ only depends on the difference $n - m$.
- We are mainly interested in the homogeneous Markov chain. For a homogeneous Markov chain, we define m-step transition probability , denoted by $p_{jk}^{(m)}$, given by
$$p_{jk}^{(m)} = \mathbb{P}\big(X_{n+m} = k \big| X_n = j\big).$$
$p_{jk}^{(m)}$ gives the probability that from the state j at nth trial, the state k is reached at $(m+n)$th trial in m steps. Define
$$p_{jk}^{(0)} = \mathbb{P}\big(X_n = k \big| X_n = j\big) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \ne k\,. \end{cases}$$
Then the transition matrix $P^{(m)} := (p_{ij}^{(m)})_{i,j \in S}$ is called the m-step transition matrix.
- The probabilities
$$p_{ij} := \mathbb{P}\big(X_{n+1} = j \big| X_n = i\big)$$
are called transition probabilities (of one step).
- The probability distribution $\pi := (\pi)_{i \in S}$ with $\pi_i = \mathbb{P}(X_0 = i)$ is called the initial distribution.
- The matrix $P := (p_{ij})$ is called the transition matrix or a stochastic matrix. Note that $p_{ij} \ge 0$ for all $i,j \in S$ and $\sum_{j \in S} p_{ij} = 1$ for all $i \in S$.

Example 7.3. *Consider the example of simple Queueing Model. Then*

$$\begin{cases} p_{ij} = \mathbb{P}\big(X_{n+1} = j \big| X_n = i\big) = \mathbb{P}\big(Y_n = j - i + 1\big) & \text{if } i > 0, \\ p_{0j} = \mathbb{P}(Y_n = j)\,. \end{cases}$$

Denote

$$p_j := \mathbb{P}(Y_n = j).$$

Then the transition matrix P of the Markov chain is given by

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \cdots \\ p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & p_0 & p_1 & p_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

Remark 7.1. The matrix of transition probabilities together with the initial distribution completely specifies the Markov chain $\{X_n : n \geq 0\}$. For example, let $\{X_n : n \geq 0\}$ is a Markov chain with three states $0, 1, 2$ with the transition matrix

$$P = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix} \quad \text{and the initial distribution } \mathbb{P}(X_0 = i) = \frac{1}{3} \quad i = 0, 1, 2\,.$$

Then we have

$$\begin{aligned} &\mathbb{P}(X_1 = 1 \big| X_0 = 2) = p_{21} = \frac{3}{4}\,, \\ &\mathbb{P}(X_2 = 2 \big| X_1 = 1) = p_{12} = \frac{1}{4}\,, \\ &\mathbb{P}(X_2 = 2, X_1 = 1 \big| X_0 = 2) = \mathbb{P}(X_2 = 2 \big| X_1 = 1)\mathbb{P}(X_1 = 1 \big| X_0 = 2) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}\,, \\ &\mathbb{P}(X_2 = 2, X_1 = 1, X_0 = 2) = \mathbb{P}(X_2 = 2, X_1 = 1 \big| X_0 = 2)\mathbb{P}(X_0 = 2) = \frac{3}{16} \cdot \frac{1}{3} = \frac{1}{16}\,. \end{aligned}$$

Markov chain as a Graphs: Let us explain how Markov chains can be described as graphs. The states of a Markov chain may be represented by the vertices (nodes) of the graph and one step transition between states by directed arcs; if $i \to j$, the vertices i and j are joined by directed arc with arrow towards j. The value of p_{ij} corresponding to the arc weight may be indicated in the directed arc. If $S = \{1, 2, \cdots, m\}$ is the set of vertices corresponding to the state of Markov chain, and "a" is the set of directed arcs between these vertices , then the graph $G = \{S, a\}$ is called transition graph of the chain. The transition graph is of great aid in visualizing a Markov chain; it is useful tool in studying the properties of the chain (e.g., irreducibility).

Remark 7.2. If for some transition, the probability of occurrence is zero, then it indicates that the transition is not possible and the corresponding arc is not drawn.

The transition graph of the previous example is given by Figure 2

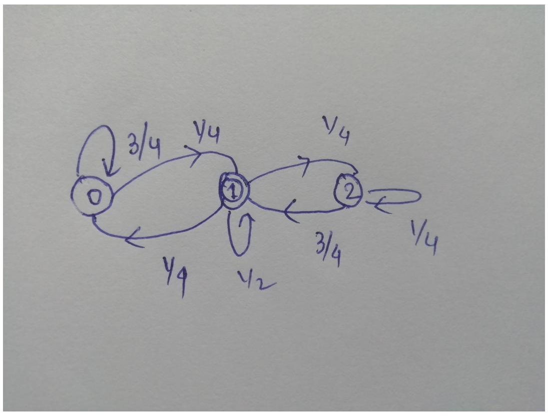


FIGURE 2. Transition garph

7.1. **Chapman-Kolmogorov equation:** The Chapman-Kolmogorov equation provide a procedure to compute the m-step transition probabilities. Before that, we have the following theorem.

Theorem 7.2. *For any Markov chain $\{X_n : n \geq 0\}$, we have for all $h, j \in S$ and $k < m < n$,*
$$\mathbb{P}(X_n = j \big| X_k = h) = \sum_{i \in S} \mathbb{P}(X_n = j \big| X_m = i)\mathbb{P}(X_m = i \big| X_k = h)\,.$$

Proof. By using Lemma 7.1, we get
$$\begin{aligned}
\mathbb{P}\big(X_k = h, X_n = j\big) &= \sum_{i \in S} \mathbb{P}\big(X_k = h, X_m = i, X_n = j\big) \\
&= \sum_{i \in S} \mathbb{P}(X_k = h)\mathbb{P}\big(X_m = i \big| X_k = h\big)\mathbb{P}\big(X_n = j \big| X_m = i\big) \\
\implies \frac{\mathbb{P}\big(X_k = h, X_n = j\big)}{\mathbb{P}(X_k = h)} &= \sum_{i \in S} \mathbb{P}\big(X_m = i \big| X_k = h\big)\mathbb{P}\big(X_n = j \big| X_m = i\big) \\
\implies \mathbb{P}\big(X_n = j \big| X_k = h\big) &= \sum_{i \in S} \mathbb{P}\big(X_m = i \big| X_k = h\big)\mathbb{P}\big(X_n = j \big| X_m = i\big)\,.
\end{aligned}$$

