Week 4

Markov chains: Limit theorems, stationary distribution and absorption probability.

10 Markov Chains: limit distribution

Consider a Markov chain $\{X_n\}_{n\geq 0}$ with $S = \{1, 2, \cdots, r\}$ and transition matrix P. Suppose that we know the probability distribution of X_0, i.e.,

$$\pi^{(0)} = \left[P(X_0 = 1), \;\; P(X_0 = 2), \;\; \cdots, \;\; P(X_0 = r)\right].$$

We have that

$$\begin{aligned} \pi^{(n+1)} &= \pi^{(n)} P, \;\; \text{for } n = 0, 1, 2, \cdots; \\ \pi^{(n)} &= \pi^{(0)} P^n, \;\; \text{for } n = 0, 1, 2, \cdots, \end{aligned}$$

where

$$\pi^{(n)} = \left[P(X_n = 1), \;\; P(X_n = 2), \;\; \cdots, \;\; P(X_n = r)\right],$$

See Lect 5.

In many applications, we would like to discuss long-term behavior of the MC. More precisely, we would like to consider the limits:

$$\pi_j := \lim_{n\to\infty} P(X_n = j), \quad j \in S,$$

(In physical terms, if the MC describes movements of some particle, this is a question about the asymptotic stability of the particle.)

The probability distribution $\pi = \{\pi_j, j \in S\}$ is called the **limiting distribution** of the Markov chain X_n if

$$\pi_j = \lim_{n \to \infty} P(X_n = j | X_0 = i)$$

for all $i, j \in S$, and we have

$$\sum_{j \in S} \pi_j = 1.$$

By the definition, when a limiting distribution exists, it does not depend on the initial state $(X_0 = i)$, so we can write

$$\pi_j = \lim_{n \to \infty} P(X_n = j), \text{ for all } j \in S.$$

Remark: not all Markov chains have a limit distribution. For example, consider a MC with the transition matrix:

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

In this case, the chain has a periodic behavior, i.e.,

$$X_{n+2} = X_n, \quad \text{for all } n$$

can be a MC with the transition matrix P. In particular,

$$X_n = \begin{cases} X_0 & \text{if } n \text{ is even} \\ X_1 & \text{if } n \text{ is odd} \end{cases} \tag{1}$$

In this case, the distribution of X_n does not converge to a single initial probability distribution. Also, the distribution of X_n depends on the initial distribution.

Question: when does a Markov chain have a limiting distribution (that does not depend on the initial probability distribution)?

Recall that

$$P(X_n = j) = \sum_{k \in S} \pi_k^{(0)} \, p_{kj}^{(n)},$$

where $\pi^{(0)} = (\pi_k^{(0)})_{k \in S}$ is the initial distribution. The existence of π depends on $\lim_{n\to\infty} p_{ij}^{(n)}$. (The asymptotic behavior of $p_{ij}^{(n)}$ depends on the nature of states i and j which is determined by the transition matrix P)

The following basic result comes from Feller (1970): An introduction to probability theory and its applications, page 321.

Basic result: Let $\{X_n, n \geq 0\}$ be a MC with state space S and transition matrix P.

(a). A state j is transient iff $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$;

(b). A state j is null recurrent iff $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ and $p_{jj}^{(n)} \to 0$, as $n \to \infty$;

(c). A state j is positive recurrent $\Longrightarrow \sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ and

(i). if j is aperiodic, then $\mu_j < \infty$ and, as $n \to \infty$,

$$p_{jj}^{(n)} \to 1/\mu_j, \quad \text{where } \mu_j = \textstyle\sum_{n=1}^{\infty} n \, f_{jj}^{(n)},$$

(ii). if j has period d, then $\mu_j < \infty$ and, as $n \to \infty$,

$$p_{jj}^{(nd)} \to d/\mu_j,$$

(while, of course, $p_{jj}^{(n)} = 0$ for all n not divisible by d).

Question: when does a Markov chain have a limiting distribution (that does not depend on the initial probability distribution)?

By using the basic result, we have the following theorem.

