STATS 3U03

Sang Woo Park

March 24, 2017

Course Outline

- Textbook: Inroduction to stochastic processes
- Requirement: 5 assignments, 2 tests, and 1 final
- Test 1: Friday, February 10th
- Test 2: Friday, March 17th

Contents

1 Introduction

1.1 Review

Definition 1.1 (Independent random variables)**.** *X and Y are independent iff for any $a, b \in \mathbb{R}$, $P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b)$*

1.2 Stochastic processes

Definition 1.2 (Stochastic process)**.** *Let T be a subset of $[0, +\infty]$. For each $t \in T$, let X_t be a random variable. Then, the collection of $\{X_t : t \in T\}$ is called a stochastic process. Simply put, a stochastic process is just a family of random variables.*

Example 1.2.1. Let $T = \{0\}$. Then, $\{X_0\}$ is a stochastic process.

Example 1.2.2. Let $T = \{1, 2, 3, \ldots, m\}$ be a set of finite natural numbers. Then, $\{X_1, X_2, X_3, \ldots, m\}$ is a stochastic process.

Example 1.2.3. Let $T = \{0, 1, 2, \ldots\}$ be a set of all non-negative integers. Then, $\{X_1, X_2, X_3, \ldots\}$ is a stochastic process.

Example 1.2.4. Let $T = [0, +\infty)$ be a set of all non-negative real numbers. Then, $\{X_t : t \geq 0\}$ is a stochastic process.

Definition 1.3 (Time index)**.** *Let T be time index. If $T = \{0, 1, 2, \ldots\}$, then the time is discrete. If $T = [0, \infty)$, then time is continuous.*

Definition 1.4 (State Space)**.** *State space, S, is the space space where the random variable takes the values.*

Given a sample space, S, and time index $t \in T$, we can define $X_t(w) \in S$, to describe a stochastic process. Here, $\{X_t : t \in T\}$ describes the dependence relation.

We can further categorize a stochastic process by considering the following two cases: countable and uncountable state space. Time index can also be categorized as follows: discrete and continuous time. Note that each stochastic process must belong to one of the four categories.

Remark. Every stochastic process can be described by the following three factors:

1. Time index
2. State space
3. Dependence relation

Example 1.2.5. Let $S = \{0, 1\}$ and $T = \{0, 1, 2, \ldots\}$. Given,

$$X_n = \begin{cases} 1 & \text{with probability of } 1/2 \\ 0 & \text{with probability of } 1/2 \end{cases}$$

$\{X_0, X_1, X_2, \ldots\}$ is a stochastic process and is often noted as Bernoulli trials.

2 Markov chains (Discrete time Markov chains)

We will only be dealing with discrete time Markov chains in chapter 1 and 2. In other words, $T = \{0, 1, 2, \dots\}$. It follows that the state space, S, will be at most countable. Finally, Markov describes the dependence relation: $X_0, X_1, X_2, \dots$.

In example 1.2.5, every trial of the Bernoulli trials was independent. On the other hand, in a Markov model, X_{n+1} depends on X_n but not on any past stats, $X_1, X_2, \dots, X_{n-1}$.

2.1 Markov property

Definition 2.1. Markov property *can be expressed as follows:*

$$\begin{aligned} &P\left(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = x_n\right) \\ &= P\left(X_{n+1} = x_{n+1} | X_n = x_n\right) \end{aligned}$$

$P\left(X_{n+1} = y | X_n = x\right)$ is noted as the transition probability and it describes the one step transition from x to y starting at time n. If

$$P\left(X_{n+1} = y | X_n = x\right) = P\left(X_1 = y | X_0 = x\right),$$

then the Markov chain is called to have stationary transition, or homogeneous.

Definition 2.2. *Let* $\{X_n : n = 0, 1, 2, \dots\}$ *be a homogeneous Markov chain. Then,*

$$P_{xy} = P\left(X_1 = y | X_0 = x\right) = P\left(X_{n+1} = y | X_n = x\right),$$

is the one-step transition probability*.*

Definition 2.3. *Following the definition 2.2, we can now define* one-step transition matrix*:*

$$\mathbb{P} = \left(P_{xy}\right)_{x, y \in S}$$

Remark. Given, X_0, $\pi_0(x) = P(X_0 = x)$ is called the *initial distribution*.

Given a Markov chain, we wish to answer the following fundamental questions:

1. Distribution of X_n for any $n \geq 1$.
2. Join distribution of $X_{n_1}, \dots, X_{x_k}$ for any $1 \leq n_1 < n_2 < \cdots < n_k \quad k, \geq 2$.
3. Long time behaviour, i.e.

$$\lim_{n \to \infty} P(X_n = x)$$

Example 2.1.1. We have the following Markov chain: $\{X_n : n = 0, 1, 2, \dots\}$ where $S = \{0, 1\}$. For this model, its initial distribution can be described as follows:

$$\begin{cases} \pi_0(0) = P(X_0 = 0) = a \\ \pi_0(1) = 1 - a \end{cases}$$

Transition probabilities can be written in a similar fashion:

$$
\begin{array}{ll}
P(X_1 = 1|X_0 = 0) = p, & P(X_1 = 0|X_0 = 0) = 1 - p \\
P(X_1 = 0|X_0 = 1) = q, & P(X_1 = 1|X_0 = 1) = 1 - q
\end{array}
$$

where $0 \leq p, q \leq 1$. For this Markov chain, we can consider the following three cases:

Case 1. $p = q = 0$**.**

This case is trivial.

Case 2. $p = q = 1$**.**

This case is also trivial.

Case 3. $0 \leq p + q \leq 2$**.**

$$
\begin{aligned}
P(X_{n+1} = 0) &= P(X_{n+1} = 0 \cap X_n = 0) + P(X_{n+1} = 0 \cap X_n = 1) \\
&= P(X_n = 0)P(X_{n+1} = 0|X_n = 0) + P(X_n = 1)P(X_{n+1} = 0|X_n = 1) \\
&= P(X_n = 0)(1 - p) + P(X_n = 1)q \\
&= P(X_n = 0)(1 - p) + (1 - P(X_n = 0))q \\
&= (1 - p - q)P(X_n = 0) + q
\end{aligned}
$$

We can further expand this as follows:

$$
\begin{aligned}
P(X_{n+1} = 0) &= (1 - p - q)P(X_n = 0) + q \\
&= (1 - p - q)\big[(1 - p - q)P(X_{n-1} = 0) + q\big] + q \\
&= (1 - p - q)^n P(X_0 = 0) + q\sum_{j=0}^{n-1}(1 - p - q)^j
\end{aligned}
$$

Note that

$$
\sum_{j=0}^{n-1}(1 - p - q)^j \quad = \frac{(1 - p - q)^n - 1}{(1 - p - q) - 1}
$$

Therefore, we have

$$
\begin{aligned}
P(X_{n+1} = 0) &= (1 - p - q)^n a + q\frac{(1 - p - q)^n - 1}{(1 - p - q) - 1} \\
&= (1 - p - q)^n a - \frac{q}{p + q}\big((1 - p - q)^n - 1\big)
\end{aligned}
$$

For this Markov chain, we find that

$$
\lim_{n \to \infty} P(X_n = 0) = \frac{q}{p + q}
$$

2.2 Transition function and initial distribution

Example 2.2.1.

$$\begin{aligned} P_{xy} &= P(X_{n+1} = y | X_n = x) \\ &= P(X_1 = y | X_0 = x) \end{aligned}$$

Definition 2.4. Transition function*,* $P(x, y) : S \times X \to [0, 1]$*, satisfies the following conditions:*

1. $P(x, y) \geq 0$.
2. $\sum_{y \in S} P(x, y) = 1$ *for all* $x \in S$.

Definition 2.5. *Given a transition function,* $P(x, y)$*, a transition matrix is defined as follows:*

$$\mathbb{P} = (P(x, y))_{x,y \in S}$$

Example 2.2.2.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Example 2.2.3.

$$\begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/8 & 1/4 & 5/8 \\ 0 & 1/4 & 3/4 \end{pmatrix}$$

Definition 2.6. Initial distribution *is a probability mass function (pmf) that is defined as follows*

$$\pi_0(x) = P(X_0 = x).$$

Note that it must satisfy the following conditions:

1. $\pi_0(x) \geq 0$
2. $\sum_{x \in S} \pi_0(x) = 1$

Theorem 2.1. *Let* $\{x_N : n = 0, 1, 2, \dots\}$ *be a Markov chain with initial distribution* $\pi_0(x)$*, and one-step transition matrix* $\mathbb{P} = (P(x, y))_{x,y \in S}$*. Then, the distribution of* X_n *is*

$$\begin{aligned} P(X_n = x_n) &= \sum_{x_0 \in S} \sum_{x_1 \in S} \cdots \sum_{x_{n-1} \in S} \pi_0(x_0) P(x_0, x_1) \dots P(x_{n-1}, x_n) \\ &= \pi_0 \underbrace{\mathbb{P}\mathbb{P} \dots \mathbb{P}}_{n} \end{aligned}$$

Proof. For any $n \geq 1, \quad x_n \in S$

$$\begin{aligned} P(X_n = x_n) &= P(X_n = x_n, x_0 \in S, X_1 \in S, \dots, X_{n-1} \in S) \\ &= \sum_{x_0 \in S} \sum_{x_1 \in S} \cdots \sum_{x_{n-1} \in S} P(X_n = x_n, X_0 = x_0, \dots, X_{n-1} = x_{n-1}) \end{aligned}$$

Note that

$$\begin{aligned}
&P(X_n = x_n, X_0 = x_0, \dots, X_{n-1} = x_{n-1}) \\
&= P(X_0 = x_0)P(X_1 = x_1|X_0 = x_0)P(X_2 = x_2|X_0 = x_0, X_1 = x_1) \\
&\cdots P(X_n = x_n|X_0 = x_0, \dots, X_{n-1} = x_{n-1})
\end{aligned}$$

Using the Markov property, it is evident that the equation above is equivalent to $P(X_0 = x_0)P(X_1 = x_1|X_0 = x_0)\cdots P(X_n = x_n|X_{n-1} = x_{n-1})$. $\square$

Example 2.2.4. Simple random walk is a Markov chain:

$$\begin{aligned}
X_0 &= 0 \\
X_1 &= \begin{cases} 1 & p \\ -1 & q \end{cases} \\
X_2 &= \begin{cases} X_1 + 1 & p \\ X_1 - 1 & q \end{cases} \\
X_n &= \begin{cases} X_{n-1} + 1 & p \\ X_{n-1} - 1 & q \end{cases}
\end{aligned}$$

where $S = \{0, \pm 1, \pm 2, \dots\}$.

