Discrete Time Markov Chain

Consider three events A, B and C in a probability space $(S, \mathbb{F}, P)$. Recall that the joint probability of A, B and C is given by

$$P\left(A\cap B\cap C\right)=P(A)P\left(B\cap C\mid A\right)$$

We can also apply the chain rule for computing this probability:

$$P\left(A\cap B\cap C\right)=P(A)P\left(B\mid A\right)P\left(C\mid A\cap B\right) \qquad (1)$$

If A, B and C are independent, then

$$P\left(A\cap B\cap C\right)=P(A)P\left(B\right)P\left(C\right)$$

Just like independence, the notion of *conditional independence* simplifies the computation of the joint probability and is discussed below.

Definition: Given A, the events B and C are called *conditionally independent* if

$$P\left(B\cap C / A\right)=P\left(B\mid A\right)P\left(C\mid A\right)$$

Or equivalently

$$P\left(C / A\cap B\right)=P\left(C\mid A\right)$$

This is because,

$$P\left(C / A\cap B\right)=\frac{P\left(A\cap B\cap C\right)}{P\left(A\cap B\right)}$$

$$=\frac{P(A)P\left(B\cap C\mid A\right)}{P\left(A\cap B\right)}$$

$$=\frac{P(A)P\left(B\mid A\right)P(C\mid A)}{P\left(A\right)P(B\mid A)} \quad \text{(using the first definition of conditional independence)}$$

$$=P(C\mid A)$$

Thus,if B and C are conditionally independent given A, the joint probability is given by

$$P\left(A\cap B\cap C\right)=P(A)P\left(B\mid A\right)P\left(C\mid A\right)$$

Note that conditional independence does not imply unconditional independence and vice versa.

Example 1

Consider the experiment of tossing two dice. The possible outcomes are tabulated as shown.

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Let A be the event that the first throw is greater than 1, B be the event that the first throw is 3 and C be the event that the first throw is 1 or the second throw is 6.

Clearly, B and C are not independent.

$$P\left(B\cap C/A\right)=\frac{1}{30},\ P\left(B/A\right)=\frac{6}{30} \text{ and } P\left(C/A\right)=\frac{5}{30}$$

Therefore, B and C are conditionally independent given A.

The conditional independence of events is used to define an important class of random process called the *Markov Process.*

Definition: Markov process

A random process $\{X(t), t\in\Gamma\}$ defined on $(S,\mathbb{F},P)$ is called a Markov process if for any sequence of time $t_1<t_2<.....<t_n<t_{n+1}\in\Gamma$,

$$P(\{X(t_{n+1})\le x \mid X(t_1)=x_1, X(t_2)=x_2,..., X(t_n)=x_n\})=P(\{X(t_{n+1})\le x \mid X(t_n)=x_n\}) \qquad (2)$$

For a Markov process $\{X(t)\}$, the conditional CDF of $X(t_{n+1})$ given the values of $X(t_1), X(t_2),..., X(t_n)$ is same as the conditional CDF of $X(t_{n+1})$ given the value of $X(t_n)$ alone. In other words, given $X(t_n)$, $X(t_{n+1})$ is conditionally independent of $X(t_1), X(t_2),..., X(t_{n-1})$. This property is known as the *Markovian property.*

Definition: Markov Chain

Suppose $X(t_n)$ takes values from a *countable* set V called the *state space*. The elements of V are called the states of the process $\{X(t_n)\}$. Since V has one-to-one correspondence with some subset of $\mathbb{Z}$, we can assume V as a set consisting of integers. Thus $\{X(t_n)=i\}$ means the event that $X(t_n)$ takes the ith state. For such a process, the Markovian property in (2) can be expressed in terms of the probability mass function and the process is called a *Markov chain* (MC). A Markov chain may be a *continuous-time Markov chain (CTMC*) or a *discrete-time Markov chain (DTMC)*.

Suppose $\Gamma=[0,\infty)$ and $\{X(t)\}$ takes values from a discrete state space $V=\{0,1,2,\cdots\}$. Then $\{X(t), t\in[0,\infty)\}$ is called a CTMC if for any $n\ge 1$ and $t_1<t_2\cdots<t_{n+1}$,

$$P\left(X\left(t_{n+1}\right)=j \mid X\left(t_0\right)=i_0, X\left(t_1\right)=i_1,\cdots, X\left(t_n\right)=i\right)=P\left(X\left(t_{n+1}\right)=j \mid X\left(t_n\right)=i\right) \qquad (3)$$

<u>Example 2:</u> Independent increment process: For any $n > 1$ and $t_0 < t_1 < ... < t_n \in \Gamma,$ we have

$$\begin{aligned} &P(X(t_n) = i_n / X(t_{n-1}) = i_{n-1}, X(t_{n-2}) = i_{n-2}, ... ,X(t_0) = i_0) \\ &= P(X(t_n) - X(t_{n-1}) = i_n - i_{n-1} / X(t_{n-1}) = i_{n-1}, X(t_{n-2}) = i_{n-2}, ... ,X(t_0) = i_0) \\ &= P(X(t_n) - X(t_{n-1}) = i_n - i_{n-1}) \quad \text{(Using the independent increment property)} \end{aligned}$$

Similarly,

$$\begin{aligned} &P(X(t_n) = i_n / X(t_{n-1}) = i_{n-1}) \\ &= P(X(t_n) - X(t_{n-1}) = i_n - i_{n-1} / X(t_{n-1}) = i_{n-1}) \\ &= P(X(t_n) - X(t_{n-1}) = i_n - i_{n-1}) \\ &\therefore P(X(t_n) = i_n / X(t_{n-1}) = i_{n-1}, X(t_{n-2}) = i_{n-2}, ... ,X(t_0) = i_0) = P(X(t_n) = i_n / X(t_{n-1}) = i_{n-1}) \end{aligned}$$

Thus $\{X(t), t \in [0, \infty)\}$ is a CTMC.

Next consider a discrete time random process $\{X_n, n \geq 0\}$ defined on $(S, \mathbb{F}, P)$ such that the index n is a non-negative integer starting at 0. *A discrete-time random process* $\{X_n, n \geq 0\}$ taking values from a countable set V is said to be a DTMC if

$$P\left(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \cdots, X_n = i\right) = P\left(X_n = j \mid X_n = i\right). \quad (4)$$

We shall discuss the DTMC first.

<u>Example 3</u>

Suppose $\{X_n, n \geq 0\}$ is a sequence of iid and integer-valued random variables. Then

$$\begin{aligned} \therefore P\left(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, ..., X_n = i\right) &= P\left(X_{n+1} = j\right) \\ &= P\left(X_{n+1} = j / X_n = i\right) \end{aligned}$$

Therefore, a sequence of integer –valued iid random variables is trivially an MC.

<u>Example 4</u>

Suppose $\{Z_n, n \geq 0\}$ is a sequence of iid and integer-valued random variables and $X_n = \sum_{i=0}^{n} Z_i$.

Then $\{X_n, n \geq 0\}$ is an MC.