□

In view of Theorem 7.2, for all $n, m \geq 0$ and $i, j \in S$, there holds
$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)}\,. \tag{7.2}$$
In matrix form, we can write the Chapman-Kolmogorov equation (7.2) as
$$P^{(n+m)} = P^{(n)} P^{(m)}.$$
Hence, we have
$$P^{(n)} = P P^{(n-1)} = P \cdot P \cdot P^{(n-2)} = P^n\,,$$
where P is the one step transition matrix.

Corollary 7.3. *Let $\{X_n : n \geq 0\}$ be a Markov chain with initial distribution π and the one step transition matrix P. Then for each $n \geq 1$ and for each $k \in S$*
$$\mathbb{P}(X_n = k) = \sum_{j \in S} p_{jk}^{(n)} \pi_j\,.$$

Proof. It is easy to see that
$$\mathbb{P}(X_n = k) = \sum_{j \in S} \mathbb{P}(X_n = k, X_0 = j) = \sum_{j \in S} \mathbb{P}(X_n = k \big| X_0 = j)\mathbb{P}(X_0 = j) = \sum_{j \in S} p_{jk}^{(n)} \pi_j\,.$$

□

Example 7.4. *Let $\{X_n : n \geq 0\}$ be a Markov chain with state space $\{0, 1\}$, initial distribution $\pi = (\frac{1}{2}, \frac{1}{2})$ and transition matrix*
$$P = \begin{pmatrix} 1/10 & 9/10 \\ 3/10 & 7/10 \end{pmatrix}$$
Then find $\mathbb{P}(X_2 = 1)$.
Solution: *Note that*
$$P^{(2)} = P \cdot P = \begin{pmatrix} 1/10 & 9/10 \\ 3/10 & 7/10 \end{pmatrix} \cdot \begin{pmatrix} 1/10 & 9/10 \\ 3/10 & 7/10 \end{pmatrix} = \begin{pmatrix} 0.28 & 0.72 \\ 0.24 & 0.76 \end{pmatrix}$$
Thus, $p_{01}^{(2)} = 0.72$ and $p_{11}^{(2)} = 0.76$, and hence
$$\mathbb{P}(X_2 = 1) = \frac{1}{2}(0.72 + 0.76) = 0.74\,.$$

Example 7.5. *The transition probability matrix of a discrete-time Markov chain* $\{X_n : n \geq 0\}$ *having state space* $S = \{1, 2, 3\}$ *is*

$$P = \begin{pmatrix} 0.3 & 0.4 & 0.3 \\ 0.6 & 0.2 & 0.2 \\ 0.5 & 0.4 & 0.1 \end{pmatrix}$$

and initial distribution is $\pi = (0.7, 0.2, 0.1)$. *Then find*

a) $\mathbb{P}(X_2 = 3)$.

b) $\mathbb{P}\big(X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2\big)$

Solution: *Let us find the* 2*-step transition matrix* $P^{(2)}$. *Observe that*

$$P^{(2)} = P \cdot P = \begin{pmatrix} 0.3 & 0.4 & 0.3 \\ 0.6 & 0.2 & 0.2 \\ 0.5 & 0.4 & 0.1 \end{pmatrix} \cdot \begin{pmatrix} 0.3 & 0.4 & 0.3 \\ 0.6 & 0.2 & 0.2 \\ 0.5 & 0.4 & 0.1 \end{pmatrix} = \begin{pmatrix} 0.48 & 0.32 & 0.2 \\ 0.4 & 0.36 & 0.24 \\ 0.44 & 0.32 & 0.24 \end{pmatrix}.$$

Now

$$\begin{aligned} &\mathbb{P}(X_2 = 3) = \sum_{j=1}^{3} p_{j3}^{(2)} \pi_j = (0.2)(0.7) + (0.24)(0.2) + (0.24)(0.1) = 0.212\,, \\ &\mathbb{P}\big(X_3 = 2, X_2 = 3, X_1 = 3, X_0 = 2\big) \\ &= \mathbb{P}(X_0 = 2)\mathbb{P}(X_1 = 3 \big| X_0 = 2)\mathbb{P}(X_2 = 3 \big| X_1 = 3)\mathbb{P}(X_3 = 2 \big| X_2 = 3) \\ &= \mathbb{P}(X_0 = 2) p_{23}\, p_{33}\, p_{32} = (0.2)(0.2)(0.1)(0.4) = 0.0016\,. \end{aligned}$$

7.2. **Classification of states:** The state $j \in S$ of a Markov chain $\{X_n : n \geq 0\}$ can be classified in a distinctive manner according to some fundamental properties of the system.

Definition 7.3. Let $\{X_n : n \geq 0\}$ be a Markov chain with state space S.

a) **Accessibility**: State j is said to be accessible from the state i if for some $n \geq 0$, $p_{ij}^{(n)} > 0$. In this case, we write $i \to j$.

b) **Communicate**: If two states i and j are such that each is accessible from the other, then we say that two states communicate, denoted by $i \leftrightarrow j$. If two states i and j communicates, then there exist integres m and n such that $p_{ij}^{(n)} > 0$ and $p_{ji}^{(m)} > 0$.

c) **Absorbing**: A state i is said to be an absorbing state if $p_{ii} = 1$.

Lemma 7.4. *The relation "$\to$" is transitive i.e., if* $i \to j$ *and* $j \to k$, *then* $i \to k$.