Th4.1. Let $\{X_n, n \geq 0\}$ be an irreducible aperiodic MC, i.e., the MC has only one class with period $d = 1$,

(a). If $\{X_n, n \geq 0\}$ is a positive recurrent chain, then the chain has a limit distribution $\pi = \{\pi_j, j \in S\}$ in which

$$\pi_j = \lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} p_{ij}^{(n)} = 1/\mu_j > 0, \tag{2}$$

where $\mu_j = \sum_{n=1}^{\infty} n\, f_{jj}^{(n)}$, the mean recurrence time of state j, and the limit distribution π is the unique solution to the equations:

$$\pi_j \;=\; \sum_{k\in S} \pi_k\, p_{kj}, \quad j \in S, \quad \sum_{j\in S} \pi_j = 1. \tag{3}$$

(c). If $\{X_n, n \geq 0\}$ is a null recurrent or a transient chain, then for all $i, j \in S$,

$$\pi_j = \lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} p_{ij}^{(n)} = 0.$$

Note: In general, a finite Markov chain can consist of several transient as well as recurrent states. As n becomes large the chain will enter a recurrent class and it will stay there forever. Therefore, when studying long-run behaviors we usually focus only on the recurrent classes.

Th4.2. If for all $i, j \in S$,

$$\lim_{n\to\infty} p_{ij}^{(n)} = \tilde{\pi}(j), \tag{4}$$

(i.e., the limit of $\lim_{n\to\infty} p_{ij}^{(n)}$ exists and is independent of i), then

(a). $\pi_j = \lim_{n\to\infty} P(X_n = j) = \tilde{\pi}(j),\ j \in S$;

(b). $\sum_{i\in S} \tilde{\pi}(i) \le 1$;

(c). $\sum_{i\in S} \tilde{\pi}(i) p_{ij} = \tilde{\pi}(j)$, $j \in S$;

(d). either $\sum_{i\in S} \tilde{\pi}(i) = 1$ or $\tilde{\pi}(j) = 0$ for any $j \in S$.

Example 4.1. Consider a Markov chain with $S = \{0, 1\}$ and the transition matrix is given by

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix},$$

where $0 \le a \le 1$ and $0 \le b \le 1$. Suppose that

$$\pi^{(0)} = \begin{bmatrix} P(X_0 = 0), & P(X_0 = 1) \end{bmatrix} = \begin{bmatrix} \alpha, & 1-\alpha \end{bmatrix},$$

where $\alpha \in [0, 1]$.

Note that

- if $a = b = 0$, then

$$P^n = P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix};$$

- if $0 < a + b < 2$, then

$$P^n = \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b)^n}{a+b} \begin{bmatrix} a & -a \\ -b & b \end{bmatrix};$$

- if $a = b = 1$, then

$$P^{2k} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad P^{2k+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We have:

- if $a = b = 0$, then

$$\big(\pi_0, \pi_1\big) \;=\; \pi^{(0)};$$

- if $0 < a + b < 2$, then

$$\big(\pi_0, \pi_1\big) \;=\; \lim_{n\to\infty} \pi^{(n)} = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}.$$

- if $a = b = 1$, $\lim_{n\to\infty} \pi^{(n)}$ does not exist.

11 Markov chain: stationary distribution

Let $\{X_n, n \geq 0\}$ be a MC with state space S and transition matrix P.

Let $\pi = \{\pi_j, j \in S\}$ be a probability distribution, i.e, $\pi_j \geq 0$ and $\sum_{j \in S} \pi_j = 1$.

If $\pi = \pi\, P$, i.e.,

$$\pi_j \;=\; \sum_{k \in S} \pi_k\, p_{kj}, \quad \text{for all } j \in S, \tag{5}$$

then π is said to be a **stationary distribution** of the MC.

Note that if the initial distribution $\pi^{(0)}$ is stationary, i.e., $\pi^{(0)} = \pi^{(0)} P$ then X_n's have the same distribution (hence the MC is a strictly stationary process), since $\pi^{(n)} = \pi^{(0)} P^n = \pi^{(0)} P^{n-1} = ... = \pi^{(0)} P = \pi^{(0)}$.

Note that limit distribution is not the exact same as stationary distribution.

Example 4.2. Let $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. $\pi = (1/2, 1/2)$ is a stationary distribution, but $\lim_{n\to\infty} p_{ij}^{(n)}$ does not exist (no limit distribution).