Solution:

We have $X_{n+1} = \sum_{i=0}^{n+1} Z_i = \sum_{i=0}^{n} Z_i + Z_{n+1} = X_n + Z_{n+1}$

$$\therefore P\left(X_{n+1}=j \mid X_n=i\right)=\frac{P\left(X_{n+1}=j, X_n=i\right)}{P\left(X_{n-1}=i\right)}$$
$$=\frac{P\left(Z_{n+1}=j-i, X_n=i\right)}{P\left(X_n=i\right)}$$
$$=\frac{P\left(Z_{n+1}=j-i\right)P\left(X_n=i\right)}{P\left(X_n=i\right)}$$
$$=P\left(Z_{n+1}=j-i\right)$$

The third step results from the fact that Z_{n+1} is independent of $Z_0, Z_1, \cdots, Z_n$ implying that Z_{n+1} is independent of X_n which is a function of independent random variables.

Again

$$P\left(X_{n+1}=j \mid X_n=i, X_{n-1}=k\right)$$
$$=\frac{P\left(X_{n+1}=j, X_n=i, X_{n-1}=k\right)}{P\left(X_n=i, X_{n-1}=k\right)}$$
$$=\frac{P\left(Z_{n+1}=j-i, Z_n=i-k, X_{n-1}=k\right)}{P\left(Z_n=i-k, X_{n-1}=k\right)}$$
$$=\frac{P\left(Z_n=j-i\right)P\left(Z_{n-1}=i-k\right)P\left(X_{n-1}=k\right)}{P\left(Z_n=i-k\right)P(X_{n-1}=k)}$$
$$=P\left(Z_n=j-i\right)$$

Arguing in the similar manner, we can show that $P\left(X_{n+1}=j \mid X_0=i_0, X_1=i_1, \cdots, X_n=i\right)=P\left(Z_n=j-i\right)$

$$\therefore P\left(X_{n+1}=j \mid X_0=i_0, X_1=i_1, \cdots, X_n=i\right)=P\left(X_{n+1}=j \mid X_n=i\right)$$

$\therefore \left\{X_n, n \geq 0\right\}$ is an MC.

When Z_n takes values from $\{-1,1\}$, then $\left\{X_n, n \geq 0\right\}$ is the simple random walk process. Thus the simple random walk process is an MC.

Theorem 1

For a discrete-time MC, the following important relation equivalently holds:

$$P\left(X_{n+m}=j \mid X_0=i_0, X_1=i_1, \cdots, X_n=i\right)=P\left(X_{n+m}=j \mid X_n=i\right), m=2,3,..$$

Proof:

We use the total probability theorem to prove the result.

Using the Markov property, we can show that

$$P\left(X_0=i_0, X_1=i_1, \cdots, X_n=i, X_{n+1}=j_1, X_{n+2}=j_2, ..., X_{n+m}=j\right)$$
$$=P(X_0=i_0)P(X_1=i_1 / X_0=i_0)P(X_2=i_2 / X_0=i_0, X_1=i_1)\cdots P(X_n=i / X_0=i_0, X_1=i_1, \cdots, X_{n-1}=i_{n-1})...P($$
$$=P(X_0=i_0)P(X_1=i_1 / X_0=i_0)P(X_2=i_2 / X_1=i_1)\cdots P(X_n=i / X_{n-1}=i_{n-1})...P(X_{n+m}=j / X_{n+m-1}=j_{n+m-1})$$

We find the marginal probability $P\left(X_0=i_0, X_1=i_1, \cdots, X_n=i, X_{n+m}=j\right)$ as

$$P\left(X_0 = i_0, X_1 = i_1, \cdots, X_n = i, X_{n+m} = j\right)$$
$$= \sum_{j_1}\sum_{j_2}\cdots\sum_{j_{n+m-1}} P(X_0 = i_0)P(X_1 = i_1 / X_0 = i_0)P(X_2 = i_2 / X_1 = i_1)\cdots P(X_n = i / X_{n-1} = i_{n-1})P(X_{n+1} = j_1 / X$$
$$= P(X_0 = i_0)P(X_1 = i_1 / X_0 = i_0)P(X_2 = i_2 / X_1 = i_1)\cdots P(X_n = i / X_{n-1} = i_{n-1})\sum_{j_1}\sum_{j_2}\cdots\sum_{j_{n+m-1}} P(X_{n+1} = j_1 / X$$
$$\therefore P\left(X_{n+m} = j / X_0 = i_0, X_1 = i_1, \cdots, X_n = i\right)$$
$$= \sum_{j_1}\sum_{j_2}\cdots\sum_{j_{n+m-1}} P(X_{n+1} = j_1 / X_n = i)\ldots P(X_{n+m} = j / X_{n+m-1} = j_{n+m-1})$$

Similarly,
$$P\left(X_{n+m} = j / X_n = i\right)$$
$$= \sum_{j_1}\sum_{j_2}\cdots\sum_{j_{n+m-1}} P(X_{n+1} = j_1 / X_n = i)\ldots P(X_{n+m} = j / X_{n+m-1} = j_{n+m-1})$$
$$P\left(X_{n+m} = j \mid X_0 = i_0, X_1 = i_1, \cdots, X_n = i\right) = P\left(X_{n+m} = j \mid X_n = i\right)$$

Evolution of a DTMC

Applying the chain rule as in (1), we get the joint probability of the states up to instant *n* as
$$P\left(X_0 = i_0, X_i = i_1, \cdots, X_n = i, X_{n+1} = j\right)$$
$$= P\left(X_0 = i_0\right)P\left(X_1 = i_1 \mid X_0 = i_0\right)P\left(X_2 = i_2 \mid X_0 = i_0, X_1 = i_1\right)\ldots$$
$$\ldots.P\left(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \cdots, X_n = i\right) \qquad (5)$$

Thus the complete description of the states of $\{X_n\}$ up to the instant n requires $P\left(X_0 = i_0\right)$ and the probabilities of the successive states conditioned to the respective past states. The computation of such probabilities is greatly simplified by the Markovian property and (5) reduces to

$$P\left(X_0 = i_0, X_i = i_1, \cdots, X_n = i, X_{n+1} = j\right)$$
$$= P\left(X_0 = i_0\right)P\left(X_1 = i_1 \mid X_0 = i_0\right)P\left(X_2 = i_2 \mid X_1 = i_1\right)\ldots.P\left(X_{n+1} = j \mid, X_n = i\right)$$

Definition: The conditional probability $p_{ij} = P\left(X_{n+1} = j \mid X_n = i\right)$ is called the *one-step transition probability* of the chain at the instant n.

Similarly, the m-step transition probability $p_{ij}^{(m)}$ is defined by

$$p_{ij}^{(m)} = P\left(X_{n+m} = j \mid X_n = i\right)$$

In general, p_{ij} may or may not depend on *n*. If p_{ij} does not depend upon n, then this transition probability is stationary and $\{X_n, n \geq 0\}$ is called *a homogeneous Markov chain*.

Thus for a homogeneous MC,

$$P\left(X_0=i_0, X_1=i_1,\cdots,X_n=i, X_{n+1}=j\right)$$
$$=P\left(X_0=i_0\right)P\left(X_1=i_1 \mid X_0=i_0\right)\cdots P\left(X_{n+1}=j \mid X_n=i\right)$$
$$=P\left(X_0=i_0\right)p_{i_0i_1}p_{i_1i_2}\cdots p_{ij}$$

In other words, the probabilistic evolution of a homogeneous Markov chain can be completely described by

(a) The initial probability $P\left(X_0=i_0\right)$

(b) The transition probability p_{ij}

The probability $p_{ij}^{(n)}=P\left(X_{n+m}=j \mid X_m=i\right)=P\left(X_n=j \mid X_0=i\right)$ is called the $n-$*step transition probability* of the homogenous MC $\left\{X_n, n\geq 0\right\}$.