Example 2.2.5 (Ehrenfest chain)**.** Suppose that we have a box and a inivisble bar that divides the box into region I and II. d balls are placed in a box. Initially, n balls are distributed in region I and $d - n$ balls are distributed in region II. You pick a ball at random. If it's from region I, you put it back in region II. If it's from region II, you put it back in region I.

First, note that this Markov chain has a state space of $S = \{0, 1, 2, \dots, d\}$. We observe that

$$\begin{aligned}
P(0, y) &= \begin{cases} 0 & y > 1 \\ 1 & y = 1 \end{cases} \\
P(1, y) &= \begin{cases} 0 & y \neq 0, 2 \\ \frac{1}{d} & y = 0 \\ 1 - \frac{1}{d} & y = 2 \end{cases}
\end{aligned}$$

In general, we have

$$P(x, y) = \begin{cases} 0 & y \neq x \pm 1 \\ 1 - \frac{x}{d} & y = x + 1 \\ \frac{x}{d} & y = x - 1 \end{cases}$$

Combining these results, we have the following transition matrix:

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & \\ \frac{1}{d} & 0 & 1 - \frac{1}{d} & & \\ & & & \ddots & \\ & & & & \end{pmatrix}_{(d+1)\times(d+1)}$$

Example 2.2.6 (Birth-Death Markov chain)**.** At each time step, one person can die and a new person can be born:

$$X_{n+1} = \begin{cases} p_x & y = X_n + 1 \\ q_x & y = X_n - 1 \\ r_x & y = X_n \\ 0 & \text{else} \end{cases}$$

Example 2.2.7 (Queuing chain)**.** At each time step, one customer is served and new customers arrive:

$$X_{n+1} = \begin{cases} y_{n+1} & \text{if } X_n = 0 \\ X_n - 1 + y_{n+1} & \text{if } X_n \geq 1 \end{cases}$$

We introduce a new notation, $x^+ = x \vee 0$, which is essentially $\max(x, 0)$. Using this notation, we can rewrite the Markov chain as follows

$$X_{n+1} = (X_n - 1)^+ + y_n$$

Example 2.2.8 (Branching Markov chain)**.** If $X_0 = 0$, then $X_n = 0$ for all $n \geq 1$. We call 0 an *absorbing state.*

Suppose $X_0 \geq 1$. An individual, i, will produce y_i number of offsprings at each generation. Then, we will have

$$X_1 = y_1^{(1)} + \cdots + y_{X_0}^{(1)}$$

Each individual in generation 1 will also produce offsprings. Then,

$$X_2 = y_1^{(2)} + \cdots + y_{X_0}^{(2)}$$

We wish to understand how the population will evolve over time. To do so, we can look at the expected value. It's clear that the population will grow if $E[y] > 1$. On the other hand, if $E[y] < 1$, the population will eventually die out.

Example 2.2.9 (Wright-Fisher Markov chain)**.** For this Markov chain, we start by make the following assumptions:

1. The population size is fixed.
2. No generation overlap.

Within the population, there are N number of individuals of two types: I and II. Let X_0 be number of type I individuals at time 0. Each individual in generation 1 pick its parent from generation 0 at random. This process is equivalent to repeating Bernoulli trials N times (also equivalent to binomial).

Therefore, we have

$$\begin{aligned}
X_1 &\sim \text{Bin}(N, \frac{X_0}{N}) \\
X_2 &\sim \text{Bin}(N, \frac{X_1}{N}) \\
&\vdots \\
X_{n+1} &\sim \text{Bin}(N, \frac{X_n}{N})
\end{aligned}$$

2.3 Joint distribution

Given a Markov chain with π_0 and $\mathbb{P}$, how do we find (1) the distribution of X_n and (2) the joint distribution of X_n and X_m where $n < m$?

From the previous section, recall that $\pi_n = \pi_0 \underbrace{\mathbb{P}\mathbb{P}\dots\mathbb{P}}_{n}$.

Example 2.3.1. Consider the following transition matrix:

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

Supose that $\pi_0 = (1, 0, 0, 0)$. Then, we have

$$\begin{aligned}
\pi_1 &= (1, 0, 0, 0)\mathbb{P} = (1/2, 1/2, 0, 0) \\
\pi_2 &= (1, 0, 0, 0)\mathbb{P}\mathbb{P} = (1/4, 1/2, 1/4, 0)
\end{aligned}$$

Eventually, all states will converge to the absorbing state and stay there.

To find the join distribution, we first note that

$$\begin{aligned}
P(X_n = x, X_m = x_m) &= P(X_n = x_n)P(X_m = x_m | X_n = x_n) \\
&= P(X_m = x_m)P(X_{m-n} = x_m | X_0 = x_n)
\end{aligned}$$

Definition 2.7. *For any interger m, m-step transition matrix is given by*

$$P^m(x, y) = P(X_m = y | X_0 = x).$$

When $m = 0$, we have

$$P^0(x, y) = \begin{cases} 1 & y = x \\ 0 & y \neq x \end{cases}$$

We can decompose m-step transition matrix as follows:

$$\begin{aligned}
P^m(x, y) &= P(X_m = y | X_0 = x) \\
&= P(X_m = y, x_1 \in S, \dots, X_{m-1} \in S | X_0 = x) \\
&= \sum_{x_1 \in S} \sum_{x_2 \in S} \cdots \sum_{x_{m-1} \in S} P(x_0, x_1) \dots P(x_{n-1}, y)
\end{aligned}$$

Then, we have

$$\begin{aligned}
P(X_m = y) &= P(X_m = y, X_m \in S) \\
&= \sum_{x_0 \in S} P(X_m = y, X_0 = x_0) \\
&= \sum_{x_0 \in S} P(X_0 = x_0) P^m(X_0, y) \\
&= \sum_{x_0 \in S} \pi_0(x_0) P^m(X_0, y)
\end{aligned}$$

Therefore, we have

$$(P(X_m = x_m))_{x_m \in S} = \pi_0 \mathbb{P}^m$$

Definition 2.8 (Hitting time)**.** *Given $A \subset S$, hitting time T_A is defined as follows:*

$$T_a = \min\{n \geq 1 : X_n \in A\}$$

If $A = \{x\}$, then we have $T_x = T_x$. Note that

- $T_A \geq 1$
- *If $x_n \notin A$ for all $n \geq 1$. we have $T_A = +\infty$*

Now, we wish to understand the distribution of T_y given thatn $X_0 = x$. First, note that we have

$$P_x(T_y = 1) = P(T_y = 1 | X_0 = x) = P(x, y)$$

Similarly, we have

$$\begin{aligned}
P_x(T_y = 2) &= P_x(x_1 \neq y, x_2 = y) \\
&= \sum_{w \neq y} P(x, w) P(w, y)
\end{aligned}$$

Generally, we have

$$\begin{aligned}
P_x(T_y = n+1) &= P_x(x_1 \neq y, \ldots, x_n \neq y, x_{n+1} = y) \\
&= \frac{P(X_0 = x, X_1 \neq y \ldots X_n \neq y, X_{n+1} = y)}{P(X_0 = x)} \\
&= \frac{P(X_0 = x, X_1 \neq y)}{P(X_0 = x)} \frac{P(X_0 = x, X_1 \neq y \ldots X_n \neq y, X_{n+1} = y)}{P(X = x, X_1 \neq y)} \\
&= \sum_{x_1 \neq y} P(x, x_1) P_{x_1}(T_y = n)
\end{aligned}$$

Note that the last result follows from the Markov property.