Th4.3. Let $\{X_n, n \geq 0\}$ be an irreducible and aperiodic MC. Suppose that the chain has a stationary distribution $\pi = \{\pi_j, j \in S\}$. Then,

(a). the MC is positive recurrent;

(b). the limit dist exists and π is the limit dist, i.e., we have

$$\pi_j = \lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} p_{ij}^{(n)} = 1/\mu_j > 0, \quad j \in S,$$

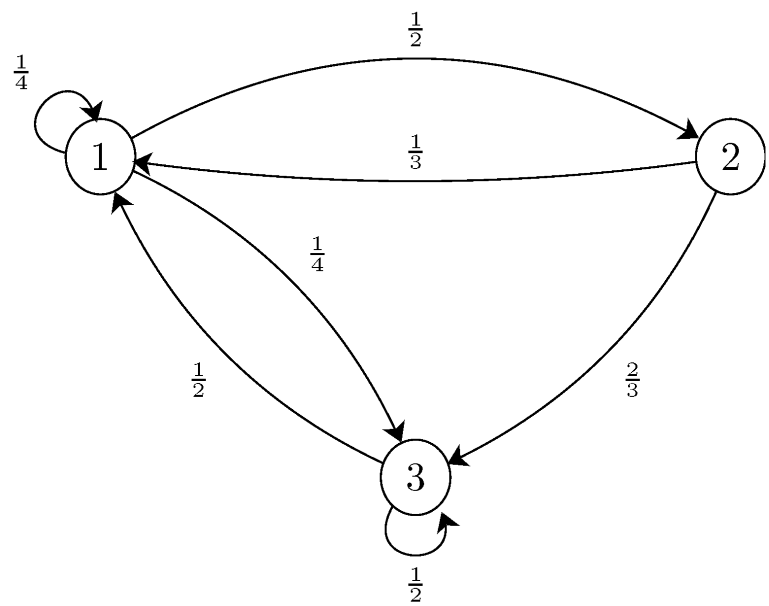
where $\mu_j = \sum_{n=1}^{\infty} n\, f_{jj}^{(n)}$;

(c). The stationary distribution is unique.

Proof. The results come from Th4.1 and Th4.2 (Lect 10). In fact, (a) follows from (c) of Th 4.1. (b) follows from (a) and Th 4.1. (c) comes from Th 4.2.

Comments: (in Feller's book) an irreducible MC with a finite state space has a unique stationary distribution. In general, a MC with a finite state space has at least one stationary distribution. It is not necessarily so if the state space is infinite. In fact, an irreducible transient (or null recurrent) chain has no stationary distribution.

Example 4.3 Consider the Markov chain shown in the following state transition diagram:



- Is this chain irreducible?
- Is this chain aperiodic?
- Find the stationary distribution for this chain.
- Is the stationary distribution a limiting distribution for the chain?
- What is the mean recurrence time for state i?

Solution:

- The chain is irreducible since we can go from any state to any other states in a finite number of steps.
- Since there is a self-transition, i.e., $p_{11} > 0$, we conclude that the chain is aperiodic.

- To find the stationary distribution $\pi = (\pi_1, \pi_2, \pi_3)$, we need to solve

$$\begin{aligned}
&\pi_1 = \frac{1}{4}\pi_1 + \frac{1}{3}\pi_2 + \frac{1}{2}\pi_3, \\
&\pi_2 = \frac{1}{2}\pi_1, \\
&\pi_3 = \frac{1}{4}\pi_1 + \frac{2}{3}\pi_2 + \frac{1}{2}\pi_3, \\
&\pi_1 + \pi_2 + \pi_3 = 1.
\end{aligned}$$

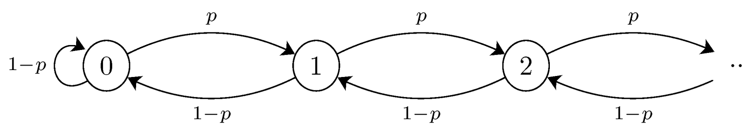
  We find

$$\pi_1 = \frac{3}{8},\ \pi_2 = \frac{3}{16},\ \pi_3 = \frac{7}{16}.$$

- Since the chain is irreducible and aperiodic, we conclude that the above stationary distribution is a limiting distribution.

- $\mu_1 = 1/\pi_1 = 8/3$.

Example 4.4. Consider the Markov chain shown in the following state diagram:



Assume that $0 < p < \frac{1}{2}$. Does this chain have a limiting distribution?