The one step transition probabilities of a Markov chain can be represented compactly in terms of the *state transition matrix* $\mathbf{P}$ given by

$$\mathbf{P}=\begin{bmatrix} p_{00} & p_{01}\cdots p_{0k}\cdots \\ p_{10} & p_{11}\cdots p_{1k}\cdots \\ & \cdots \\ p_{k0} & p_{k1}\cdots p_{kk}\cdots \\ & \vdots \end{bmatrix}$$

$\mathbf{P}$ is a square matrix and its rows and columns are indexed by the elements of V . The matrix $\mathbf{P}$ is called a *stochastic matrix* in the sense that $\forall (i,j)\in V\times V$

(1) $p_{ij}\geq 0$

(2) $\sum_{j\in V}p_{ij}=1$

Thus the sum of elements in each row of $\mathbf{P}$ is one. If the sum of elements in each column of $\mathbf{P}$ is also one, then $\mathbf{P}$ is *a doubly stochastic matrix.* It is easy to verify that unity is an eigen value of $\mathbf{P}$.

<u>Example 5</u>

The identity matrix is a stochastic matrix where $p_{ij}=\begin{cases}1, & \text{for } i=j \\ 0 & \text{therwise}\end{cases}$.

<u>Example 6</u>

For the simple random walk process $\left\{X_n, n\geq 0\right\}$ the state transition matrix $\mathbf{P}$ is given by

$$\mathbf{P}=\begin{bmatrix} \cdots 1-p & 0 & p & 0\cdots \\ \vdots & 1-p & 0 & p\cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

which is doubly infinite matrix.

<u>Example 7</u> Random Walk with Barriers

Consider the random walk process $\{X_n, n \geq 0\}$ with the finite state space $V = \{0,1,...,N\}$. Suppose the process is at state i where $0 < i < N$. The process moves to state $i+1$ in the next instant with a probability *p* and to a state $i-1$with a probability *1-p*. Once the process reaches the sate 0, it stays in the same state in the next instant with a probability *a* and moves to the neighbouring state with a probability *1-a.* If *a=1*, the state 0 is the *absorbing barrier*, if *a=0,* it is the *reflecting barrier and* therwise it is an *elastic barrier*. Similar is in the case of sate N where the respective probabilities are *b* and 1-*b*. The state transition matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \begin{bmatrix} a & 1-a & 0 & ... & 0 & 0 & 0 \\ 1-p & p & 0 & ... & 0 & 0 & 0 \\ \cdots & \cdots & & \cdots & \cdots & & \\ 0 & 0 & 0 & ... & 1-q & 0 & 0 \\ 0 & 0 & 0 & ... & 0 & 1-b & b \end{bmatrix}$$

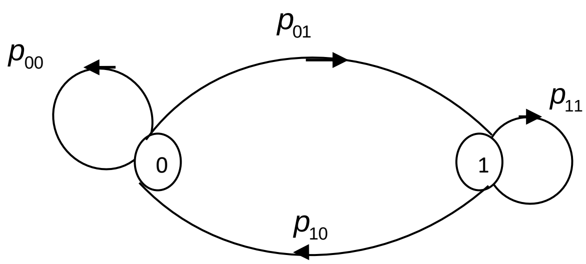
.

The n-step transition probabilities can be represented by the matrix $\mathbf{P}^{(n)}$ given by

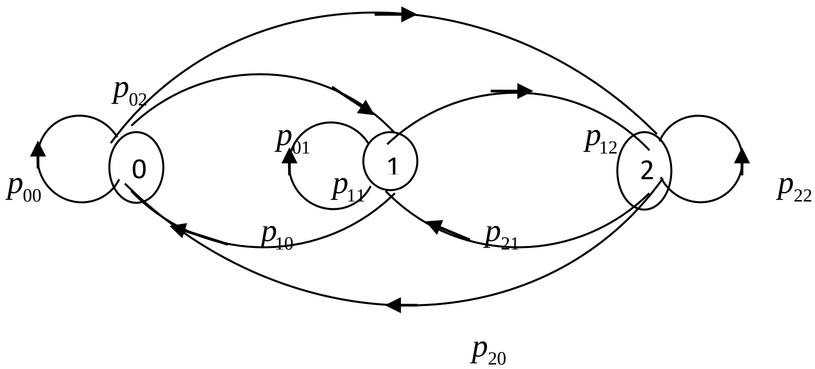
$$\mathbf{P}^{(n)} = \begin{bmatrix} p_{00}^{(n)} & p_{01}^{(n)} & \cdots & p_{0k}^{(n)} & \cdots \\ p_{10}^{(n)} & p_{11}^{(n)} & \cdots & p_{1k}^{(n)} & \cdots \\ \vdots & & & & \\ p_{k0}^{(n)} & p_{k1}^{(n)} & \cdots & p_{kk}^{(n)} & \cdots \\ \vdots & & & & \end{bmatrix}$$

State Transition Graph

The states and the state transition probabilities of an MC can be pictorially represented with the help of a *state transition graph* where each node denotes a state and each directed edge denotes the one-step transition probability. Figure 1 shows the state transition graph of a 2-state MC and a 3 -state MC.



(a)



(b)

Figure 1: Examples of state transition diagrams for (a) two-state and (b) three-state MCs

The n-step transition probabilities of a homogeneous MC can be computed from the one-step transition probabilities by using the following important theorem.

Theorem 2 Chapman-Kolmogorov Equation

For a homogenous MC $\{X_n, n \geq 0\}$,

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

Proof:

We have

$$P(X_0 = i, X_{m+n} = j)$$
$$= P\bigcup_k (X_0 = i, X_m = k, X_{m+n} = j)$$
$$= \sum_k P(X_0 = i, X_m = k, X_{m+n} = j)$$
$$= \sum_k P(X_0 = i) P(X_m = k \mid X_0 = i) P(X_{m+n} = j \mid X_0 = i, X_m = k) \quad \text{(using the Total Probability theorem).}$$
$$= \sum_k P(X_0 = i) P(X_m = k \mid X_0 = i) P(X_{m+n} = j \mid X_0 = i) \quad \text{(using the Markov property)}$$
$$= \sum_k P(X_0 = i)\, p_{ik}^{(m)} p_{kj}^{(n)}$$

Dividing by $P(X_0 = i)$, we get

$P\left(X_{n+m}=j \mid X_0=i\right)=\sum_k p_{ik}^{(m)} p_{kj}^{(n)}$

Particularly with $m=n-1$ and $n=1$, $p_{ij}^{(n)}=\sum_k p_{ik}^{(n-1)} p_{kj} \quad (i,j) \in V \times V$

The above expression can be compactly represented in the matrix notation

$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1)}\mathbf{P}$

Applying Chapman Kolmogorov equation for $p_{ij}^{(n-1)}$, we can write

$\mathbf{P}^{(n-1)} = \mathbf{P}^{(n-2)}\mathbf{P}$

Through mathematical induction we can show that $\mathbf{P}^{(n)} = \underbrace{\mathbf{PP}...\mathbf{P}}_{n-times} = \mathbf{P}^n$

Thus for every n,

$\mathbf{P}^{(n)} = \mathbf{P}^n$

which is the *Chapman-Kolmogorov equation in the matrix form*.