Lemma 2.1. $P^m(x, y) = \sum\limits_{k=1}^{m} P_x(T_y = k) P^{m-k}(y, y)$

Proof.

$$\begin{aligned}
P^m(x,y) &= P(X_m = y|X_0 = x) \\
&= P(X_m = y, T_y \leq m|x_0 = x) \\
&= \sum_{k=1}^{m} P(X_m = y, T_y = k|X_0 = x) \\
&= \sum_{k=1}^{m} \frac{P(X_0 = x, T_y = k, X_m = y)}{P(X_0 = x)} \\
&= \sum_{k=1}^{m} \frac{P(X_0 = x, T_y = k)}{P(X_0 = x)} \frac{P(X_0 = x, T_y = k, X_m = y)}{P(X_0 = x, T_y = k)} \\
&= \sum_{k=1}^{m} P_x(T_y = k) P(X_m = y|X_0 = x, x \neq y, x_k = y) \\
&= \sum_{k=1}^{m} P_x(T_y = k) P(X_m = y|x_k = y)
\end{aligned}$$

□

2.4 Recurrence

Before we define recurrent and transient states, we introduce the following notation:

$$\rho_{xy} = P_x(T_y \leq \infty) = \sum_k P_x(T_y = k).$$

Definition 2.9 (Recurrent and Transient states)**.** *A state x is called* recurrent *if $\rho_{xx} = 1$. Ohterwise, it is called* transient.

We introduce more notations:

- $I_x(y) = \begin{cases} 1 & y = x \\ 0 & \text{else} \end{cases}$ (indicator function of x).
- $N(y) = \sum_{n=1}^{\infty} I_y(X_n)$

Theorem 2.2.

1. $P_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1}$
2. $P_x(N(y) = m) = \rho_{xy}\rho_{yy}^{m-1}(1 - \rho_{yy})$
3. $P_x(N(y) = 0) = 1 - \rho_{xy}$

Proof. First, assume that theorem 1 is true. Then, we have

$$\begin{aligned} P_x(N(y)=m) &= P_x(N(y) \geq m) - P_x(N(y) \geq m+1) \\ &= \rho_{xy}\rho_{yy}^{m-1} - \rho_{xy}\rho_{yy}^{m} \\ &= \rho_{xy}\rho_{yy}^{m-1}(1-\rho_{yy}) \end{aligned}$$

Now, we want to prove theorem 3:

$$\begin{aligned} P_x(N(y)=0) &= 1 - P_x(N(y) \geq 1) \\ &= 1 - \rho_{xy} \end{aligned}$$

Finally, we just have to prove theorem 1:

$$\begin{aligned} P_x(N(y) \geq m) &= P_x(\text{The Markov chain visits state } y \text{ at least } m \text{ times}) \\ &= \sum_{n_1 \geq 1} \cdots \sum_{n_m \geq 1} P_x(T_y = n_1)P_y(T_y = n_2)\cdots P_y(T_y = n_y) \\ &= \sum_{n_1 \geq 1} P_x(T_y = n_1) \sum_{n_2 \geq 1} P_y(T_y = n_2) \cdots \sum_{n_m \geq 1} P_y(T_y n_m) \\ &= P_x(T_y < \infty)P_y(T_y < \infty)\cdots P_y(T_y < \infty) \\ &= \rho_{xy}\rho_{yy}^{m-1} \end{aligned}$$

□

Before looking at the next theorem, we introduce another notation: $E_x[\cdot]$ is the expectation given the initial state of x. Then, we have

$$\begin{aligned} E_x[I_y(X_n)] &= P_x(I_y(X_n) = 1) \\ &= P_x(X_n = y) \\ &= P^n(x,y) \end{aligned}$$

Furthermore, we introduce the notation, G:

$$\begin{aligned} G(x,y) &= E_x[N(y)] \\ &= E_x\left[\sum_y I_y(x_n)\right] \\ &= \sum_y E_x[I_y(x_n)] \\ &= \sum_y P^n(x,y) \end{aligned}$$

Theorem 2.3.

1. *If y is transient, then for any $x \in S$, $P_x(N(y) < \infty) = 1$ and $G(x,y) = \frac{P_{xy}}{1-P_{yy}} < \infty$.*

2. *If y is recurrent, then for any $x \in S$, $P_x(N(y) = \infty) = 1$ and $G(y, y) = \infty$. Furthermore, we have $P_x(N(y) = \infty) = \rho_{xy}$ and*

$$G(x,y) = \begin{cases} \infty & \text{if } \rho_{xy} > 0 \\ 0 & \text{if } \rho_{xy} = 0 \end{cases}$$

Proof. Suppose y is transient. Then, we have $\rho < 1$. For any x, we have

$$\begin{aligned} P_x(N(y) = \infty) &= P_x\left(\bigcap_{m=1}^{\infty} \{N(y) \geq m\}\right) \\ &= \lim_{m\to\infty} P_x(N(y) \geq m) \\ &= \lim_{m\to\infty} \rho_{xy}\rho_{yy}^{m-1} \\ &= 0 \end{aligned}$$

Therefore, we have

$$P_x(N(y) < \infty) = 1 - P_x(N(y) = \infty) = 1 - 0 = 1$$

Furthermore,

$$\begin{aligned} G(x,y) &= \sum_{m=1}^{\infty} m P_x(N(y) = m) \\ &= \sum_{m=1}^{\infty} m\rho_{xy}\rho_{yy}^{m-1}(1 - \rho_{yy}) \\ &= \rho_{xy}(1 - \rho_{yy}) \sum_{m=1}^{\infty} m\rho_{yy}^{m-1} \\ &= \rho_{xy}(1 - \rho_{yy}) \left(\sum_{m=1}^{\infty} \rho_{yy}^{m}\right)' \\ &= \rho_{xy}(1 - \rho_{yy}) \frac{1}{(1 - \rho_{yy})^2} \end{aligned}$$

Let's prove the second statement. If y is recurrent, then $\rho_{yy} = 1$. For any x, we have

$$\begin{aligned} P_x(N(y) = \infty) &= P_x(\bigcap \{N(y) \geq m\}) \\ &= \lim_{m\to\infty} P_x(N(y) \geq m) \\ &= \lim_{m\to\infty} \rho_{xy}\rho_{yy}^{m-1} \\ &= \rho_{xy} \end{aligned}$$

Then, we have

$$
\begin{aligned}
G(x,y) &= \sum m\rho_{xy}\rho_{yy}^{m-1} \\
&= \sum m\rho_{xy} \\
&= \rho_{xy}\sum m \\
&= \infty
\end{aligned}
$$

□

Example 2.4.1. Let y be a transient state. Find

$$\lim_{n\to\infty} P^n(x,y).$$

Recall that $G(x,y) = \sum_{n=1}^{\infty} P^n(x,y) = \frac{\rho_{xy}}{1-\rho_{yy}} < \infty$. Since the series converges, it is easy to see that $\lim\limits_{n\to\infty} P^n(x,y) = 0$

Example 2.4.2. Let $\{X_n, n \geq 0\}$ be a two state Markov $S = \{0,1\}$. Can both be transient?

We start by noting that $P_x(X_n \in S) = 1$. If both are transient, we have

$$\lim_{n\to\infty} P_x(X_n \in S) = \lim_{n\to\infty} P^n(x,0) + P^n(x,1) = 0,$$

yielding a contradiction.

Definition 2.10. *A Markov chain is* recurrent *if all states are recurrent, and the chain is* transient *of all states are transient.*

Definition 2.11. *A state x leads to state y of $\rho_{xy} > 0$ denoted $x \to y$.*

Remark. It is possible that $x \not\to x$.

Lemma 2.2.

1. *$x \to y$ iff there exists $n \geq 1$ such that $P^n(x,y) > 0$.*
2. *If $x \to y$ and $y \to z$, then $x \to z$.*

Proof. By definition, $x \to y$ ifff $\rho_{xy} > 0$. In other words, $P_x(T_y < \infty) > 0$. Then,

$$0 < P_x(T_y < \infty) = \sum_{n=1}^{\infty} P_x(T_y = n) \leq \sum_{n=1}^{\infty} P^n(x,y)$$

Therefore, $P^n(x,y) > 0$ for some n. Conversely, if $P^n(x,y) > 0$ for some $n \geq 1$, we can define

$$n_0 = \min\{n \geq 1, P^n(x,y)\}.$$

Clearly, $0 < P^n(x,y) \leq= P_x(T_y = n_0) < \rho_{xy}$.

To prove the second statement, note that $x \to y$ iff $\exists n_1 \geq 1$ such that $P^{n_1}(x,y) \geq 0$. Similarly, $y \to z$ iff $\exists n_2 \geq 1$ such that $P^{n_2}(y,z) \geq 0$. Then,

$$P^{n_1+n_2}(x,z) \geq P^{n_1}(x,y)P^{n_2}(y,z) > 0$$

Therefore, $x \to z$ □

Theorem 2.4. *If x is recurrent and $x \to y$, then y is recurrent and $\rho_{xy} = \rho_{yx} = 1$.*

Proof. To yield contradiction, suppose $\rho_{yx} \neq 1$. Then,

$$1 - \rho_y > 0.$$

Furthermore, if $x \to y$, there exists n_1 such that $P^{n_1}(x,y) > 0$. This implies that

$$P^{n_1}(x,y)(1 - \rho_{yx}) > 0$$

The first part is the probability that x reaches y in n_1 steps. However, the second part says that y never goes to x, contradicting the assumption that x is recurrent. Therefore,

$$\rho_{yx} = 1.$$

To prove that y is recurrent, we first note that if $x \to y$, there exists n_1 such that $P^{n_1}(x,y) > 0$. Similarly, if $\rho_{yx} = 1$, $y \to 1$ and there exists n_2 such that $P^{n_1}(y,x) > 0$. Then,

$$\begin{aligned}
G(y,y) &= \sum_{n=1}^{\infty} P^n(y,y) \\
&\geq \sum_{m=1}^{\infty} P^{n_1+n_2+m}(y,y) \\
&\geq \sum_{m=1}^{\infty} P^{n_2}(y,x) P^m(x,x) P^{x,y} \\
&= P^{n_2}(y,x) \underbrace{\left(\sum_{n=1}^{\infty} P^m(x,x) \right)}_{G(x,x)=\infty} P^{n_1}(x,y)
\end{aligned}$$

Finally, to prove that $\rho_{xy} = 1$, we note that y is recurrent. Then, by following the proof of the first statement, we can prove that $\rho_{xy} = 1$. $\square$

Definition 2.12. *If $x \to y$ and $y \to x$, we write*

$$x \leftrightarrow y$$

and say that x communicates *with y*

Definition 2.13. *A subset C is* closed *if for any $x \in C$ and $y \notin C$, $x \not\to y$ ($\rho_{xy} \leq 0$).*

Definition 2.14. *A closed subset C is* irreducible *if every $x, y \in C$ communicate with each other.*

We can further define *closed and irreducible set* where (1) $x, y \in C$, $x \leftrightarrow y$, and (2) $x \in C$, $z \notin C$, $\rho_{xz} = 0$. *Closed, irreducible, and recurrent set* is then defined as (1) $x, y \in C$, $x \leftrightarrow y$, $\rho_{xy} = \rho_{yx} = 1$, and (2) $x \in C, z \notin C, \rho_{xz} = 0$.