Solution: This chain is irreducible since all states communicate with each other. It is also aperiodic since it includes a self-transition, $p_{00} > 0$. Let's write the equations for a stationary distribution. For state 0, we can write

$$\pi_0 = (1-p)\pi_0 + (1-p)\pi_1,$$

which results in

$$\pi_1 = \frac{p}{1-p}\pi_0.$$

For state 1, we can write

$$\begin{aligned}\pi_1 &= p\pi_0 + (1-p)\pi_2 \\ &= (1-p)\pi_1 + (1-p)\pi_2,\end{aligned}$$

which results in

$$\pi_2 = \frac{p}{1-p}\pi_1.$$

Similarly, for any $j \in \{1, 2, \cdots\}$, we obtain

$$\pi_j = \alpha\pi_{j-1},$$

where $\alpha = \frac{p}{1-p}$. Note that since $0 < p < \frac{1}{2}$, we conclude that $0 < \alpha < 1$. We obtain

$$\pi_j = \alpha^j \pi_0, \quad \text{for } j = 1, 2, \cdots .$$

Finally, we must have

$$\begin{aligned}
1 &= \sum_{j=0}^{\infty} \pi_j \\
&= \sum_{j=0}^{\infty} \alpha^j \pi_0, && \text{(where } 0 < \alpha < 1\text{)} \\
&= \frac{1}{1-\alpha} \pi_0 && \text{(geometric series).}
\end{aligned}$$

Thus, $\pi_0 = 1 - \alpha$. Therefore, the stationary distribution is given by

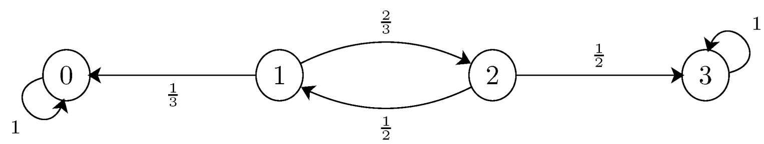
$$\pi_j = (1-\alpha)\alpha^j, \quad \text{for } j = 0, 1, 2, \cdots .$$

Since this chain is irreducible and aperiodic and we have found a stationary distribution, we conclude that all states are positive recurrent and $\pi = [\pi_0, \pi_1, \cdots]$ is the limiting distribution.

12 Markov chains: Absorption probabilities

Example 4.5 Consider a MC with the following state transition matrix:

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



There are three classes: $\{0\}, \{1, 2\}$ and $\{3\}$.

- $\{1, 2\}$ is a transient class.
- $\{0\}$ and $\{3\}$ are **absorbing**, which are recurrent states, i.e., the process will eventually get absorbed in one of them.

We are interested in finding the absorption probabilities: for $i = 0, 1, 2, 3$,

$$a_i = P(\text{absorption in } 0 \mid X_0 = i) = P\big(\cup_{n=1}^{\infty} (X_n = 0) \mid X_0 = i\big),$$

and/or

$$b_i = P(\text{absorption in } 3 \mid X_0 = i) = P\big(\cup_{n=1}^{\infty} (X_n = 3) \mid X_0 = i\big).$$

Solution. Note that

$$a_i = P(\text{absorption in } 0 \mid X_0 = i) = P(\text{absorption in } 0 \mid X_1 = i).$$

It follows from the law of total probability that, for $i = 0, 1, 2, 3$,

$$\begin{aligned} a_i &= \sum_{k=0}^{3} P(\text{absorption in } 0, X_1 = k \mid X_0 = i) \\ &= \sum_{k=0}^{3} a_k \, p_{ik}. \end{aligned}$$

This yields that $a = P\,a$, where $a = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$. Hence,

$$\begin{aligned} a_0 &= a_0 = 1, \\ a_1 &= \frac{1}{3}a_0 + \frac{2}{3}a_2, \\ a_2 &= \frac{1}{2}a_1 + \frac{1}{2}a_3, \\ a_3 &= a_3 = 0. \end{aligned}$$

Solving a_i, we obtain

$$a_0 = 1, \quad a_1 = \frac{1}{2}, \quad a_2 = \frac{1}{4}, \quad a_3 = 0.$$

Since $a_i + b_i = 1$, we conclude

$$b_0 = 0, \quad b_1 = \frac{1}{2}, \quad b_2 = \frac{3}{4}, \quad b_3 = 1.$$

Absorption probabilities: in general

Let T be a set of transient states of a MC with state space S. Let $C_m, m \geq 1$ be closed recurrent classes of a MC. Recall that

$$S = T \cup (\cup_{m=1}^{\infty} C_m).$$

If a MC starts at a recurrent state i in C_m, then since C_m is closed, it will **never** leave C_m, i.e.,

$$a_i(C_m) \;=\; P(\text{absorption in } C_m \mid X_0 = i) = 1, \quad i \in C_m.$$