Proof. Since $i \to j$ and $j \to k$, there exist integers r and l such that $p_{ij}^{(r)} > 0$ and $p_{jk}^{(l)} > 0$. From the Chapman-Kolmogorov equation, we have

$$p_{ik}^{(l+r)} = \sum_{j \in S} p_{ij}^{(r)} p_{jk}^{(l)} \geq p_{ij}^{(r)} p_{jk}^{(l)} > 0\,.$$

Hence $i \to k$. □

Remark 7.3. One can check easily that the relation $\leftrightarrow$ is transitive, and $i \leftrightarrow j$ is an equivalence relation over S.

Define the equivalence class

$$C(i) := \{j \in S : i \leftrightarrow j\} \quad i \in S\,.$$

Then, it is easy to show that $C(i)$ forms a partition of S.

Definition 7.4 (Irreducible). A Markov chain is said to be **irreducible** if the state space consists of only one class, i.e., all states are communicate with each others.

Example 7.6. *Let $\{X_n : n \geq 0\}$ be a Markov chain with state space $S = \{1, 2, 3\}$ with initial distribution $\pi = (1, 0, 0)$ and transition matrix $P = \begin{pmatrix} 0 & 3/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1 & 0 & 0 \end{pmatrix}$. Then $\{X_n : n \geq 0\}$ is an irreducible Markov chain.*

To see, let us first look at the transition graph of this Markov chain and then calculate equivalence classes. The transition graph is given in Figure 3. Now from the graph and the transitivity

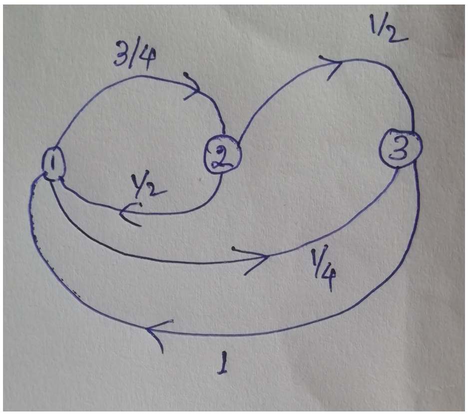


FIGURE 3. Transition graph

of $\to$, we see that $2 \to 3$, $3 \to 1$ and $1 \to 2$ and hence $3 \to 2$. Thus $3 \leftrightarrow 2$. It is easy to see that

$$C(1) = \{j \in S : 1 \leftrightarrow j\} = \{1, 2, 3\}, \quad C(2) = \{1, 2, 3\} = C(3)\,.$$

Hence $\{X_n : n \geq 0\}$ is an irreducible Markov chain.

Definition 7.5. Let $i \in S$ be fixed.

a) **Return state:** The state i such that $p_{ii}^{(n)} > 0$ for some $n \geq 1$ is called return state.
b) The period of i is defined by

$$\lambda(i) = \gcd\{n : p_{ii}^{(n)} > 0\}\,.$$

If $p_{ii}^{(n)} = 0$ for all $n \geq 1$, we define $\lambda(i) = 0$.
c) The state $i \in S$ is called **aperiodic** if $\lambda(i) = 1$.
d) A Markov chain with all states aperiodic is called an **aperiodic Markov chain**.

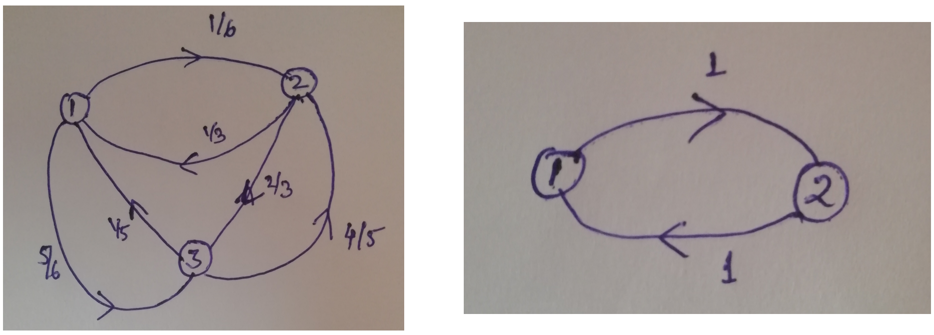


FIGURE 4. Transition graphs

Example 7.7. *Consider the Markov chains whose transition graphs are given in Figure 4. Then first one is aperiodic Markov process, while the period of any element of the state space of Markov chain (right) is* 2.

Theorem 7.5. *If $i \leftrightarrow j$, then $\lambda(i) = \lambda(j)$.*

Proof. Since $i \to j$ and $j \to i$, there exist n_1 and n_2 such that $p_{ij}^{(n_1)} > 0$ and $p_{ji}^{(n_2)} > 0$. Thus, $p_{ii}^{(n_1+n_2)} > 0$. Again, since $j \to j$, there exists $n_3 \geq 0$ such that $p_{jj}^{(n_3)} > 0$. Hence

$$p_{ii}^{(n_1+n_3+n_2)} \geq p_{ij}^{(n_1)} p_{jj}^{(n_3)} p_{ji}^{(n_2)} > 0 \,.$$

Hence $\lambda(i)$ divides both $n_1 + n_2$ and $n_1 + n_2 + n_3$. Thus, $\lambda(i)$ divides n_3 whenever $p_{jj}^{(n_3)} > 0$. Therefore, $\lambda(i)$ divides $\lambda(j)$. A similar argument yields that $\lambda(j)$ divides $\lambda(i)$. Thus, $\lambda(i) = \lambda(j)$. ☐

First visit and mean passage time:

Definition 7.6. For any states i and j, define $f_{ij}^{(n)}$ to be the probability that starting in i , the first transition into state j occurs at time n. In other words,

$$\begin{cases} f_{ij}^{(n)} := \mathbb{P}\big(X_n = j, X_k \neq j, 1 \leq k \leq n-1 \big| X_0 = i\big), \quad n \geq 1\,, \\ f_{ij}^{(0)} = 0\,. \end{cases}$$

Let

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}\,.$$

Then f_{ij} denotes the probability of ever making a transition into state j given that the process starts at state i.