Then, we can decompose a state space, S, into a set of recurrent and transient state:

$$S = C_R \cup C_T$$

Theorem 2.5. *If for $x, y \in C_R$, $C_x \cap C_y \neq \varnothing$. Then, $C_x = C_y$.*

Proof. Let $w \in C_x \cap C_y$. Then, $w \leftrightarrow x$ and $w \leftrightarrow y$. For any $z \in C_x$, we have

$$z \leftrightarrow x \leftrightarrow w \leftrightarrow y$$

and

$$z \in C_y,$$

implying that $C_x \subset C_y$. By symmetry, $C_y \subset C_x$. Therefore, $C_x = C_y$. $\square$

Theorem 2.6. *The state space S of a Markov chain can be decomposed as two union of C_R and C_T. Furthermore, C_R can be decomposed into the union of at most countable number of closed, irreducible, recurrent sets.*

Note that you have to stay in a recurrent set if you start from a recurrenut set. On the other hand, if you start from a transient set, you have to move to a recurrent state if the set contains finite elements. If the set contains infinite number of elements, it is possible to stay in the transient set forever.

Example 2.4.3. Let $\{X_n, n = 0, 1, 2, \dots\}$ be a Markov Chain with $S = \{0, 1, 2, 3, 4, 5\}$ and the following one step transient matrix:

$$\mathbb{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 1/5 & 2/5 & 1/5 & 0 & 1/5 \\ 0 & 0 & 0 & 1/6 & 1/3 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 1/3 \\ 0 & 0 & 0 & 1/4 & 0 & 3/4 \end{pmatrix}$$

(a) Find C_R and C_T

Note that 0 is an absorbing state. If you start from state 1 or 2. you have a positive probability of going to state 0. Therefore, state 1 and 2 are transient. On the other hand, we have $3 \to 4 \to 5 \to 3$, implying that

$$3 \leftrightarrow 4 \leftrightarrow 5.$$

Then, $\{3, 4, 5\}$ form a closed, irreducible, and recurrent state. Therefore,

$$C_R = \{0, 3, 4, 5\}$$
$$C_T = \{1, 2\}$$

(b) Decompose C_R

Clearly, $C_R = \{0\} \cup \{3, 4, 5\}$ and two subsets are irreducible.

Remark. All closed, irreducible, finite set are recurrent set.

Example 2.4.4. Let $S = \{0, 1, 2, 3, \dots\}$. Given the following transition matrix,

$$\mathbb{R} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ \vdots & & & & \end{pmatrix}$$

Then, each state is an absorbing state and we have

$$C_R = \bigcap_{i=0}^{\infty} \{i\}.$$

Example 2.4.5. Consider the following transtion matrix:

$$\mathbb{R} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Then, since $1 \to 2 \to 3$, we have

$$C_R = \{1, 2, 3\}$$

Example 2.4.6. Consider the following transition matrix:

$$\mathbb{R} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Then, we have $C_T = \{1, 2\}$ and $C_R = \{3\}$.

Example 2.4.7. Consider the following transition matrix:

$$\mathbb{R} = \begin{pmatrix} 0 & 1 & 0 \\ a & 0 & 1-a \\ 0 & 0 & 1 \end{pmatrix}$$

For all $0 \leq a < 1$, decomposition of the state space does not change. Higher a only implies that it will take longer to get to the absorbing state.

2.5 Absorption probabilities

Definition 2.15 (Absorption probabilities)**.** *Let C be a recurrent, irreducible, closed set. For $x \in C_T$, probability of x being abosrbed by C is given by*

$$\rho_C(x) = \rho_x(T_c < \infty)$$

To calculate the absorption state, we must solve

$$\rho_C(x) = \sum_{y \in C} P(x, y) + \sum_{y \in C_T} P(x, y)\rho_C(y).$$

This is in fact a system of linear equations. We are interested the uniqueness of the solution.

Theorem 2.7. *If C_T is finite, then the system*

$$w_x = \sum_{y \in C} P(x, y) + \sum_{y \in C_T} w_y$$

has a unique solution $w_x = \rho_C(x)$.

Proof. Let $\{w_x : x \in C_T\}$ be any solution. Then,

$$\begin{aligned}
w_x &= \sum_{y \in C} P(x, y) + \sum_{y \in C_T} P(x, y)w_y \\
&= \sum_{y \in C} P(x, y) + \sum_{y \in C_T} P(x, y) \left[\sum_{z \in C} P(y, z) + \sum_{y \in C_T} P(y, z)w_z \right] \\
&= \sum_{y \in C} P(x, y) + \sum_{y \in C_T} \sum_{z \in C_T} P(x, y)P(y, z)w_z + \sum_{y \in C_T} \sum_{z \in C} P(x, y)P(y, z) \\
&= \sum_{y \in C} P(x, y) + \sum_{z \in C_T} P^2(x, z)w_z + \sum_{y \in C_T} \sum_{z \in C} P(x, y)P(y, zsss) \\
&= P_x(T_C \leq 2) + \sum_{z \in C_T} P^2(x, z)w_z \\
&= \cdots \\
&= P_x(T_c \leq n) + \sum_{z \in C_T} P^2(x, z)w_z
\end{aligned}$$

Now, we can take the limit as n goes to infinity:

$$\begin{aligned}
w_x &= \lim_{n \to \infty} \left(P_x(T_c \leq n) + \sum_{z \in C_T} P^2(x, z)w_z \right) \\
&= P_x(T_c \leq \infty) + \sum_{z \in C_T} \lim_{n \to \infty} P^n(x, z)w_z
\end{aligned}$$

Since C_T is finite, $\lim_{n \to \infty} P^n(x, z) = 0$, and therefore, $w_x = P_x(T_c \leq \infty)$. $\square$

Example 2.5.1. Let $\{X_n, n = 0, 1, 2, \dots\}$ be a Markov chain with $S = \{1, 2, 3, 4\}$ and

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Let $C = \{1\}$. Find $\rho_C(2), \rho_C(3)$

First, note that we can decompose the set as follows:

$$C_R = \{1, 4\}, C_T = \{2, 3\}$$

Since C_T is finite, we have

$$w_x = \sum_{y \in C} P(x, y) + \sum_{y \in C_T} P(x, y) w_y$$

Then, we have

$$\begin{aligned} w_x &= P(3, 1) + P(2, 2) w_2 + P(2, 3) w_3 \\ &= \frac{1}{4} + \frac{3}{4} w_3 \end{aligned}$$

Similarly, we have

$$w_3 = \frac{1}{3} w_2 + \frac{1}{3} w_3$$

Therefore, we have

$$w_3 = \frac{1}{5}, w_2 = \frac{2}{5}$$

Theorem 2.8. *If for any $x, y \in S$, $x \leftrightarrow y$, then the chain is irreducible. Then, it follows that a finite state of irreducible Markov Chain is recurrent.*

Remark. Infinite, irreducible Markov chain can be transient. Irreducibility doesn't imply recurrence.

Then, when will an inifinite state, irreducible Markov chain be reccurent? We look at the birth-death Markov chain to understand this idea.

2.6 Birth-Death Markov Chain

Definition 2.16. *A Markov Chain $\{X_n, n = 1, 2, \dots\}$ is called a birth-death Markov chain if*

1. $S = \{0, 1, 2, \dots, d\}$ *where d can be either finite or infinite. When $d = \infty$,* $S = \{0, 1, 2, \dots\}$.
2. $P(x, y) = \begin{cases} p_x & y = x + 1 \\ q_x & y = x - 1 \\ r_x & y = x \\ 0 & \textit{else} \end{cases}$.

Note that if $p_x > 0, q_x > 0$, for $1 \leq x \leq d-q$ and $p_0 > 0, q_d > 0$, then the chain irreducible. If the chain is irreducible and $d < \infty$, then the birth-death chain is recurrent.