If the starting state i is in T, the chain can be **absorbed** into one of the C_m's. We have

$$\begin{aligned} a_i(C_m) \;&:=\; P(\text{absorption in } C_m \mid X_0 = i) = \sum_{k \in S} p_{ik}\, a_k(C_m) \\ &=\; \sum_{k \in C_m} p_{ik} + \sum_{k \in T} p_{ik}\, a_k(C_m), \quad i \in T. \end{aligned} \tag{6}$$

In particular, if $C_m = \{m\}$ is absorbing, the **absorption probabilities** $a_i(m) = P(\text{absorption in } m \mid X_0 = i)$ satisfy the equations:

$$a_i(m) \;=\; p_{im} + \sum_{k \in T} p_{ik}\, a_k(m), \quad i \in T. \tag{7}$$

The following results are also well-known:

- If T is finite, then (6) has a unique solution. It is not necessary if T is infinite.

Mean absorbtion time

Let $A = \cup_{m=1}^{\infty} C_m$ and $\tau = \min\{n \geq 1 : X_n \in A\}$ or $\tau = \min\{n \geq 1 : X_n \notin T\}$, where $T = S - A$ is the transient set.

τ is called the **"absorbtion time"**. Define

$$m_i = E(\tau \mid X_0 = i), \quad i \in T = S - A.$$

Then m_i is the **"mean absorbtion time"** (starting from the transient i).

Assume that $P(\tau < \infty \mid X_0 = i) = 1$ for all $i \in T = S - A$ (This is the case if the state space S is finite). We have

$$m_i = 1 + \sum_{k \in T} p_{ik}\, m_k, \quad i \in T = S - A. \tag{8}$$

Proof. Indeed, by letting $m_k = 0$ if $k \in A$ and noting

$$\begin{aligned} E(\tau \mid X_1 = k) &= 1 + E(\tau \mid X_0 = k) = 1 + m_k, \quad k \in S, \\ E[\tau I_{X_1 = k} | X_0 = i] &= p_{ik}\, E[\tau | X_1 = k], \quad k \in S, \end{aligned}$$

it follows that

$$\begin{aligned} E[\tau | X_0 = i] &= \sum_{k \in S} E[\tau I_{X_1 = k} | X_0 = i] = \sum_{k \in S} \big[1 + E(\tau \mid X_0 = k)\big] p_{ik} \\ &= 1 + \sum_{k \in T} p_{ik}\, m_k, \end{aligned}$$

yielding (8). $\square$

In general: Mean Hitting Time (Adv)

Consider a finite Markov chain $\{X_n\}_{n\geq 0}$ with state space S. Let $A \subset S$ be a set of states. Let T_A be the first time the chain visits (hits) a state in A, i.e.,

$$T_A = \min\{n \geq 0 : X_n \in A\}.$$

For all $i \in S$, define the mean hitting time:

$$t_i(A) \;=\; E[T_A | X_0 = i].$$

By the definition, $t_i(A) = 0$ for all $i \in A$ and

$$t_i(A) = E[T_A | X_0 = i] = E[\widetilde{T}_A | X_0 = i]$$

for all $i \notin A$, where $\widetilde{T}_A = \min\{n \geq 1 : X_n \in A\}$.

Similarly to (8), the unknown values of $t_i(A)$'s follows from the equations:

$$t_i(A) \;= 1 + \sum_{k\in S} t_k(A) p_{ik} = 1 + \sum_{k\in S-A} t_k(A) p_{ik}, \tag{9}$$

for $i \in S - A$.

Mean Return (recurrence) Time (Adv)

Recall notation used in Section 9:

$$\mu_i = E[T_i|X_0 = i], \quad \text{where } T_i = \min\{n \geq 1 : X_n = i\},$$

μ_i is the mean return (recurrence) time to state i, i.e., the expected time (number of steps) needed until the chain returns to state i.