Remark 7.4. It can be shown that

$$\lambda(i) = \gcd\{m : f_{ii}^{(m)} > 0\}\,.$$

A relation between $f_{jk}^{(n)}$ and $p_{jk}^{(n)}$ is as follows.

Theorem 7.6 (First Entrance Theorem)**.** *There holds*

$$p_{jk}^{(n)} = \sum_{r=0}^{n} f_{jk}^{(r)} p_{kk}^{(n-r)}, \quad n \geq 1\,.$$

Thus, $f_{jk}^{(n)} = p_{jk}^{(n)} - \sum_{r=1}^{n} f_{jk}^{(r)} p_{kk}^{(n-r)}$ *with* $f_{jk}^{(1)} = p_{jk}$.

We define the **mean recurrence time/ mean passage time** as

$$\mu_{jk} := \sum_{n=1}^{\infty} n f_{jk}^{(n)}.$$

We are mainly interested in the case $j = k$. In that case $f_{jj}^{(n)}$ represents the distribution of the recurence time of state j. Instead of μ_{jj}, we simple use μ_j. One can view μ_j in the following sense: consider the random variable T_j, the first return time to state j, given by

$$T_j := \inf\{n \geq 1 : X_n = j \big| X_0 = j\}.$$

Then μ_j can be given by

$$\mu_j = \mathbb{E}[T_j].$$

μ_j represents the expected return time of the chain to state j given that the chain started from state j.

Definition 7.7. a) A state $j \in S$ is said to be **recurrent** if $f_{jj} = 1$ i.e., the return to the state j is certain. Otherwise it is **transient**.

b) The recurrent state j is said to be **positive recurrent** if $\mu_j < \infty$. It is said to be **null-recurrent** if $\mu_j = \infty$.

One can check that

$$f_{jj} < 1 \Leftrightarrow \mathbb{P}(T_j = \infty) > 0.$$

Example 7.8. *Consider the Markov chain with state space $S = \{0, 1, 2, 3\}$ and transition matrix*

$$P = \begin{pmatrix} 0.8 & 0 & 0.2 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0 & 0.3 \end{pmatrix}.$$

We want to classify the states of the given Markov chain. The transition graph of the Markov chain is given in Figure 5. Observe from the transition graph,

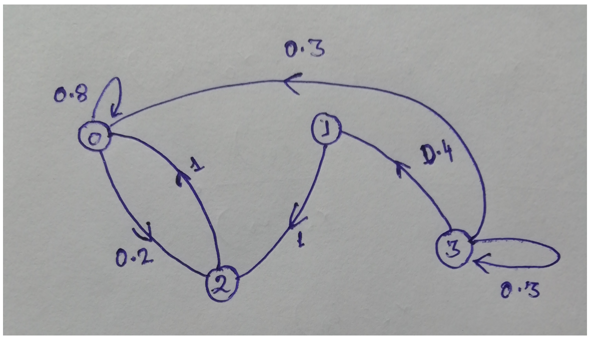


FIGURE 5. Transition graph

$$f_{00}^{(1)} = p_{00} = 0.8\,.$$
$$f_{00}^{(2)} = (0.2)(1.0)\,.$$
$$f_{00}^{(n)} = 0, \quad n \geq 3\,.$$

Thus, $f_{00} = f_{00}^{(1)} + f_{00}^{(2)} = 1$. Since $f_{11}^{(n)} = 0$, we have $f_{11} = 0$. Let us consider f_{22}. Notice that

$$\begin{aligned} f_{22} &= f_{22}^{(1)} + f_{11}^{(2)} + f_{11}^{(3)} + f_{11}^{(4)} + \cdots \\ &= 0 + (0.2)(0.1) + (1.0)(0.8)(0.2) + (1.0)(0.8)^2(0.2) + \cdots = 1 \end{aligned}$$

Again, since $f_{33}^{(1)} = 0.3$ and $f_{33}^{(n)} = 0$ for all $n \geq 2$, we see that $f_{33} = 0.3$. Hence the states 0 and 2 are recurrent and the states 1 and 3 are transient.

Example 7.9. *Consider the Markov chain with state space $S = \{0, 1, 2, 3\}$ and transition matrix*

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

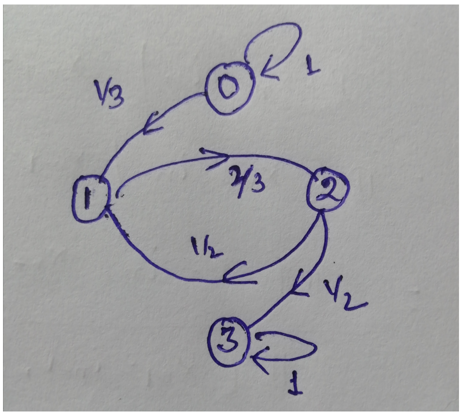


FIGURE 6. Transition graph

Note that $p_{00} = 1 = p_{33}$. *Thus the states* 0 *and* 3 *are absorbing states. Note from the transition graph (see; Figure 6) that* $f_{00}^{(1)} = 1$, *and* $f_{00}^{(n)} = 0$ *for all* $n \geq 2$. *Thus,* $f_{00} = 1$. *Hence the state* 0 *is recurrent. Similarly state* 3 *is recurrent. Consider the state* 1. *Observe that*

$$f_{11}^{(1)} = 0, \; f_{11}^{(2)} = 2/3 \cdot 1/2 = 1/3, \; f_{11}^{(n)} = 0, \forall\, n \geq 3\,.$$

$$\implies f_{11} = \sum_{n=1}^{\infty} f_{11}^{(n)} = 1/3 \neq 1\,.$$

Thus the state 1 *is transient. Similarly, we can show that the state* 2 *is also transient.*

Example 7.10. *Consider the Markov chain as stated in Example 7.9. We have seen that the state* 0 *is recurrent. Since* $f_{00}^{(1)} = 1$, *and* $f_{00}^{(n)} = 0$ *for all* $n \geq 2$, *it is easy to see that* $\mu_0 = \sum_{n=1}^{\infty} f_{00}^{(n)} = 1 < +\infty$. *Hence the state* 0 *is positive recurrent.*