Marginal Distribution of the chain at instant n

Suppose $p_i^{(0)}, i \in V$ is the initial probability distribution. We can represent the probability of the states at $n=0$ by the row vector

$\mathbf{p}^{(0)} = \left[p_0^{(0)}\ p_1^{(0)} \ldots\ p_k^{(0)} \ldots . \right]$

We have

$$
\begin{aligned}
& P\left(X_0=i, X_n=j\right) \\
& = P\left(X_0=i\right) P\left(X_n=j \mid X_0=i\right) \\
& = p_i^{(0)} p_{ij}^{(n)}
\end{aligned}
$$

The marginal probability

$P\left(X_n=j\right)=\sum_i P\left(X_0=i, X_n=j\right)=\sum_i p_i^{(0)} p_{ij}^{(n)}$

In matrix notation,

$$
\begin{aligned}
\mathbf{p}^{(n)} &= \mathbf{p}^{(0)}\mathbf{P}^{(n)} \\
&= \mathbf{p}^{(0)}\mathbf{P}^{n}
\end{aligned}
$$

Example 5

Suppose $\{X_n, n \geq 0\}$ is a 3-state MC with $V=\{0,1,2\}$. Given

$\mathbf{P} = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.4 & 0 & 0.6 \\ 0 & 0.4 & 0.6 \end{bmatrix}$ and $\mathbf{p}^{(0)} = \left[0.2\ 0.6\ \ 0.2\right]$, find out the 2-step transition probability and the distribution of the state at (a) $n=2$ and (b) $n=20$.

Solution:

a) $n=2$

$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{bmatrix} 0.4 & 0.24 & 0.36 \\ 0.16 & 0.48 & 0.36 \\ 0.16 & 0.24 & 0.6 \end{bmatrix}$

$\mathbf{p}^{(2)} = \mathbf{p}^{(0)}\mathbf{P}^{(2)} =$ [0.256 0.384 0.36]

(c) $n=20$

(d) $\mathbf{P}^{(20)} = \mathbf{P}^{20} = \begin{bmatrix} 0.2105 & 0.3158 & 0.4737 \\ 0.2105 & 0.3158 & 0.4737 \\ 0.2105 & 0.3158 & 0.4737 \end{bmatrix}$

and

$\mathbf{p}^{(20)} = \mathbf{p}^{(0)}\mathbf{P}^{(20)} = [0.2105 \quad 0.3158 \quad 0.4737]$

Observe the elements of $\mathbf{P}^{(20)}$ and $\mathbf{p}^{(20)}$. These values illustrate the behavior of the chain as $n \to \infty$.

<u>Steady state probability distribution of states</u>

We have

$$\begin{aligned} \mathbf{p}^{(n)} &= \mathbf{p}^{(0)}\mathbf{P}^{n} \\ &= \mathbf{p}^{(0)}\mathbf{P}^{(n-1)}\mathbf{P} \\ \therefore \mathbf{p}^{(n)} &= \mathbf{p}^{(n-1)}\mathbf{P} \end{aligned}$$

If the steady state solution of the above system of linear difference equations exist and $\lim_{n\to\infty}\mathbf{p}^{(n)} = \lim_{n\to\infty}\mathbf{p}^{(n-1)} = \boldsymbol{\pi}$,then

$$\begin{aligned} &\lim_{n\to\infty}\mathbf{p}^{(n)} = \lim_{n\to\infty}\mathbf{p}^{(n-1)}\mathbf{P} \\ &\Rightarrow \boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} \\ &\Rightarrow \boldsymbol{\pi}(\mathbf{I}-\mathbf{P}) = \mathbf{0} \end{aligned}$$

where $\mathbf{I}$ is the identity matrix. A non-trivial solution of the above homogeneous equation exists only if $(\mathbf{I}-\mathbf{P})$ is singular.

It is convenient analyse the steady-state probabilities when they exist through the eigen-value decomposition of the matrix $\mathbf{P}$. Suppose $\mathbf{P}$ is an $N \times N$ stochastic matrix.

Theorem 3 Eigen values of the finite stochastic matrix

$\lambda = 1$ is one eigen value of $\mathbf{P}$ with an eigen vector $[1\,1 ... 1]'$. All other eigen values have magnitudes less than or equal to unity.

Proof:

Noting that $\mathbf{P}$ is a stochastic matrix, the row sum of each row is one. Therfore

$$\therefore \mathbf{P}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}=\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}=1\times\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}$$

Therefore, $\lambda=1$ is an eigen value of $\mathbf{P}$ with an eigen vector $[1\ 1 ... 1]'$.

Suppose λ is an eigen value with the corresponding eigen vector $\mathbf{v}=\begin{bmatrix}v_1\\v_2\\\vdots\\v_N\end{bmatrix}$.

Then, $\mathbf{Pv}=\lambda\mathbf{v}$.

Suppose $V_{\max}=\max(|v_1|,|v_2|,....,|v_N|)$. Since $\sum_{j=0}^{N-1}p_{ij}=1$, we have

$$\begin{aligned}
V_{\max} &= V_{\max}\sum_{j=0}^{N-1}p_{ij}\ ,\ \ i=0,1,...,N-1\\
&\geq \sum_{j=0}^{N-1}|v_j|\,p_{ij}\\
&= \sum_{j=0}^{N-1}|v_jp_{ij}|\\
&\geq \left|\sum_{j=0}^{N-1}v_jp_{ij}\right|,\ \ i=0,1,...,N-1\\
&= |\lambda v_i|,\ \ i=0,1,...,N-1
\end{aligned}$$

$$\therefore V_{\max}\geq|\lambda v_i|,\ i=0,1,...,N-1$$
$$\Rightarrow V_{\max}\geq|\lambda|V_{\max}$$
$$\Rightarrow|\lambda|\leq 1$$

Steady state probabilities using eigen decomposition of P matrix

Assuming that the eigen values are distinct (one eigen value is one and rest less than 1), we can diagonalize $\mathbf{P}$ as

$$\mathbf{M}^{-1}\mathbf{PM}=\mathbf{\Lambda}$$

where $\mathbf{M}$ is the modal matrix of eigen vectors and $\mathbf{\Lambda}$ is the diagonal matrix of eigen values.

$$\therefore \mathbf{P}=\mathbf{M\Lambda M}^{-1}$$

Suppose $\mathbf{\Pi} = \lim_{n\to\infty} \mathbf{P}^n$

Then

$$\begin{aligned}\mathbf{\Pi} &= \lim_{n\to\infty} \left(\mathbf{M\Lambda M^{-1}}\right)^n \\ &= \lim_{n\to\infty} \mathbf{M\Lambda}^n\mathbf{M^{-1}}\end{aligned}$$

$$\begin{aligned}\lim_{n\to\infty} \mathbf{\Lambda}^n &= \lim_{n\to\infty} \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_N \end{pmatrix}^n \\ &= \lim_{n\to\infty} \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & \lambda_2^n & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_N^n \end{pmatrix} \quad \because \lambda_i < 1,\ i \geq 2\end{aligned}$$