Theorem 2.9. *For any $a, b \in S$ and $a < b$. Let $u(x) = P_x(T_a < T_b)$ for $a \leq x \leq b$, $u(a) = 1, u(b) = 0$. Also, define*

$$\Gamma_0 = 1, \Gamma_k \frac{q_1 q_2 \cdots q_k}{p_1 p_2 \cdots p_k}, k \geq 1$$

Then,

$$u(x) = \frac{\sum_{r=x}^{b-1} \Gamma_r}{\sum_{r=1}^{b-1} \Gamma_r}$$

Proof. First, note that

$$\begin{aligned}
u(x) =& P_x(T_a < T_b) \\
=& P_x(x_1 = x \text{ or } x+1 \text{ or } x_1 = x-1, T_a < T_b) \\
=& P_x(x_1 = x, T_a < T_b) + P_x(x_1 = x+1, T_a < T_b) \\
&+ P_x(x_1 = x_2, T_a < T_b) \\
=& P_x(x_1 = x) P_x(T_a < T_b) + P_x(x_1 = x+1) P_{x+1}(T_a < T_b) \\
&+ P_x(x_1 = x-1) P_{x-1} P(T_a < T_b) \\
=& r_x u(x) + p_x u(x+1) + q_x u(x-1)
\end{aligned}$$

Rearranging, we get

$$\begin{aligned}
(1 - r_x) u(x) &= p_x u(x+1) + q_x u(x-1) \\
(p_x + q_x) u(x) &= p_x u(x+1) + q_x u(x-1) \\
p_x(u(x+1) - u(x)) &= q_x(u(x) - u(x-1))
\end{aligned}$$

Now, we can use this formula recursively:

$$\begin{aligned}
u(x+1) - u(x) &= \frac{q_x}{p_x}(u(x) - u(x-1)) \\
&= \frac{q_x}{p_x} \frac{q_{x-1}}{p_{x-1}}(u(x-1) - u(x-2)) \\
&= \frac{q_x}{p_x} \cdots \frac{q_{a+1}}{p_{a+1}}(u(a+1) - u(a)) \\
&= \frac{\frac{q_1 \cdots q_a}{p_1 \cdots p_a}}{\frac{q_1 \cdots q_a}{p_1 \cdots p_a}} \frac{q_x}{p_x} \cdots \frac{q_{a+1}}{p_{a+1}}(u(a+1) - u(a)) \\
&= \frac{\Gamma_x}{\Gamma_a}(u(a+1) - u(a)).
\end{aligned}$$

By definition, we know that $u(b) = 0$ and $u(a) = 1$. It is then trivial that $u(b) - u(a) = -1$. Finally, we can apply *telescoping* to achieve the desired result:

$$\begin{aligned}
-1 &= u(b) - u(b-1) + u(b-1) - u(b-2) + \cdots + u(a+1) - u(a) \\
&= \frac{\Gamma_{b-1}}{\Gamma_a}(u(a+1) - u(a)) + \frac{\Gamma_{b-2}}{\Gamma_a}(u(a+1) - u(a)) \\
&\quad + \cdots + (u(a+1) - u(a))
\end{aligned}$$

Thus, we have

$$u(a) - u(a+1) = \frac{\Gamma_a}{\sum_{r=1}^{b-1} \Gamma_r}$$

If we put everything together, we have

$$\begin{aligned} u(x) - u(x+1) &= \frac{\Gamma_x}{\Gamma_a}(u(a+1) - u(a)) \\ &= \frac{\Gamma_x}{\sum_{r=1}^{b-1} \Gamma_r} \end{aligned}$$

for all $a < x < b$.

Finally, since $u(x) = u(x) - u(b)$, we can apply telescoping again:

$$\begin{aligned} u(x) &= u(x) - u(x+1) + u(x+1) - u(x+1) + \cdots + u(b-1) - u(b) \\ &= \frac{\sum_{r=x}^{b-1} \Gamma_r}{\sum_{r=a}^{b-1} \Gamma_r} \end{aligned}$$

We have now derived a major result for the birth and death Markov chain. □

Lemma 2.3. $\rho_{00} = P(0,0) + P(0,1)\rho_{10}$.

Proof.

$$\begin{aligned} \rho_{00} &= P_0(T_0 < \infty) \\ &= P_0(X_1 = 0, T_0 < \infty) + P_0(X_1 = 1, T_0, \infty) \\ &= P_0(X_1 = 0) + P(0,1)P_1(T_0 < \infty) \\ &= P(0,0) + P(0,1)\rho_{10} \end{aligned}$$

□

Theorem 2.10. *The birth and death Markov chain is recurrent iff* $\sum\limits_{r=0}^{\infty} \Gamma_r = \infty$.

Proof. Let $a = 0$, $b = n$, and $x = 1$. Observe that

$$u(1) = P_1(T_0 < T_n) = \frac{\sum_{r=1}^{n-1} \Gamma_r + \Gamma_a - \Gamma_a}{\sum_{r=0}^{n-1} \Gamma_r} = 1 - \frac{1}{\sum_{r=0}^{n-1} \Gamma_r}.$$

Then, since

$$\rho_{10} = P_1(T_0 < \infty) = \lim_{n \to \infty} P_1(T_0 < n) = \lim_{n \to \infty} \left(1 - \frac{1}{\sum_{r=0}^{n-1} \Gamma_r}\right).$$

Clearly, $\rho_{10} = 1$ iff $\sum\limits_{r=0}^{\infty} \Gamma_r = \infty$. When $\rho_{10} = 1$, we have $\rho_{00} = P(0,0) + P(0,1) = 1$ and 0 becomes a recurrent state. Since the chain is irreducible, it is reccurent. □

Example 2.6.1. Consider a birth-death Markov Chain whose state is a set of all non-negative integers. For each state, it has a probability of going up of 0.51 and probability of going down of 0.49. Then,

$$\begin{aligned}\Gamma_r &= \frac{q_1 \cdots q_r}{p_1 \cdots p_r} \\ &= \left(\frac{0.49}{0.51}\right)^r\end{aligned}$$

Clearly, $\sum\limits_{k=0}^{\infty} \Gamma_k$ is a converging geometric series. Therefore, this is a transient Markov Chain.

Example 2.6.2. Cnosider the following chain:

$$P(x,y) = \begin{cases} p_0 & x=0, y=x+1 \\ 0 & x=0, y=x-1 \\ r_0 & x=0, y=0 \\ p_x & x \geq 1, y=x+1 \\ r_x & x \geq 1, y=x \\ q_x & x \geq 1, y=x-1 \end{cases}$$

We may define

$$p_x = \frac{x+2}{2(x+1)}, q_x = \frac{x}{2(x+1)}.$$

Then, it follows that $p_x + q_x = 1$ and $r_x = 0$.

We wish to know if this Chain is transient or not. First, observe that

$$\Gamma_1 = \frac{q_1}{p_1} = \frac{1}{3}$$

In general, we have

$$\begin{aligned}\Gamma_x = \frac{q_1 \cdots q_x}{p_1 \cdots p_x} &= \frac{\frac{1}{2(1+1)} \frac{2}{2(2+1)} \cdots \frac{x}{2(x+1)}}{\frac{1+2}{2(1+1)} \frac{2+2}{2(2+1)} \cdots \frac{x+2}{2(x+1)}} \\ &= \frac{1 \cdot 2 \cdots x}{(1+2)(2+2) \cdots (x+2)} \\ &= \frac{1 \cdot 2}{(x+1)(x+2)}\end{aligned}$$

Then, we see that

$$\begin{aligned}
\sum_{x=0}^{\infty}\Gamma_x &= \Gamma_0+\Gamma_1+\sum_{x=2}^{\infty}\Gamma_x \\
&= 1+\frac{1}{3}+2\sum_{x=2}^{\infty}\left(\frac{1}{(x+1)(x+2)}\right) \\
&= 1+\frac{1}{3}+2\left(\frac{1}{x+1}-\frac{1}{x+2}\right) \\
&= 1+\frac{1}{3}+\frac{2}{3}=2<\infty
\end{aligned}$$

Therefore, this chain is transient.

2.7 Branching process

In the branching process, offspring of each individual follows a distribution ψ whose probability mass is given by $P(x)$. Then, we have

$$X_{n+1}=\sum_{i=1}^{X_n}\psi_i^{n+1}$$

with $X_1 = \psi_1^1$. We will be looking at the case where $0 < P(0) < 1$ and $P(0) + P(1) < 1$.

For this Markov Chain, state space is defined as $S = \{0, 1, 2, \dots\}$, and 0 is the absorbing state. Since all the other states are transient, we define ρ as the probability of extinction.

Definition 2.17. *Let $\mu = E[\psi]$. The model is called subcritical if $\mu < 1$; critical if $\mu = 1$; supercritical if $\mu > 1$; and explosive if $\mu = \infty$.*

Theorem 2.11. $\rho = 1$ *iff* $\mu \leq 1$.

3 Stationary distribution

3.1 Stationary distribution

Definition 3.1. *Consider a Markov Chain $\{X_n,\ n = 1, 2, 3, \dots\}$ with state space S. A probability π on S is called a stationary distribution of the chain if*

$$\sum_{x \in S} \pi(x) P(x, y) = \pi(x), \text{ for all } x \in S,$$

where $\mathbb{P} = (P(x, y))$ is the one-step transition matrix.