Example 7.11. *Consider a Markov chain with state space* $\mathbb{N}$ *and transition probabilities*

$$p_{01} = 1, p_{i,0} = \frac{1}{i+1}, p_{i,i+1} = \frac{i}{i+1}, \quad i \geq 1\,.$$

Consider the state 0. *It is easily verified that*

$$f_{00}^{(1)} = 0, \; f_{00}^{(n)} = \frac{1}{n(n-1)}, \quad n \geq 2\,,$$

$$\implies f_{00} = \sum_{n=1}^{\infty} f_{00}^{(n)} = 1\,, \quad \text{and} \quad \mu_0 = \sum_{n=1}^{\infty} n f_{00}^{(n)} = \sum_{n=1}^{\infty} \frac{1}{n} = \infty\,.$$

Hence, 0 *is a null recurrent state.*

We now discuss the necessary and sufficient conditions for a state i to be recurrent.

Theorem 7.7. *A state* $i \in S$ *is recurrent if and only if*

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

As an immediate consequence of the above theorem, we arrive at the following corollary.

Corollary 7.8. *Let* $i \leftrightarrow j$. *Then* i *is recurrent if and only if* j *is recurrent.*

Proof. Since $i \leftrightarrow j$, there exist $m, n \geq 0$ such that $p_{ij}^{(n)} > 0$ and $p_{ji}^{(m)} > 0$. Then, we have

$$\sum_{k=0}^{\infty} p_{jj}^{(n+m+k)} \geq \sum_{k=0}^{\infty} p_{ji}^{(m)} p_{ii}^{(k)} p_{ij}^{(n)} = p_{ji}^{(m)} p_{ij}^{(n)} \sum_{k=0}^{\infty} p_{ii}^{(k)}\,.$$

Hence if $\sum_{k=0}^{\infty} p_{ii}^{(k)}$ diverges, then $\sum_{k=0}^{\infty} p_{jj}^{(k)}$ also diverges. This completes the proof. □

Remark 7.5. We have the followings:

a) One can easily see that

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} = \sum_{n=0}^{\infty} \mathbb{P}\big(X_n = j \big| X_0 = i\big) = \sum_{n=0}^{\infty} \mathbb{E}\big[\mathbf{1}_{\{X_n=j\}} \big| X_0 = i\big] = \mathbb{E}\Big[\sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=j\}} \big| X_0 = i\Big]$$
$$= \text{expected number of visits that the chain makes to state } j \text{ starting at } i\,.$$

Therefore, previous theorem states that a state j is recurrent if and only if the expected number of returns is infinite.

b) By using First Entrance Theorem, one can show that **for any transient state** j**,**

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty \quad \forall i \in S \implies \lim_{n\to\infty} p_{ij}^{(n)} = 0 \quad \forall i \in S.$$

The following question arises naturally: what happens to $\lim_{n\to\infty} p_{ij}^{(n)}$ when the state j is recurrent? We now state an important theorem (without proof) which gives limiting properties of n-th step transition probability:

Theorem 7.9 (Limit Theorem)**.** *The following holds:*

- *If state j is positive recurrent, then as $n \to \infty$*
  - *a)* $p_{jj}^{(n\lambda(j))} \to \dfrac{\lambda(j)}{\mu_j}$*, when j is periodic with period $\lambda(j)$.*
  - *b)* $p_{jj}^{(n)} \to \dfrac{1}{\mu_j}$*, when j is aperiodic. Moreover,* $\dfrac{1}{n}\displaystyle\sum_{r=1}^{n} p_{jj}^{(r)} \to \dfrac{1}{\mu_j}$.
  - *c) If j is aperiodic, then for every l,*

$$\lim_{n\to\infty} p_{lj}^{(n)} = \frac{f_{lj}}{\mu_j}, \quad \textit{where} \quad f_{lj} = \sum_{n=1}^{\infty} f_{lj}^{(n)}\,.$$

- *If the state j is null recurrent, then $p_{jj}^{(n)} \to 0$. Moreover, for every $l \in S$,* $\lim_{n\to\infty} p_{lj}^{(n)} = 0$.

Remark 7.6. If the state j is recurrent and $i \leftrightarrow j$, then $f_{ij} = 1$.

Remark 7.7. We have seen that a state i is recurrent if and only if $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$. This leads to the conclusion that **in a finite state Markov chain, not all states are transient**. To see this, let $S = \{0, 1, 2, \ldots, m\}$ and all states are transient. Then after a finite amount of time, say t_0, state 0 will never be visited, and after a time, say t_1, the state 1 will never be visited, and so on. Thus, after a finite time $t_{\max} = \max\{t_0, t_1, \ldots, t_m\}$ no states will be visited. But as the process must be in some state at time $t_{\max}$, we arrive at a contradiction–which then implies that at least one of the states must be recurrent.

Theorem 7.10. *In an irreducible chain, all states are of the same type. They are either all transient, all null recurrent, or all positive recurrent.*

Proof. Since the chain is irreducible, every state can be reached from every other state. Thus, if i and j be two state,

$$p_{ij}^{(n_1)} := a > 0, \quad p_{ji}^{(n_2)} := b > 0, \text{ for some } n_1, n_2 \geq 1\,.$$

Thus, from Chapman-Kolmogorov equation, we have

$$\begin{aligned} p_{ii}^{(n+n_1+n_2)} &\geq p_{ij}^{(n_1)} p_{jj}^{(n)} p_{ji}^{(n_2)} = ab p_{jj}^{(n)}\,, \\ p_{jj}^{(n+n_1+n_2)} &\geq p_{ji}^{(n_2)} p_{ii}^{(n)} p_{ij}^{(n_2)} = ab p_{ii}^{(n)}\,. \end{aligned}$$

Thus, the two series $\sum_n p_{ii}^{(n)}$ and $\sum_n p_{jj}^{(n)}$ converge or diverge together. Thus, the two states i and j are either both transient or both recurrent.