We have

$$= \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

Noting that one eigen vector is $[11\ldots1]'$, we get

$$\begin{aligned}\mathbf{\Pi} = \lim_{n\to\infty} \mathbf{M\Lambda}^n\mathbf{M^{-1}} &= \begin{bmatrix} 1 & m_{12} & m_{13}\ldots & m_{1N} \\ 1 & m_{22} & m_{23}\ldots & m_{2N} \\ \vdots & \vdots & & \vdots \\ 1 & m_{N2} & m_{N3}\ldots & m_{NN} \end{bmatrix} \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} m'_{11} & m'_{12} & m'_{13}\ldots & m'_{1N} \\ m'_{21} & m'_{22} & m'_{23}\ldots & m'_{2N} \\ \vdots & \vdots & & \vdots \\ m'_{N1} & m'_{N2} & m'_{N3}\ldots & m'_{NN} \end{bmatrix} \\ &= \begin{pmatrix} \pi_0 & \pi_1 & \ldots & \pi_n \\ \pi_0 & \pi_1 & \ldots & \pi_n \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \ldots & \pi_n \end{pmatrix}\end{aligned}$$

$$\lim_{n\to\infty}\mathbf{p}^{(n)} = \begin{bmatrix} p_0 & p_1 & \cdots & p_n \end{bmatrix}\begin{pmatrix} \pi_0 & \pi_1 & \cdots & \pi_n \\ \pi_0 & \pi_1 & \cdots & \pi_n \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \cdots & \pi_n \end{pmatrix}$$

$$= [\pi_0 \; \pi_1 ... \pi_n]$$

Therefore, steady state probabilities are independent of initial probabilities assuming that steady state probabilities exist.

Example 6 2-State Markov Chain:

Given $\mathbf{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$ and $\mathbf{p}^{(0)} = \begin{bmatrix} p & 1-p \end{bmatrix}$

Find the steady state probabilities assuming that they exists.

Solution

We have

$$\det\left|\lambda I - \mathbf{P}\right| = 0$$

$$\Rightarrow \det\begin{bmatrix} \lambda-1+a & -a \\ -b & \lambda-1+b \end{bmatrix} = 0$$

The roots are $\lambda_1 = 1$ and $\lambda_2 = 1-a-b$.

The corresponding eigen vectors are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ \frac{-b}{a} \end{pmatrix}$.

When $a \neq b = 1$,

$$\begin{aligned}
\mathbf{\Pi} &= \lim_{n\to\infty}\mathbf{P}^n \\
&= \lim_{n\to\infty}\mathbf{M}\mathbf{\Lambda}^n\mathbf{M}^{-1} \\
&= \begin{pmatrix} 1 & 1 \\ 1 & \dfrac{-b}{a} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{a}{a+b} & \dfrac{-a}{a+b} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{a}{a+b} & \dfrac{-a}{a+b} \end{pmatrix} \\
&= \begin{pmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\therefore \lim_{n\to\infty}\mathbf{p}^{(n)} &= \lim_{n\to\infty}\mathbf{p}\mathbf{P}^{(n)} \\
&= \begin{bmatrix} p & 1-p \end{bmatrix}\begin{pmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{pmatrix} \\
&= \begin{bmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \end{bmatrix}
\end{aligned}$$

When $a = b = 1$

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

$$\Rightarrow \mathbf{P}^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \mathbf{I} \text{ and}$$

$$\mathbf{P}^3 = \mathbf{P}^2\mathbf{P} = \mathbf{I}\mathbf{P} = \mathbf{P} \quad \text{and so on}$$

$$\therefore \mathbf{P}^n = \begin{cases} \mathbf{I}, & \text{when } n \text{ is even} \\ \begin{bmatrix} 0\,1 \\ 1\,0 \end{bmatrix}, & \text{when } n \text{ is odd} \end{cases}$$

$$\therefore \mathbf{p}^n = \begin{cases} \begin{bmatrix} p & 1-p \end{bmatrix}, & \text{when } n \text{ is even} \\ \begin{bmatrix} 1-p & p \end{bmatrix}, & \text{when } n \text{ is odd} \end{cases}$$

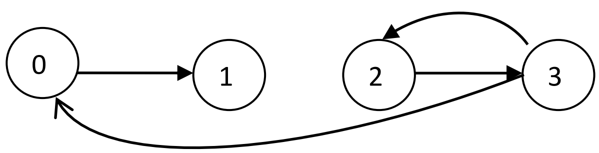
Classification states

The eigen decomosition method to find the steady state probabilities of the states requires the state space to be finite and the such decomposition may be difficult when the state space is large. The steady state distribution of these probabilities also may not exist. To anlyse a general MC, we need to classify the states.

Definition: A state j is said to be *accessible* from state i if the chain can reach j at some instant staring from the state i. In other words, $p_{ij}^{(n)} > 0$ for some $n \geq 0$. This property is denoted by $i \to j$.

The states i and j are said to *communicate* with each other if $i \to j$ and $j \to i$. We write $i \leftrightarrow j$ to denote that *i* communicates with *j*. Thus $i \leftrightarrow j$ if $i \to j$ and $j \to i$

Example 7:



In the above Markov chain, state 2 is not accessible from state 0. States 2 and 3 are communicative with each other.

Theorem 4: communication is an equivalence relation in the sense that

(1) it is reflective $i \leftrightarrow i$
(2) it is symmetric $i \leftrightarrow j \Rightarrow j \leftrightarrow i$
(3) it is transitive $i \leftrightarrow j$, $j \leftrightarrow k \Rightarrow i \leftrightarrow k$

The first property follows definition. Since $i \leftrightarrow j$, $p_{ii}^{(r)} \neq 0$ for some *r*. The symmetry property is obvious from the definition of communication.

To prove the third property first assume that *j* is accessible from *i* and *k* is accessible from *j*. Then, $\exists m, n \geq 0$ such that

$p_{ij}{}^{(m)} > 0$ and $p_{jk}{}^{(n)} > 0$

Therefore using CK equation:

$$\begin{aligned} p_{ik}{}^{(m+n)} &= \sum_{l} p_{il}{}^{(m)} p_{lk}{}^{(n)} \\ &\geq p_{ij}{}^{(m)} p_{jk}{}^{(n)} \\ &> 0 \end{aligned}$$

Thus *k* is accessible from *i.* Similarly *i* can be shown to be accessible from *k.*

Therefore,

$$i \leftrightarrow k$$

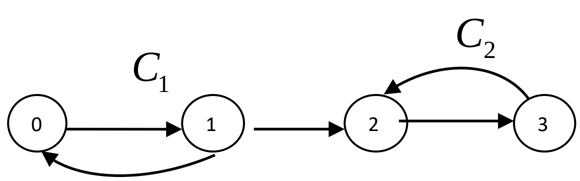
Definition: The relation $\leftrightarrow$ partitions the state space **V** of an MC into distinct subsets called *communicating classes*. The states belonging to a communicating class *C*, communicate with one another.

A communicating class C is called *closed* if any $j \in C'$ is not accessible from $i \in C$. Thus C is closed if $\forall i \in C$ and $j \in C'$

$p_{ij} = 0$.

Thus, $\sum_{j \in C} p_{ij} = 1$ and $\sum_{j \in C'} p_{ij} = 0$ for each $i \in C$.

Example 8:



In this example, C_1 is not closed but C_2 is closed.

Definition Irreducible Markov chain: The state transition matrix **P** is called irreducible if **V** is a communicating class and there is no other communicating class in **V**. The corresponding Markov Chain is called *irreducible.* An irreducible MC is always closed.