Lemma 3.1. *If $\pi_0 = \pi$, then $P(X_n = x) = \pi(x)$ for all n.*

Proof. If $n = 0$, $\pi_0 = \pi$. Now, assume $n = k$ is true. Then,

$$\begin{aligned} P(X_{k+1} = x) &= P(X_k = S | X_{k+1} = x) \\ &= \sum_{z \in S} P(X_k = z) P(X_{k+1} = x | X_k = z) \\ &= \sum_{z \in S} \pi(z) P(z, x) \\ &= \pi(x) \end{aligned}$$

By induction, the proof is complete[1]. □

Definition 3.2. *Consider a Markov Chain $\{X_n, n = 1, 2, 3, \dots\}$ with state space S. A probability π on S is called a steady state of the chain if*

$$\lim_{n \to \infty} P^n(x, y) = \pi(y), \text{ for all } x \in S$$

Lemma 3.2. *Let π be a steady state distribution of the Markov chain. Then, for any initial distribution π_0,*

$$\lim_{n \to \infty} P(X_n = y) = \pi(y)$$

Proof. Let $\pi_0(x) = p(x_0 = x)$. Then,

$$\begin{aligned} P(X_n = y) &= \sum_{x \in S} \pi_0(x) P^n(x, y) \\ \lim_{n \to \infty} P(X_n = y) &= \lim_{n \to \infty} \sum_{x \in S} \pi_0(x) P^n(x, y) \\ &= \sum_{x \in S} \lim_{n \to \infty} \pi_0(x) P^n(x, y) \\ &= \sum_{x \in S} \pi_0(x) \lim_{n \to \infty} P^n(x, y) \\ &= \left(\sum_{x \in S} \pi_0(x) \right) \pi(y) = \pi(y) \end{aligned}$$

[1] Since $\pi \mathbb{P} = \pi$, π is the eigenvector of the matrix $\mathbb{P}$ whose eigenvalue is 1.

$\square$

Example 3.1.1. Let $\{X_n,\, n = 0, 1, 2, \dots\}$ be a two state Markov chain with $S = \{0, 1\}$ and $\mathbb{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Since $\pi\mathbb{P} = \pi$ for any π, any distribution is a stationary distribution.

Example 3.1.2. If $\mathbb{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\pi = \begin{pmatrix} 1/2 & 1/2 \end{pmatrix}$ is the only stationary distribution.

Example 3.1.3. Let $\mathbb{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$. To find the stationary distribution, we must solve

$$\begin{pmatrix} \pi(0) & \pi(1) \end{pmatrix} \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} = \begin{pmatrix} \pi(0) & \pi(1) \end{pmatrix}$$

Then, we get

$$(1-p)\pi(0) + q\pi(1) = \pi(0)$$
$$p\pi(0) + (1-q)\pi(1) = \pi(1)$$

Therefore,

$$\begin{cases} \pi(0) = \frac{q}{p+q} \\ \pi(1) = \frac{p}{p+q} \end{cases}$$

Note that

$$\mathbb{P}^n = \begin{pmatrix} \frac{q}{p+q} + (1-p-q)^n \frac{p}{p+q} & \frac{p}{p+q} - (1-p-q)^n \frac{p}{p+q} \\ \frac{q}{p+q} - (1-p-q)^n \frac{q}{p+q} & \frac{p}{p+q} + (1-p-q)^n \frac{q}{p+q} \end{pmatrix}$$

As $n \to \infty$, we get

$$\begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}$$

Therefore, we conclude that this is both stationary and steady state distribution.

Example 3.1.4. Consider a Markov chain characterized by the following transition matrix:

$$\mathbb{P} = \begin{pmatrix} 1/4 & 3/4 \\ 1/3 & 2/3 \end{pmatrix}$$

Clearly, the chain is more likely to be at state 1 than 0. Then, we have

$$P(X_n = 0) \to = \pi(0) = \frac{1/3}{3/4 + 1/3} = \frac{4}{13}$$
$$P(X_n = 1) \to = \pi(1) = \frac{3/4}{3/4 + 1/3} = \frac{9}{13}$$

Example 3.1.5. Let $\{X_n, n = 0, 1, 2, \dots\}$ be a Markov chain with $S = \{0, 1, 2\}$ and

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

Find the stationary distribution of the chain.

First, let $\pi = (\pi(0), \pi(1), \pi(2))$. Since $\pi\mathbb{P} = \pi$, we have

$$\begin{cases} \frac{1}{2}\pi(0) + \frac{1}{3}\pi(1) = \pi(0) \\ \frac{1}{2}\pi(0) + \frac{1}{3}\pi(1) + \frac{1}{2}\pi(2) = \pi(1) \\ \frac{1}{3}\pi(1) + \frac{1}{2}\pi(2) = \pi(2) \end{cases}$$

Then, we find that

$$\pi = (2/7, 2/7, 3/7)$$

Example 3.1.6. Let $\{X_n, n = 1, 2, 3, \dots\}$ be a Birth-death Markov chain with $S = \{0, 1, 2, \dots, d\}$ and

$$\mathbb{P} = \begin{pmatrix} r_0 & p_0 & 0 & \cdots & 0 \\ q_1 & r_1 & p_1 & 0 & \cdots \\ & & & & \\ \vdots & & \ddots & & \\ & & & & \\ \cdots & 0 & q_{n-1} & r_{n-1} & p_{n-1} \\ 0 & \cdots & 0 & q_n & r_n \end{pmatrix}$$

Find the stationary distribution of the chain.

Once again, we use the fact that $\pi\mathbb{P} = \pi$. Then, we end up with the following set of linear equations:

$$\begin{cases} r_0\pi(0) + q_1\pi(1) = \pi(0) \\ p_0\pi(0) + r_1\pi(1) + q_2\pi(2) = \pi(1) \\ \cdots \\ p_{k-1}\pi(k-1) + r_k\pi(k) + q_{k+1}\pi(k+1) = \pi(k) \\ \cdots \\ p_{d-1}\pi(d-1) + r_n\pi(d) = \pi(d) \end{cases}$$

First, we observe that $\pi(1) = \frac{p_0}{q_0}\pi(0)$. Then, we have

$$\begin{aligned} p_0\pi(0) + (1 - p_1 - q_1)\pi(1) + q_2\pi(2) &= \pi(1) \\ p_0\pi(0) - p_1\pi(1) - q_1\pi(1) + q_2\pi(2) &= \pi(1) \\ q_2\pi(2) &= p_1\pi(1) \end{aligned}$$

Then, we have

$$\pi(2) = \frac{p_1}{q_2}\pi(1) = \frac{p_1 p_0}{q_2 q_1}\pi(0)$$

By recursion, we have

$$\pi(k) = \frac{p_0 p_1 \cdots p_{k-1}}{q_1 q_2 \cdots q_k} \pi(0)$$

Since $\pi(0) + \pi(1) + \cdots + \pi(d) = 1$, we have

$$\begin{aligned}
1 &= \pi(0) + \frac{p_0}{q_1}\pi(0)\pi(1) + \frac{p_0 p_1}{q_1 q_2}\pi(0) + \cdots + \frac{p_0 \cdots p_{d-1}}{q_1 \cdots q_d} \\
1 &= \pi(0)\left(1 + \sum_{i=1}^{d} \frac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i}\right) \\
\pi(0) &= \frac{1}{1 + \sum\limits_{i=1}^{d} \frac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i}}
\end{aligned}$$

Therefore,

$$\pi(k) = \frac{\frac{p_0 p_1 \cdots p_{k-1}}{q_1 q_2 \cdots q_k}}{1 + \sum_{i=1}^{d} \frac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i}}$$

Remark. If $d = \infty$, the birth-death chain has a unique stationary distribution iff

$$\sum_{i=1}^{d} \frac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i} < \infty$$

Example 3.1.7. Suppose we have d balls in each of the two urns. Total number of red balls is d (total number of blue balls is also d). Let X_0 be the number of red balls in total in urn 1. We pick a ball from each urn at random and switch. Then, X_1 will be the number of red balls after first switching. We want to find $\mathbb{P}$ and find its stationary distribution.

$P(i, i)$ occurs when we pick red balls or red balls from both urns. Then,

$$P(i, i) = 2\frac{i(d-i)}{d^2}$$

Likewise, we have

$$P(i, i+1) = \frac{(d-i)^2}{d^2}, P(i, i-1) = \frac{i^2}{d^2}$$

Note the boundary conditions:

$$P(0,0) = 0, P(0,1) = 1, P(d,d) = 0, P(d,d-1) = 1.$$

Finally, we can write the transition matrix:

$$\mathbb{P} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{d^2} & \frac{2(d-1)}{d^2} & \frac{(d-1)^2}{d^2} & \cdots & 0 \\ & & & & \\ & & \ddots & & \\ & & & & \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

Since this chain is equivalent to birth-death Markov chain, we know that

$$\pi(k) = \frac{\dfrac{p_0 p_1 \cdots p_{k-1}}{q_1 q_2 \cdots q_k}}{1 + \sum_{i=1}^{d} \dfrac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i}}$$

Observe that

$$\begin{aligned} \frac{p_0 \cdots p_{k-1}}{q_1 \cdots q_k} &= \frac{P(0,1)P(1,2)\cdots P(k-1,k)}{P(1,0)P(2,1)\cdots P(k,k-1)} \\ &= \frac{\dfrac{d^2}{d^2} \cdot \dfrac{(d-1)^2}{d^2} \cdots \dfrac{(d-(k-1))^2}{d^2}}{\dfrac{1}{d^2} \cdot \dfrac{2^2}{d^2} \cdots \dfrac{k^2}{d^2}} \\ &= \frac{(d(d-1)\cdots(d-k+1))^2}{(1 \cdot 2 \cdots k)^2} \\ &= \binom{d}{k}^2 \end{aligned}$$

Then,

$$1 + \sum_{i=1}^{d} \binom{d}{i}^2 = \sum_{i=0}^{d} \binom{d}{i}^2 = \sum_{i=0}^{d} \binom{d}{i}\binom{d}{d-i} = \binom{2d}{d}$$

Therfore,

$$\pi(k) = \frac{\binom{d}{k}\binom{d}{d-k}}{\binom{2d}{d}}$$

3.2 Positive recurrence

We introduce a new notation:

$$m_x = E_x[T_x] = \sum_{k=1}^{\infty} kP(T_x = k | X_0 = x),$$

where $T_x = \min\{n \geq 1, X_n = x\}$.