Now suppose i is null recurrent. Then $\lim_{n\to\infty} p_{ii}^{(n)} = 0$. Hence from the above inequality, $\lim_{n\to\infty} p_{jj}^{(n)} = 0$, so j is null recurrent.

□

Corollary 7.11. *In a finite irreducible Markov chain all the states are positive recurrent.*

Proof. In view of the above remark, any finite state Markov chain must have a recurrent state. Again, since in an irreducible chain, all the states are of same type, we conclude that all the states are recurrent. Suppose if that the state i is null recurrent. Then all other states are null-recurrent. This implies that $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all $j \in S$. Now, $\sum_{j\in S} p_{ij}^{(n)} = 1$ for all n. Thus, since S is finite, we lead to a contradiction. Hence all states must be positive recurrent. □

Definition 7.8. A state $j \in S$ of a Markov chain is said to be **ergodic** if j is positive recurrent and $\lambda(j) = 1$. Markov chain is said to be **ergodic Markov chain** if all the states are ergodic.

Example 7.12. *Consider the Markov chain as stated in Example 7.9. We have seen that the state* 0 *is positive recurrent. Moreover, from the transition graph (see Figure 6), it is easy to see that period of the state* 0 *is* 1*. Thus, the state* 0 *is ergodic.*

7.3. **Finite Markov chain and absorption probabilities:** The transition matrix P of a finite Markov chain can always be written in the following *canonical form*:

$$P = \begin{pmatrix} R & \mathbf{O} \\ A & B \end{pmatrix}$$

where

i) R is a submatrix of probability of transition between recurrent states,
ii) A is a submatrix of probability of transition from transient states to recurrent states,
iii) B is a submatrix of probability of transition between transient states,
iv) $\mathbf{O}$ is the zero matrix.

Example 7.13. *Find the canoniacl form the of the given transition matrix*

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

of a Markov chain with state space $S = \{0, 1, 2, 3, 4\}$.
Solution: *Check that* $1, 2$ *and* 3 *are transient states and* 0 *and* 4 *are recurrent states. So,*

$$S_R = \{0, 4\}, \quad S_T := \{1, 2, 3\}.$$

Therefore, we have

$$R = \begin{pmatrix} p_{00} & p_{04} \\ p_{40} & p_{44} \end{pmatrix}, \quad B = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}, \quad A = \begin{pmatrix} p_{10} & p_{14} \\ p_{20} & p_{24} \\ p_{30} & p_{34} \end{pmatrix}.$$

Thus, the required canonical form of P is given by

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

In the study of finite Markov chains, we are interested in answering questions:

a) What is the probability that starting from a transient state i, the chain reaches the recurrent state j at least once?
b) Given that the chain is in a transient state, what is the mean number of visits to another transient state j before reaching a recurrent state?

Definition 7.9 (Fundamental Matrix)**.** Let $\{X_n : n \geq 0\}$ be a finite Markov chain with transition matrix P. The matrix

$$M = (\mathbf{I} - B)^{-1}$$

is called **fundamental matrix** for P.

The entry m_{ij} of M gives the expected number of times that the process is in transient state j if it is started in the transient state i.

Example 7.14. *Consider the Example 7.13. The fundamental matrix M of P is given by*

$$M = (\mathbf{I} - B)^{-1} = \begin{pmatrix} 1 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix}.$$

Therefore, if the process start at state 2, then the expected number of times in state $1, 2$ and 3 before being absorbed are $1, 2$ and 1 respectively.

Absorption probabilities

Theorem 7.12. *Let g_{ij} be the probability that a finite Markov chain reaches the recurrent state j at least once if the chain starts from the transient state i. Let G be the matrix with entries g_{ij}. Then G is given by*

$$G = MA$$

where M is the fundamental matrix and A is in the canonical form.

Example 7.15. *Consider the finite Markov chain as in Example 7.13. Find the probability of absorption in state 0 if the chain starts from state 1.*
Solution: *In this case*

$$M = \begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix}, \quad A = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad S_R = \{0, 4\}, \quad S_T = \{1, 2, 3\}.$$

Hence the matrix G is given by

$$G = \begin{pmatrix} 3/2 & 1 & 1/2 \\ 1 & 2 & 1 \\ 1/2 & 1 & 3/2 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}.$$

Thus, the probability of absorption in state 0 if the chain starts from state 1 is $3/4$.

7.4. **Invariant measure and Limiting distribution of DTMC.**

Definition 7.10. Let $\{X_n : n \ge 0\}$ be a Markov chain with state space S. A probability measure $\Pi = (\Pi_i)_{i\in S}$ over S is called **invariant or stationary** if

$$\Pi_i \ge 0, \ \sum_{i\in S} \Pi_i = 1, \ \text{and } \Pi_j = \sum_{i\in S} \Pi_i p_{ij} \quad \forall\, j \in S\,. \tag{7.3}$$

In matrix notation, Π is invariant if $\Pi = \Pi P$.

Let $\{X_n : n \ge 0\}$ be a Markov chain such that the initial distribution $(\pi_i)_{i\in S}$ is stationary i.e.,

$$\pi_j = \sum_{i\in S} \pi_i p_{ij} \quad \text{where} \quad \pi_i = \mathbb{P}(X_0 = i)\,.$$

Now, the law of total probability gives

$$\mathbb{P}(X_1 = j) = \sum_{i\in S} \mathbb{P}\big(X_1 = j \big| X_0 = i\big)\mathbb{P}(X_0 = i) = \sum_{i\in S} \pi_i p_{ij} = \pi_j$$

and by induction, we get

$$\mathbb{P}(X_n = j) = \sum_{i\in S} \mathbb{P}\big(X_n = j \big| X_{n-1} = i\big)\mathbb{P}(X_{n-1} = i) = \sum_{i\in S} \pi_i p_{ij} = \pi_j\,.$$

Thus, $\{X_n\}$ has the same distribution for all n. In fact since $\{X_n : n \ge 0\}$ is a Markov chain, it follows from this fact that for each $m \ge 0$, $X_n, X_{n+1}, \ldots, X_{n+m}$ will have same joint distribution for each n. In other words, $\{X_n : n \ge 0\}$ will be a stationary process.