If a closed class C consist of only one state i, then i is called an *absorbing state*. Thus the sate *i* with $p_{ii} = 1$ is an absorbing state

Example 9: Gambler's ruin problem

A man starts to gamble with 2 units of money at hand which he wants to double. Each time he wins a game, he gets 1 unit and if he loses a game, he loses 1unit. The probability of the gambler winning a particular game is p and the play stops if the gambler wins 2 units or he losses2 units.

(a) Show that state transition matrix
(b) Find all the communicating classes

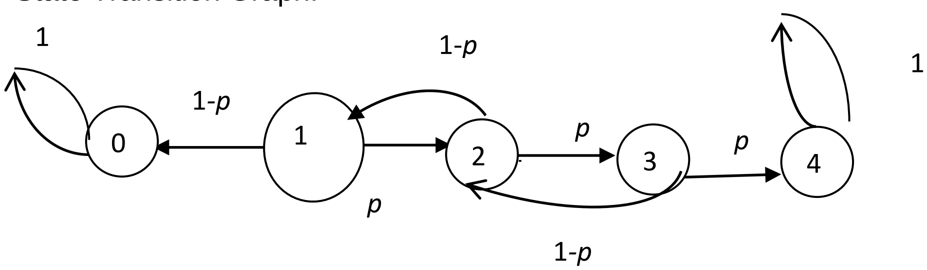
(c) Find the absorbing states.

<u>Solution:</u> $\{X_n, n \geq 0\}$ is the MC that represents the total fortune the gambler has after *n*th game. Claealy, the state space is given by $V = \{0,1,2,3,4\}$.

(a) State Transition Matrix:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1-p & 0 & p & 0 & 0 \\ 0 & 1-p & 0 & p & 0 \\ 0 & 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(b) State Transition Graph:



(c) Communicating classes: $\{0\},\{4\},\{1,2,3\}$

1-*p*

Closed classes:

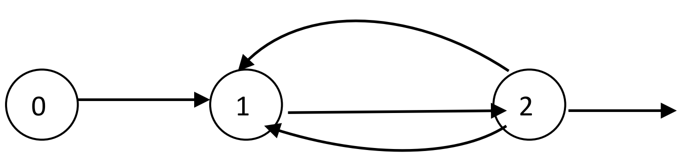
$\{0\},\{4\}$

Not closed:

$\{1,2,3\}$

'0'and '4'are absorbing states.

<u>Periodic and Aperiodic Markov Chains:</u> A state *i* is said to be *periodic* with a *period d*, if $p_{ii}^{(n)} = 0$, unless *n* is a multiple of *d* and *d* is the greatest integer with this property. Thus a path leading back to a state *i* has length in multiple of *d*. Thus in the following graph, each state is periodic with a period 2.



Thus,

$$d = \gcd\{n \mid p_{ii}^{(n)} > 0\}$$

Every d^{th} term of $p_{ii}^{(n)}$ is non-zero and the rest terms are zero. If $d=1$, then the chain is aperiodic .

Recurrence Property of States

This property is very helpful to determine the steady-state solution of the MC.
We define the terms:

First passage probability:
$f_{ij}^{(n)} = P\left(X_n = j, X_1 \neq j, X_2 \neq j, \cdots, X_{n-1} \neq j \mid X_0 = i\right)$
First return probability:
$f_{ii}^{(n)} = P\left(X_n = i, X_1 \neq i, X_2 \neq i, \cdots, X_{n-1} \neq i \mid X_0 = i\right)$
These probabilities must be distinguished from the n-step transition probability $p_{ij}^{(n)}$
We define
$f_{i,j}{}^{(0)} = 0$

Also

$f_{i,j}{}^{(1)} = p_{ij}$

$f_{i,j}{}^{(\infty)} = P\left(X_n \neq j / X_0 = i\right) \text{ for } \quad n \geq 1$

For each pair of states (i, j), we can define
$f_{ij} =$ the probability that the chain will pass through state j given that it is at i at $n=0$.

$$\therefore\ f_{ij} = \sum_{n=1}^{\infty} P\left(X_n = j, X_k \neq j, k = 1, 2, \cdots, n-1 \mid X_0 = i\right)$$
$$= \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

$f_{ij}^{(n)}$ is a *defective probability distribution* in the sense that $f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$ may not be equal to 1.
Similarly, the probability that the chain will not return to state i given that it has started at state *i.*
$$f_{ii} = \sum_{n=1}^{\infty} P\left(X_n = i, X_k \neq i, k = 1, 2, \cdots, n-1 \mid X_0 = i\right)$$
$$= \sum_{n=1}^{\infty} f_{ii}^{(n)}$$
As f_{ii} may not be equal to 1,
$1 - f_{ii}$ = probability that the chain will not return to state i.
We can derive a relationship between $p_{ij}^{(n)}$ and $f_{ij}^{(n)}$. This relationship is helpful to find out the steady state behavior of an MC.

Theorem 5:
The first passage probabilities and the first return probabilities are related to the state transition probabilities by

$$p_{ij}^{(n)} = \sum_{m=1}^{n} f_{ij}^{(m)} p_{jj}^{(n-m)}$$

$$p_{ii}^{(n)} = \sum_{m=1}^{n} f_{ii}^{(m)} p_{ii}^{(n-m)}$$

Proof:
The event that state j can be reached from the state i in n steps can be written as the union of disjoint events as:

$$\{X_n = j\} = \bigcup_{m=1}^{n} \{X_n = j\} \cap \{X_m = j, X_{m-k} \neq j; k = 1, 2, \cdots, m-1\}$$

$$\therefore\ p_{ij}^{(n)} = P\left(X_n = j \mid X_0 = i\right)$$

$$= \sum_{m=1}^{n} P\left(X_n = j, X_m = j, X_{m-k} \neq j, k = 1, 2, \cdots, m-1 \mid X_0 = i\right)$$

$$= \sum_{m=1}^{n} P\left(X_m = j, X_{m-k} \neq j, k = 1, 2, \cdots, m-1 \mid X_0 = i\right) P\left(X_n = j \mid X_m = j, X_{m-k} \neq j, k = 1, 2, \cdots, m-1, X_0 = i\right)$$

$$= \sum_{m=1}^{n} f_{ij}^{(m)} P\left(X_n = j \mid X_m = j\right) \qquad \text{(using Markov property)}$$

$$= \sum_{m=1}^{n} f_{ij}^{(m)} p_{jj}^{(n-m)}$$

Thus,

$$p_{ij}^{(n)} = \sum_{m=1}^{n} f_{ij}^{(m)} p_{jj}^{(n-m)}$$

Similarly,

$$p_{ii}^{(n)} = \sum_{m=1}^{n} f_{ii}^{(m)} p_{ii}^{(n-m)} \qquad \text{for } n = 1, 2, \cdots$$

The above relations are convolution sums of two discrete sequences and can be solved using the probability generating function (z transform)

Probability Generating Functions of the state transition and first passage probabilities

Let us define $P_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}^{(n)} z^n$

where the region of convergence is given by $|z| < 1$.