Definition 3.3. *Let x be a recurrent state. If $m_x < \infty$, then x is called* positive recurrent. *If $m_x = \infty$, then x is called* null recurrent.

Theorem 3.1. *If x is transient, then $m_x = \infty$.*

Proof. If x is transient, $\rho_{xx} < 1$. In other words, $P_x(T_x = \infty) = 1 - \rho_{xx} > 0$. Therefore,

$$m_x = \sum_{k=1}^{\infty} kP(T_x = k|X_0 = x) \geq \infty \cdot P_x(T_x = \infty) = \infty$$

□

Recall that

$$G(x,y) = E_x[N(y)] = E_x\left[\sum_{n=1}^{\infty} I_{\{y\}}(X_n)\right] = \sum_{n=1}^{\infty} P^n(x,y).$$

For any $n \geq 1$, let

$$N_n(y) = \sum_{k=1}^{n} I_{\{y\}}(X_k) \leq n$$
$$G_n(x,y) = E_x[N_n(y)] = \sum_{k=1}^{n} P^k(x,y)$$

Theorem 3.2.

1. *If y is transient, then*

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = 0, \lim_{n\to\infty} \frac{G_n(x,y)}{n} = 0$$

for all x.

2. *If y is recurrent, then*

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{I\{T_y < \infty\}}{m_y}, \lim_{n\to\infty} \frac{G_n(x,y)}{n} = \frac{\rho_{xy}}{m_y}$$

Corollary. Let C be an irreducible set of recurrent states. Then,

$$\lim_{n\to\infty} \frac{G_n(x,y)}{n} = \lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{1}{m_y}, \forall x,y \in C$$

Theorem 3.3. *If x is positive recurrent and $x \to y$, then y is positive recurrent.*

Proof. If x is positive recurrent and $x \to y$, then $x \leftrightarrow y$. In other words, there exists $n_1 \geq 1, n_2 \geq 1$ such that $P^{n_1}(x,y) > 0, P^{n_2}(xy) > 0$. Observe that

$$P^{n_1+n+n_2}(y,y) \geq P^{n_2}(y,x)P^n(x,x)P^{n_1}(x,y)$$

Observe that,

$$\sum_{k=1}^{n} P^{n_1+n+n_2+k}(y,y) \geq P^{n_2}(y,x)\left[\sum_{k=1}^{n} P^k(x,x)\right] P^{n_1}(y,x)$$

Then,

$$\begin{aligned}
&\sum_{m=n_1+n_2+1}^{n+n_1+n_2} P^m(y,y) = \sum_{k=1}^{n} P^{n_1+n_2+k}(y,y) \\
&= -\sum_{m=1}^{n_1+n_2} P^m(y,y) + \sum_{m=1}^{n_1+n_2} P^m(y,y) + \sum_{m=n_1+n_2+1}^{n+n_1+n_2} P^m(y,y) \\
&= G_{n+n_1+n_2}(y,y) - G_{n_1+n_2}(y,y) \\
&\geq P^{n_2}(y,x)G_n(x,x)P^{n_1}(x,y)
\end{aligned}$$

Then,

$$\frac{G_{n+n_1+n_2}(y,y) - G_{n_1+n_2}(y,y)}{n} \geq P^{n_2}(y,x)\frac{G_n(x,x)}{n}P^{n_1}(x,y)$$

Since $G_{n_1+n_2}(y,y) \to 0$ as $n \to \infty$, we have[2]

$$\frac{1}{m_y} \geq P^{n_1}(x,y)P^{n_2}(y,x)\frac{1}{m_x} > 0$$

If m_x is finite, then m_y must be finite as well. □

Theorem 3.4. *Let C be a finite irreducible set of recurrent states. Then, every state in C is positive recurrent.*

Proof. Clearly, given $x \in C$,

$$\sum_{y \in C} P^k(x,y) = 1,$$

[2] Above result is derived from the following:

$$\lim_{n\to\infty} \frac{G_{n_1+n_2+n}}{n} = \lim_{n\to\infty} \frac{(n_1+n_2+n)G_{n_1+n_2+n}}{n(n_1+n_2+n)} = 1 \cdot \frac{1}{m_y}$$

for all positive integer k. Then,

$$\begin{aligned}
n &= \sum_{k=1}^{n} \sum_{y \in C} P^n(x, y) \\
&= \sum_{y \in C} \sum_{k=1}^{n} P^k(x, y) \\
&= \sum_{y \in C} G_n(x, y)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\sum_{y \in C} \frac{G_n(x, y)}{n} = 1 \\
\implies &\lim_{n \to \infty} \sum_{y \in C} \frac{G_n(x, y)}{n} = 1 \\
\implies &\sum_{y \in C} \lim_{n \to \infty} \frac{G_n(x, y)}{n} = 1
\end{aligned}$$

Now, it follows that there exists $z \in C$ such that

$$\lim_{n \to \infty} \frac{G_n(x, y)}{n} > 0,$$

implying that $m_z < \infty$ and z is positive recurrent. Since every state in C communicated with z, every state in C is positive recurrent. $\square$

Remark. Let $\{X_n, n = 1, 2, \dots\}$ be a Markov chain with *finite* state space S. Then, all recurrent states are positive recurrent.

Theorem 3.5. *Let π be a stationary distribution of a Markov chain. If y is transient or null recurrent, then $\pi(y) = 0$.*

Proof. Let π be a stationary distribution of the chain. Then, $P(X_k = y) = \pi(y)$ if the initial distribution is π. Then,

$$\begin{aligned}
P(X_k = y) &= P(X_0 \in S, X_k = y) \\
&= \sum_{x \in S} P(X_0 = x, X_k = y) \\
&= \sum_{x \in S} P(X_0 = x) P(X_k = y | X_0 = x) \\
&= \sum_{x \in S} \pi(x) P^k(x, y) = \pi(y)
\end{aligned}$$

Then, we get

$$\begin{aligned}\pi(y) &= \frac{1}{n}\sum_{k=1}^{n}\sum_{x\in S}\pi(x)P^k(x,y)\\ &= \sum_{x\in S}\frac{1}{n}\sum_{k=1}^{n}\pi(x)P^k(x,y)\\ &= \sum_{x\in S}\pi(x)\frac{1}{n}\sum_{k=1}^{n}P^k(x,y)\\ &= \sum_{x\in S}\pi(x)\frac{G_n(x,y)}{n}\end{aligned}$$

Thus,

$$\begin{aligned}\pi(y) &= \lim_{n\to\infty}\sum_{x\in S}\pi(x)\frac{G_n(x,y)}{n}\\ &= \sum_{x\in S}\pi(x)\lim_{n\to\infty}\frac{G_n(x,y)}{n} = 0\end{aligned}$$

□

Example 3.2.1. Does a Markov Chain with no positive recurrent states have a stationary distribution?

Let π be a stationary distribution. By theorem 3.5, $\pi(y) = 0$ for all $y \in S$. This yields a contradiction and a Markov Chain with no positive recurrent states cannot have a stationary distribution.

Theorem 3.6. *An irreducible positive recurrent Markov Chain has a unique stationary distribution given by*

$$\pi(x) = \frac{1}{m_x}$$

Example 3.2.2. Does a finite state Markov Chain have a stationary distribution?

Consider the following decomposition of the state space:

$$\begin{aligned}C &= C_R \cup C_T\\ &= C_{PR} \cup C_T\\ &= C_1 \cup \cdots \cup C_j \cup C_T\end{aligned}$$

Each C_i is an irreducible positive recurrent class. By the theorem, there is a stationary on each C_i. Then,

$$\begin{cases}\pi_i(x) = \frac{1}{m_x} & x \in C_i\\ \pi_i(x) = 0 & x \notin C_i\end{cases}$$

is a stationary distribution.

Example 3.2.3. When does a finite state Markov Chain have a unique stationary distribution?

All recurrent states must communicate.

Example 3.2.4. Can a finite state Markov Chain have exactly two stationary distributions?

Assume π_2, π_2 are two different stationary distributions:

$$\pi_1 \mathbb{P} = \pi_1, \pi_2 \mathbb{P} = \pi_2$$

Then, for any $0 \leq \lambda \leq 1$, $\lambda \pi_1 + (1-\lambda)\pi_2$ is also a stationary distribution. Therefore, a Markov Chain cannot have exactly two stationary distributions.

Theorem 3.7. *Let $\{X_n, n = 0, 1, 2, \dots\}$ be a Markov Chain with state space S. Let N denote the nnumber of stationary distributions of the chain. Then,*

$$N = \begin{cases} 0 & \textit{if there is no positive recurrent state} \\ 1 & \textit{if there is one irreducible positive recurrent class} \\ \infty & \textit{else} \end{cases}$$

Example 3.2.5. Consider the Birth-Death Markov chain. If

$$\sum \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} < \infty,$$

then there is a stationary distribution. This condition is in fact equivalent to positive recurrence of a chain.