Definition 7.11 (Limiting distribution). Let $\{X_n : n \ge 0\}$ be a Markov chain with state space S. We say that chain has a limiting distribution, if there exists an invariant measure $\Pi = (\Pi_i)_{i\in S}$ over S such that $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \Pi_j$ for all $j \in S$.

Theorem 7.13. *Let $\{X_n : n \ge 0\}$ be an irreducible, aperiodic Markov chain with state space S. Then there exists an invariant probability measure over S if and only if the chain is positive recurrent. The determined probability measure is unique and satisfies the condition (7.3).*

Proof. With out loss of generality, we assume that $S = \mathbb{N}$. let us assume that there exists an invariant measure Π on S i.e., $\Pi_j = \sum\limits_{i\in S} \Pi_i p_{ij}$. Since $P^{(n)} = P \cdot P \ldots P$ (n times) and $\Pi = \Pi P$, we observe that $\Pi = \Pi P^{(n)}$ and hence $\Pi_j = \sum\limits_{i\in S} \Pi_i p_{ij}^{(n)}$. If the states of an irreducible Markov chain are transient or null recurrent, then $\lim\limits_{n\to\infty} p_{ij}^{(n)} = 0$, and hence we get that $\Pi_j = 0$ for all j–a contradiction as $\sum\limits_{j=1}^{\infty} \Pi_j = 1$. Thus, all states are positive recurrent.

Conversely assume that the chain is positive recurrent. Hence $\mu_j = \sum\limits_{n=1}^{\infty} n f_{jj}^{(n)} < +\infty$. Define

$$\Pi_j := \frac{1}{\mu_j} \quad j \in S\,.$$

Since $P^{(n)}$ is stochastic matrix, we have

$$1 = \sum_{j=1}^{\infty} p_{ij}^{(n)} \ge \sum_{j=1}^{m} p_{ij}^{(n)}\,.$$

Since the chain is irreducible $i \leftrightarrow j$, and as j is positive recurrent and aperiodic, we get

$$\begin{aligned}
&\lim_{n\to\infty} p_{ij}^{(n)} = \frac{f_{ij}}{\mu_j} = \frac{1}{\mu_j} = \Pi_j \\
&\implies \sum_{j=1}^{m} p_{ij}^{(n)} \to \sum_{j=1}^{m} \Pi_j \\
&\implies \sum_{j=1}^{\infty} \Pi_j \le 1\,.
\end{aligned}$$

On the other hand, using the Chapman-Kolmogorov equation, $p_{ij}^{(n+1)} \ge \sum_{k=1}^{m} p_{ik}^{(n)} p_{kj}$. Taking limit as $n \to \infty$, we have

$$\Pi_j \ge \sum_{k=1}^{m} \Pi_k p_{kj} \implies \sum_{k=1}^{\infty} \Pi_k p_{kj} \le \Pi_j \quad \forall\, j \in S\,.$$

We now show that

$$\sum_{k=1}^{\infty} \Pi_k p_{kj} = \Pi_j\,.$$

If not, there exists j_0 such that $\sum_{k=1}^{\infty} \Pi_k p_{kj_0} < \Pi_{j_0}$. Then

$$\begin{aligned}
&\sum_{j=1}^{\infty} \Pi_j > \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \Pi_k p_{kj} = \sum_{k=1}^{\infty} \Pi_k \sum_{j=1}^{\infty} p_{kj} = \sum_{k=1}^{\infty} \Pi_k \text{—a contradiction} \\
&\implies \sum_{k=1}^{\infty} \Pi_k p_{kj} = \Pi_j \quad \text{i.e.,} \quad \Pi = \Pi P \implies \Pi_j = \sum_{k=1}^{\infty} \Pi_k p_{kj}^{(n)} \quad \forall j \in S\,, \forall\, n \ge 1\,.
\end{aligned}$$

Thus, for given $\varepsilon > 0$, there exists $n_0 \ge 1$ such that $\displaystyle\sum_{k=n_0+1}^{\infty} \Pi_k p_{kj}^{(n)} < \varepsilon$. Hence

$$\Pi_j \le \sum_{k=1}^{n_0} \Pi_k p_{kj}^{(n)} + \varepsilon \to \sum_{k=1}^{n_0} \Pi_k \Pi_j + \varepsilon$$

Since ε is arbitrary, from the above inequality, we conclude that

$$\Pi_j = \Pi_j \sum_{k=1}^{\infty} \Pi_k \implies \sum_{k=1}^{\infty} \Pi_k = 1\,.$$

Uniqueness: Suppose there exists another stationary distribution $(r_k)_{k\in S}$. Then the similar argumentation gives that $r_j = \sum_{k=1}^{\infty} r_k p_{kj}^{(n)}$. Now taking limit as $n \to \infty$, we obtain

$$r_j = \sum_{k=1}^{\infty} r_k \Pi_j = \Pi_j \sum_{k=1}^{\infty} r_k = \Pi_j \quad \forall\, j \in S\,.$$

This gives the uniqueness. $\square$

Example 7.16. *Let $\{X_n : n \ge 0\}$ be a Markov chain with state space $S = \{1, 2, 3\}$ and transition matrix* $P = \begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix}$. *Explain whether there exists a invariant measure over S or not. If so, then find it.*

Solution: *The transition diagram of the given Markov chain is given in Figure 7.*