Similarly,

$$F_{ij}(z) = \sum_{n=0}^{\infty} f_{ij}^{(n)} z^n, \qquad |z| < 1$$

$$= \sum_{n=1}^{\infty} f_{ij}^{(n)} z^n \quad (\because f_{ij}^{(0)} = 0)$$

We also note the following powerful theorems:

Abel's theorem

Suppose $A(z) = \sum_{n=0}^{\infty} a_n z^n$ is convergent for $|z| < 1$ and $a_n \geq 0, \ \forall n$. Then

$$\lim_{z \to 1^-} A(z) = \sum_{n=0}^{\infty} a_n$$

irrespective of whether $\sum_{n=0}^{\infty} a_n$ converges or diverges. A consequence of the Abel's theorem is the final value theorem:

Final value theorem

If the sequence $\{a_n\}$ converges, then

$$\lim_{n \to \infty} a_n = \lim_{z \to 1^-} (1-z) A(z)$$

Relations in the z-domain

Let us first consider the relation

$$p_{ii}^{(n)} = \sum_{m=1}^{n} f_{ii}^{(m)} p_{ii}^{(n-m)} \qquad n = 1, 2, \cdots$$

For $n = 0,\ p_{ii}^{(0)} = 1$

Taking the z-transform, we get

$$P_{ii}(z) = 1 + F_{ii}(z) P_{ii}(z)$$

$$\therefore P_{ii}(z) = \frac{1}{1 - F_{ii}(z)}$$

Similarly, considering

$$p_{ij}^{(n)} = \sum_{m=1}^{n} f_{ij}^{(m)} p_{jj}^{(n-m)},$$

we can show that

$$P_{ij}(z) = F_{ij}(z) P_{jj}(z)$$

$$= \frac{F_{ij}(z)}{1 - F_{jj}(z)}$$

Thus ,

$$P_{ii}(z) = \frac{1}{1-F_{ii}(z)}$$

and

$$P_{ij}(z) = \frac{F_{ij}(z)}{1-F_{jj}(z)}$$

Steady State analysis:

Once we have the probability generating function, we can get the $p_{ii}^{(n)}$ sequence. To find the steady state values of $p_{ij}^{(n)}$, one can apply the final value theorem. Here we exclude the case of periodic states where we cannot apply the same analysis.

From $P_{ii}(z) = \frac{1}{1-F_{ii}(z)}$, we have

$$\lim_{n\to\infty} p_{ii}^{(n)} = \lim_{z\to 1^-}(1-z)P_{ii}(z)$$

$$= \lim_{z\to 1^-}(1-z)\frac{1}{1-F_{ii}(z)}$$

and

$$\lim_{n\to\infty} p_{ij}^{(n)} = \lim_{z\to 1^-}(1-z)P_{ij}(z)$$

$$= \lim_{z\to 1^-}(1-z)\frac{F_{ij}(z)}{1-F_{jj}(z)}$$

If $F_{ii}(1) = f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)} < 1$,

we see that

$$\lim_{n\to\infty} p_{ii}^{(n)} = 0$$

In fact, for the condition $f_{ii} < 1$, we can show that $\lim_{n\to\infty} p_{kj}^{(n)} \to 0 \;\; \forall k$

For $f_{ii} = 1$, the above procedure fails.

We can apply the L'Hospital rule to find $\lim_{z\to 1^-}(1-z)\frac{1}{1-F_{ii}(z)}$

$$\therefore \lim_{n\to\infty} p_{ii}^{(n)} = \frac{1}{F_{ii}^{'}(1)} = \frac{1}{\mu_{ii}}$$

where $\mu_{ii} = \sum_{n=1}^{\infty} n f_{ii}^{(n)} = F_{ii}^{/}(1)$ is the mean recurrence time for state *i*.

Similarly, we can apply the L'Hospital rule to find $\lim_{z\to 1^-}(1-z)\frac{F_{ij}(z)}{1-F_{ii}(z)}$

$$\therefore \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{z\to 1^-}(1-z)\frac{F_{ij}(z)}{1-F_{ii}(z)}$$

$$= \frac{F_{ij}(1)}{F_{ii}^{'}(1)}$$

$$= \frac{f_{ij}}{\mu_{ii}}$$

where $F_{ij}(1) = f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$

We have seen that f_{ii} plays an important role in the steady-state behavior of the chain. The states can be classified on the basis of the value of f_{ii}.

Definition (Transient and Recurrent State):
A state i is called *transient* if $f_{ii} < 1$. Otherwise it is called *a recurrent* state.

Theorem 6: (a) If a state i is transient then $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$.

(b) If a state i is recurrent, then $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.

Proof: (a)

$f_{ii} < 1$

$\Rightarrow \lim_{z\to 1^-} F_{ii}(z) < 1$

$\text{Now } \lim_{z\to 1^-} P_{ii}(z) = \lim_{z\to 1^-} \frac{1}{1-F_{ii}(z)} < \infty$

$\Rightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$

(b) We have $\lim_{z\to 1^-} \frac{1}{1-F_{ii}(z)} = \infty$, so that

$\lim_{z\to 1^-} P_{ii}(z) = \infty$

$\Rightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$

Corrolary: A state i is transient, then $\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty$.

We have

$P_{ij}(z) = F_{ij}(z)P_{ii}(z)$

$\therefore P_{ij}(1^-) = F_{ij}(1^-)P_{ii}(1^-)$

$\Rightarrow \sum_{n=0}^{\infty} p_{ij}^{(n)} = f_{ii}\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$

Corrolary: A state i is recurrent, then $\sum_{n=0}^{\infty} p_{ij}^{(n)} = \infty$.

We have

$$P_{ij}(z) = F_{ij}(z)P_{jj}(z)$$

$$\therefore P_{ij}(1^-) = F_{ij}(1^-)P_{jj}(1^-)$$

$$\Rightarrow \sum_{n=0}^{\infty} p_{ij}^{(n)} = f_{ii}\sum_{n=0}^{\infty} p_{jj}^{(n)} = \infty$$

Theorem 5: All the states in a communicating class are of same type:

(a) all transient,
(b) all recurrent and
(c) all periodic,

Proof:

Suppose C is a communicating class, $i, j \in C$ and $i \leftrightarrow j$. Then there exists $m > 0$ and $r > 0$ such that $p_{ij}^{(m)} > 0$ and $p_{ji}^{(r)} > 0$.

Now for any $n > 0$,

$$\begin{aligned} p_{ii}^{(m+n+r)} &= \sum_k p_{ik}^{(m+n)} p_{ki}^{r} \quad (\text{Using C} - \text{K equation})) \\ &\geq p_{ij}^{(m+n)} p_{ji}^{r} \\ &= \sum_l p_{il}^{(m)} p_{lj}^{(n)} p_{ji}^{(r)} \end{aligned}$$

$$\therefore p_{ii}^{(m+n+r)} \geq \sum_l p_{il}^{(m)} p_{lj}^{(n)} p_{ji}^{(r)}$$

Changing the order of *l* and *j*, we also get

$$p_{jj}^{(m+n+r)} \geq p_{ji}^{(r)} p_{ii}^{(n)} p_{ij}^{(m)}$$

(a) Suppose *i* is transient. Then,

$$\sum_n p_{ii}^{(n)} < \infty$$

Now,

$$\begin{aligned} \sum_n p_{ii}^{(n)} &\geq \sum_n p_{ii}^{(m+n)} \\ &\geq \sum_n p_{ij}^{(m)} p_{ji}^{(r)} p_{jj}^{(n)} \end{aligned}$$

$$\therefore \sum_n p_{ij}^{(m)} p_{ji}^{(r)} p_{jj}^{(n)} < \sum_n p_{ii}^{(n)} < \infty$$
$$\Rightarrow \sum_n \alpha\, p_{jj}^{(n)} < \infty, \text{ where } \alpha = p_{ij}^{(m)} p_{ji}^{(r)} > 0$$
$$\Rightarrow \sum_n p_{jj}^{(n)} < \infty$$

Therefore, the state *j* is also transient

(b) Now suppose *i* is recurrent. Then,

$$\sum_n p_{ii}^{(n)} \to \infty$$

Now,

$$\begin{aligned}\sum_n p_{j,j}^{(n)} &\geq \sum_n p_{j,j}^{(m+n+r)} \\ &\geq \sum_n p_{ij}^{(m)} p_{ii}^{(n)} p_{ji}^{(r)} \\ &= \sum_n \alpha\, p_{ii}^{(n)}, \text{ where } \alpha \text{ is as defined above.}\end{aligned}$$
$$\therefore \sum_n p_{jj}^{(n)} \to \infty$$

Therefore, the state *j* is also recurrent.