We may use (9) to calculate μ_i. Indeed, as in the proof of (9), we have

$$\mu_i = 1 + \sum_{k \in S-i} p_{ik}\, t_k(i), \tag{10}$$

where $t_k(i)$ is the expected time until the chain hits state i given $X_0 = k$, i.e., $t_k(i) = E(T_i \mid X_0 = k)$.

Note that $t_k(i)$ satisfies (9), i.e., $t_k(i) = 0$ for $k = i$ and

$$t_k(i) = 1 + \sum_{j \in S-i} p_{kj} t_j(i), \qquad \text{for } k \neq i.$$

Proof. By noting $E[T_i \mid X_1 = i] = 1$ and for $i \neq k$,

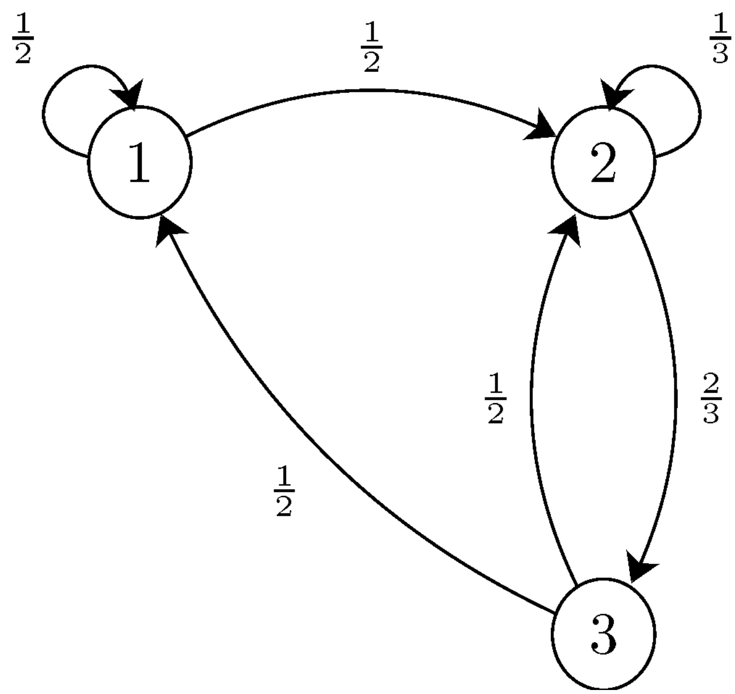
$$E[T_i \mid X_1 = k] = 1 + E[T_i \mid X_0 = k]$$

it follows that

$$\begin{aligned}
E[T_i|X_0 = i] &= \sum_{k \in S} E[T_i I_{X_1 = k}|X_0 = i] = \sum_{k \in S} E[T_i \mid X_1 = k] p_{ik} \\
&= p_{ii} + \sum_{k \in S-i} \big(1 + E[T_i \mid X_0 = k]\big) p_{ik} \\
&= \sum_{k \in S} p_{ik} + \sum_{k \in S-i} E(T_i \mid X_0 = k)\, p_{ik} \\
&= 1 + \sum_{k \in S-i} E(T_i \mid X_0 = k)\, p_{ik},
\end{aligned}$$

as required.

Example 4.6. Consider the Markov chain shown in the following graph:



Let t_k be the expected number of steps until the chain hits state 1 for the first time, given that $X_0 = k$. Clearly, $t_1 = 0$. Also, let μ_1 be the mean return time to state 1.

- Find t_2 and t_3.
- Find μ_1.

Solution: To find t_2 and t_3, it follows from (9) that

$$\begin{aligned} t_2 &= 1 + p_{22}t_2 + p_{23}t_3 = 1 + \frac{1}{3}t_2 + \frac{2}{3}t_3, \\ t_3 &= 1 + p_{32}t_2 + p_{33}t_3 = 1 + \frac{1}{2}t_2. \end{aligned}$$

Solving the equations, we obtain

$$t_2 = 5, \quad t_3 = \frac{7}{2}.$$

To find μ_1, it follows from (10) that

$$\begin{aligned}
\mu_1 &= 1 + p_{12}t_2 + p_{13}t_3 \\
&= 1 + \frac{1}{2} \cdot 5 + 0 \cdot 7/2 \\
&= \frac{7}{2}.
\end{aligned}$$

In comparison: Note that this is an irreducible and aperiodic MC. We can find the stationary dist $\pi = (\pi_1, \pi_2, \pi_3)$ satisfying $\pi = \pi P$, and then $\mu_1 = 1/\pi_1$.