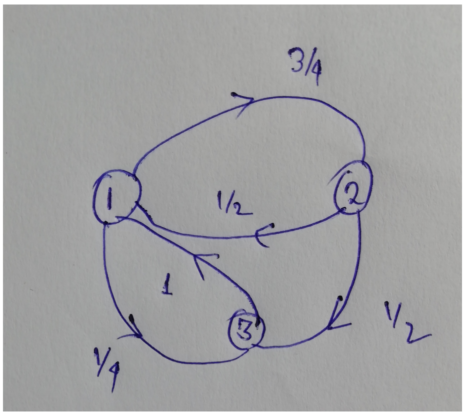


FIGURE 7. Transition graph

From the diagram, it is easy to see that the chain is irreducible and aperiodic. Since S is finite , the chain is a positive recurrent. Thus, there exists an invariant measure $\Pi = (\Pi_i)_{i\in S}$ over S with

$$\Pi_i \geq 0, \quad \sum_{i\in S} \Pi_i = 1, \quad \Pi_j = \sum_{i\in S} \Pi_i p_{ij}\,.$$

such that

$$\lim_{n\to\infty} p_{ij}^{(n)} = \Pi_j\,.$$

Hence, we have

$$\Pi_1 + \Pi_2 + \Pi_3 = 1, \;\; \Pi_1 = \frac{1}{2}\Pi_2 + \Pi_3, \;\; \Pi_2 = \frac{3}{4}\Pi_1, \;\; \Pi_3 = \frac{1}{4}\Pi_1 + \frac{1}{2}\Pi_2\,.$$

Solving these, we have

$$\Pi_1 = \frac{8}{19}, \quad \Pi_2 = \frac{6}{19}, \quad \Pi_3 = \frac{5}{19}\,.$$

Since $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \Pi_j$, we infer that for some sufficiently large n, the probability that the chain is in state 1 given that it started from state i is equal to $\frac{8}{19}$, independent of the initial state i.

Example 7.17. *Consider the simple Queueing model with countable state space $S = \{0, 1, 2, 3, \ldots\}$ and transition probability matrix $P = \begin{pmatrix} p_0 & p_1 & p_2 & p_3 & \cdots \\ p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & p_0 & p_1 & p_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$ where $p_j := \mathbb{P}(Y_n = j)$ and $\sum p_k = 1$with $p_i > 0$. Under what condition, chain has invariant measure over S.*

Solution: *Since $p_i > 0$ for all i, the chain is irreducible and aperiodic. To have existence of invariant measure we need to show that the chain is positive recurrent. Let us solve the system of equations*

$$\Pi_j = \sum_{i\in S} \Pi_i p_{ij} \quad (j \geq 0)\,.$$

To solve this, we introduce the generating functions

$$\Pi(s) = \sum_{j=0}^{\infty} \Pi_j s^j, \quad P(s) = \sum_k p_k s^k\,.$$

Note that

$$\Pi_k = p_k \Pi_0 + p_k \Pi_1 + p_{k-1}\Pi_2 + \ldots + p_0 \Pi_{k+1}$$

$$\implies \Pi_k s^k = s^k\big(p_k\Pi_0 + p_k\Pi_1 + p_{k-1}\Pi_2 + \ldots + p_0\Pi_{k+1}\big)$$
$$\implies \Pi(s) = \sum_{k=0}^{\infty}\Pi_k s^k = \sum_{k=0}^{\infty} s^k\big(p_k\Pi_0 + p_k\Pi_1 + p_{k-1}\Pi_2 + \ldots + p_0\Pi_{k+1}\big)$$
$$= \Pi_0 P(s) + \Pi_1 P(s) + s\Pi_2 P(s) + s^2\Pi_3 P(s) + \ldots$$
$$= P(s)\big\{\Pi_0 + \frac{\Pi(s)-\Pi_0}{s}\big\}$$
$$\implies \Pi(s) = \frac{\Pi_0(1-s)P(s)}{P(s)-s}\,.$$

To compute Π_0*, we let* $s \to 1$*. Define* $\rho := \sum_{k=0}^{\infty} kp_k < 1$*. By Abel's lemma, we have*

$$\lim_{s\to 1-} P(s) = \sum_{i=0}^{\infty} p_i = 1, \qquad \lim_{s\to 1-} P'(s) = \rho < 1\,.$$

Hence by L'hospital's rule, we have

$$\lim_{s\to 1-}\Pi(s) = \Pi_0 \lim_{s\to 1-}\frac{P(s)}{1-P'(s)} = \Pi_0\big(1-\rho\big)^{-1}\,.$$

Since $\lim_{s\to 1-}\Pi(s) = \sum_{i=0}^{\infty}\Pi_i$*, we have*

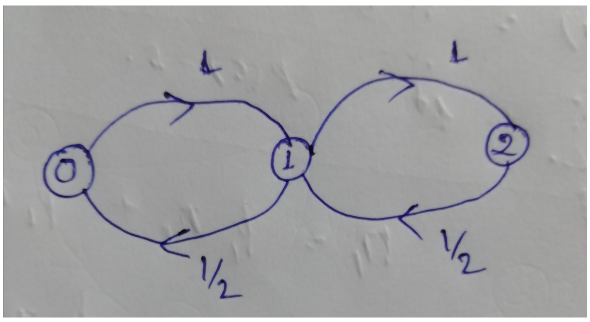
$$\sum_{i=0}^{\infty}\Pi_i = \Pi_0(1-\rho)^{-1}\,.$$

Hence the stationary probabilities exist if and only if $\rho < 1$*. In this case* $\Pi_0 = 1-\rho$*, and*

$$\Pi(s) = \frac{(1-\rho)(1-s)P(s)}{P(s)-s}\,.$$

Example 7.18. *Let* $\{X_n : n \geq 0\}$ *be a Markov chain with state space* $S = \{0,1,2\}$*, the transition matrix* $P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$*. Examine whether there exists a limiting distribution for this Markov chain.*

Solution: *From the transition diagram it is clear that the chain is irreducible. Hence there exists*



only one class. Since the chain is finite, it is positive recurrent. Therefore, limiting distribution for this chain will exist if the chain is aperiodic. Note here that the periodicity of the chain is 2*. Hence there does not exist any limiting distribution.*