(c) Suppose *i* is a periodic state with the period d_i.

Now

$$\begin{aligned}p_{ii}^{(m+r)} &= \sum_k p_{ik}^{(m)} p_{ki}^{(r)} \\ &\geq p_{ij}^{(m)} p_{ji}^{(r)}\end{aligned}$$

$$\therefore p_{ii}^{(m+r)} > 0$$

Thus (*m*+*r*) must be a multiple of d_i.

Finally,

$$p_{ii}^{(m+n+r)} \geq p_{ij}^{(m)} p_{jj}^{(n)} p_{ji}^{(r)}$$
$$\Rightarrow p_{jj}^{(n)} \leq \frac{p_{ii}^{(m+n+r)}}{\alpha} \text{ where } \alpha \text{ is as defined above}$$

Now $p_{ii}^{(m+n+r)} = 0$ unless *n* is multiple of d_i as *m+r* is already a multiple of d_i.

Therefore, $p_{jj}^{(n)} = 0$ at all *n* which are not multiple of d_i.

Suppose $n = k_1 d_i > 0$ at which $p_{jj}^{(n)} > 0$.

Then $P_{jj}^{(2kd_i)} \geq p_{jj}^{kd_i} p_{jj}^{kd_i} > 0$

Similarly $P_{jj}^{(3kd_i)} > 0$ and so on. Thus the state *j is* periodic and the period is

$d_j = k_1 d_i$

The role of *i* and *j* now can be interchanged and we can show that

$d_i = k_2 d_j$

where k_2 is another positive integer.

From the above two results

$d_j = k_1 k_2 d_j$

which happens only when $k_1 = k_2 = 1$.

$\therefore d_i = d_j$

Theorem 6:

If an MC is irreducible and recurrent, then for all (i,k), $f_{ik} = 1$.

Proof:

The chain is recurrent. Therefore,

$$f_{ii} = 1, \quad \forall i$$

The chain is irreducible. Therfore, each *k*, there exists *m* such that $p_{ik}^{(m)} > 0$.Now $f_{ii} = 1$ implies that

$$
\begin{aligned}
1 &= P(X_n = i \text{ for some } n > m / X_0 = i) \\
&= \sum_k P(X_n = i \text{ for some } n > m, X_m = k / X_0 = i) \\
&= \sum_k P(X_m = k / X_0 = i) P(X_n = i \text{ for some } n > m / X_m = k, X_0 = i) \\
&= \sum_k p_{ik}^{(m)} P(X_n = i \text{ for some } n > m / X_m = k) \quad (\text{ Using the Markov property}) \\
&= \sum_k p_{ik}^{(m)} f_{ki} \\
\therefore & \sum_k p_{ik}^{(m)} f_{ki} = 1
\end{aligned}
$$

We also have,

$$
\begin{aligned}
&\sum_k p_{ik}^{(m)} = 1 \\
&\therefore \sum_k (1 - f_{ki}) p_{ik}^{(m)} = 0
\end{aligned}
.
$$

$\therefore f_{ki} = 1$, wherever $p_{ik}^{(m)} > 0$. As the chain is irreducible, there exists m for each (i, k) such that $p_{ik}^{(m)} > 0$. Thus $f_{ki} = 1$.

Definition Null recurrent and positive recurrent state

A recurrent state is called *null recurrent* if $\mu_{ii} = \infty$. The recurrent state is *positive recurrent* if $\mu_{ii} < \infty$.

Theorem: In an irreducible positively recurrent MC, $p_{ij}^{(n)} = \dfrac{1}{\mu_{jj}}$.

Proof:

We have earlier established that for an aperiodic irreducible recurrent state,

$$
\begin{aligned}
&\lim_{n\to\infty} p_{ij}^{(n)} = \frac{f_{ij}}{\mu_{jj}} \\
&\therefore \lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_{jj}}
\end{aligned}
$$

Definition Ergodic Markov Chain

A positive recurrent aperiodic MC is called ergodic. For an ergodic state with finite state space

$$\Pi_j = \lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_{jj}}$$

Theorem 7 Consider an MC with a finite state space $V = \{0,1,...,N-1\}$ .If the chain is aperiodic, irreducible and positive recurrent,then

$$\Pi_j = \frac{1}{\mu_{jj}}, \; j = 0,1,...,N-1$$

constitute a probability distribution and are the solutions of the equations

$$\Pi_j = \sum_{k=0}^{N-1} \Pi_k p_{kj}, \; j = 0,1,...,N-1$$

Proof: The above relation is true only for positive recurrent states as we know that $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for transient and null-recurrent states. For any state *j* of an MC,

$$\sum_{j=0}^{N-1} p_{ij}^{(n)} = 1$$

$$\therefore \lim_{n\to\infty} \sum_{j=0}^{N-1} p_{ij}^{(n)} = 1$$

$$\therefore \sum_{j=0}^{N-1} \lim_{n\to\infty} p_{ij}^{(n)} = 1$$

$$\Rightarrow \sum_{j=0}^{N-1} \Pi_j = 1$$

In the third step, we have taken the limit inside the summation as the summation is over a finite number of terms.

Let us consider C-K equation

$$p_{ij}{}^{(n+1)} = \sum_{k=0}^{N-1} p_{ik}^{(n)} p_{kj}$$

Taking limits of both sides,

$$\lim_{n\to\infty} p_{ij}{}^{(n+1)} = \sum_{k=0}^{N-1} \lim_{n\to\infty} p_{ik}^{(n)} p_{kj}$$

$$\Rightarrow \Pi_j = \sum_{k=0}^{N-1} \Pi_k p_{kj}$$

The proof of the above theorem required the finiteness of the state space. However, the result is more general and we state the following theorem without proof.

Theorem: Consider an aperidic and irreducible MC with a countable state space $V = \{0,1,...\}$. The necessary and sufficient condition for $\Pi_j = \lim_{n\to\infty} p_{ij}^{(n)},\ j = 0,2,...$ to be a probability distribution is that the MC is positive recurrent. In this case, Π_j s are the solutions of the equations

$$\Pi_j = \sum_{k=0}^{\infty 1} \Pi_k p_{kj},\ j = 0,1,...$$

and related to the mean recurrence time by the relation

$$\Pi_j = \frac{1}{\mu_{jj}},\ j = 0,1,...$$