Previously, we have shown that an irreducible Birth-Death Markov Chain with $X = \{0, 1, 2, \dots\}$ (1) is recurrent iff

$$\sum_{x=1}^{\infty} \frac{q_1 \cdots q_x}{p_1 \cdots p_x} = \infty,$$

and (2) has a unique stationary distribution iff

$$\sum_{x=1}^{\infty} \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} < \infty.$$

These results lead to following theorems:

Theorem 3.8.

1. *The chain is transient iff*

$$\sum_{x=1}^{\infty} \frac{q_1 \cdots q_x}{p_1 \cdots p_x} < \infty$$

2. *The chain is null recurrent iff*

$$\sum_{x=1}^{\infty} \frac{q_1 \cdots q_x}{p_1 \cdots p_x} = \infty, \sum_{x=1}^{\infty} \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} = \infty$$

3. *The chain is positive recurrent iff*

$$\sum_{x=1}^{\infty} \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} < \infty$$

Example 3.2.6. A Birth-Death Markov Chain with $p_x = q_x = a$ is null recurrent.

Example 3.2.7. A Birth-Death Markov Chain with $p_x = p$ and $q_x = q$. Notice that

$$\sum_{x=1}^{\infty} \frac{q_1 \cdots q_x}{p_1 \cdots p_x} = \sum_{x=1}^{\infty} \left(\frac{q}{p}\right)$$
$$\sum_{x=1}^{\infty} \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} = \sum_{x=1}^{\infty} \left(\frac{p}{q}\right)$$

Clearly, if $p \neq q$, one of them will converge and the other will diverge. Thus, when $p > q$, the chain is transient and when $p < q$, the chain is positive recurrent.

4 Long time behaviour

4.1 Period of a state

Definition 4.1. *Let I be a set of positive integres. The* greatest common divisor *of I, denoted by* $\gcd(I)$ *is defined as*

$$\gcd(I) = \min\{n,\ n|m \textit{ for all } m \in I\}$$

Definition 4.2. *The* period *of a state x is defined as*

$$d_x = \gcd\{n : P^n(x,x) > 0\}$$

Example 4.1.1. Let $S = \{1,2,3\}$. Consider

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

Since $P^2(3,3) > 0$ and $P^3(3,3) > 0$, its period is 1.

Example 4.1.2. Consider

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

Clearly, $1 \to 2 \to 3 \to 1$. Therefore, period of 1 is 3.

Example 4.1.3. Consider a pure birth Markov chain with $P(x, x+1) = 1$. Then, $d_0 = \infty$.

Definition 4.3. *A state x is called periodic if $d_x = 1$.*

Theorem 4.1. *Let $I_x = \{n \geq 1 : P^n(x,x) > 0\}$. By definition, $d_x = \gcd I_x$. If $1 \in I_x$, then $d_x = 1$.*

Example 4.1.4. Consider a Markov Chain characterized by the following transition matrix:

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

Find the period of each state.

1. For $x = 0$, we have $I_0 = \{2,4,6,\dots\}$ because $0 \to 1 \to 0$. Therefore, $d_0 = 2$.

2. For $x = 1$, we have $I_1 = \{2,4,6,\dots\}$ for a similar reason. Therefore, $d_0 = 2$.

3. For $x = 2$, notice that 2 is an absorbing state. Since $2 \to 2$, we have $1 \in I_2$. Therefore, $d_2 = 1$.

Theorem 4.2. *If $x \leftrightarrow y$, then $d_x = d_y$.*

Proof. Let

$$I_x = \{n \geq 1 : P^n(x,x) \geq 0\},$$
$$I_y = \{n \geq 1 : P^n(y,y) \geq 0\}.$$

Since $x \leftrightarrow y$, there exists $n_1 \geq 1, n_2 \geq 1$ such that $P^{n_1}(x,y) > 0$ and $P^{n_1}(x,y) > 0$. Thus, $P^{n_1+n_2} > 0$.

Since $n_1 + n_2 \in I_x$, d_x is a divisor of $n_1 + n_2$. For any $m \in I_y$, we have $P^m(y,y) > 0$. Then,

$$P^{n_1+n_2+m}(x,x) \geq P^{n_1}(x,y)P^m(y,y)P^{n_2}(y,x) > 0$$

This implies that $n_1 + n_2 + m \in I_x$. Since d_x is a divisor of $n_1 + n_2 + m$ and of $n_1 + n_2$, it is follows that that d_x is also a divisor of m. Thus, $d_x \leq d_y$. Likewise, we can prove that $d_y \leq d_x$. Therefore, $d_x = d_y$. □

Recall that we can decompose any state space as follows:

$$\Big(\bigcup_i C^i_{PR}\Big) \cup \Big(\bigcup_j C^j_{NR}\Big) \cup C_T$$

By the previous theorem, we can conclude that each class has exactly one period.

4.2 Long time behaviour

Theorem 4.3. *If y is null recurent, then*

$$\lim_{n\to\infty} P^n(x,y) = 0$$

for all x.

Theorem 4.4. *If $\{X_n, n = 0, 1, \dots\}$ is an irreducible positive recurrent Markov chian with period d,*

1. *If the chain is periodic, then*

$$\lim_{n\to\infty} P^n(x,y) = \pi(y)$$

2. *If the period of the chain is greater than or equal to 2, then for any $x, y \in S$ there exists an integer $0 \leq r < d$ such that*

$$\lim_{m\to\infty} P^{md+r}(x,y) = d\pi(y)$$

Further, $P^n(x,y) = 0$ if $n \neq md + r$.

Theorem 4.5 (Long time behaviour)**.**

$$\lim_{n\to\infty} P^n(x,y) = \begin{cases} 0 & \textit{if } y \textit{ is transient or null recurrent} \\ \pi(y) & \textit{if } y \textit{ is positive recurrent and } d_y = 1 \\ \begin{cases} d\pi(y) & \textit{for } n = md + r \\ 0 & \textit{else} \end{cases} & \textit{if } \rho_{xy} = 1, y \textit{ is PR, and } d_y \geq 2 \end{cases}$$

Example 4.2.1. Consider a Markov chain characterized by the following transition matrix:

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

Then,

$$\lim_{n\to\infty} P^n(x,y) = \begin{cases} 0 & \text{if } y = 0,1 \\ 1 & \text{if } x = y = 2 \end{cases}$$

Example 4.2.2. Let $\{X_n, n = 0, 1, 2\}$ be a Markov chain with $S = \{1, 2, 3, 4\}$ and

$$\mathbb{P} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/6 & 1/2 & 1/3 & 0 \\ 0 & 1/2 & 1/2 & 1/6 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}$$

Find $\lim_{n\to\infty} \mathbb{P}^n$.

We want to compute

$$\mathbb{P}^n = \begin{pmatrix} P^n(1,1) & P^n(1,2) & P^n(1,3) & P^n(1,4) \\ P^n(2,1) & P^n(2,2) & P^n(2,3) & P^n(2,4) \\ P^n(3,1) & P^n(3,2) & P^n(3,3) & P^n(3,4) \\ P^n(4,1) & P^n(4,2) & P^n(4,3) & P^n(4,4) \end{pmatrix}$$

First, notice that

$$1 \to 2 \to 3 \to 4 \to 1$$

so all state communicate with each other. Since we have finite state space, we have one close, irreducible, positive recurrent class. Also, we have

$$d_1 = \gcd\{n \geq 1, P^n(1,1) > 0\} = 1$$

since $P^1(1,1) = 1/2 > 0$. Since all states are positive recurrent and have period of 1, we can conclude that

$$\lim_{n\to\infty} P^n(x,y) = \pi(y)$$

for all states in S. Therefore,

$$\lim_{n\to\infty} \mathbb{P}^n = \begin{pmatrix} 1/8 & 3/8 & 3/8 & 1/8 \\ 1/8 & 3/8 & 3/8 & 1/8 \\ 1/8 & 3/8 & 3/8 & 1/8 \\ 1/8 & 3/8 & 3/8 & 1/8 \end{pmatrix}$$

Example 4.2.3. Let $\{X_n, n = 0, 1, 2\}$ be a Markov chain with $S = \{1, 2, 3, 4\}$ and

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

Find $\lim_{n\to\infty} \mathbb{P}^{2n}$ and $\lim_{n\to\infty} \mathbb{P}^{2n+1}$.

Clearly, this chain is irreducible and positive recurrent. We also know that its stationary distribution is given by

$$\pi = \left(\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}\right)$$

Observe that

$$\begin{aligned} d_1 &= \gcd\{n \geq 1, P^n(1,1) > 0\} \\ &= \gcd\{2, 3, 6, \dots\} \\ &= 2 \end{aligned}$$

So there must exist r such that $\lim_{n\to\infty} P^n(x, y) = d\pi(y)$ for $n = md + r$.

Notice that $P^{2n+1}(1, 1) = 0$. Then, we must have

$$\lim_{n\to\infty} P^{2n}(1,1) = 2\pi(1) = \frac{1}{4}.$$

Likeiwse, we have $P^{2n}(2, 1) = 0$. Therefore, we must have

$$\lim_{n\to\infty} P^{2n+1}(2,1) = 2\pi(1) = \frac{1}{4}.$$

Therefore, we can conclude that

$$\mathbb{P}^{2n} \to \begin{pmatrix} 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \end{pmatrix}$$

$$\mathbb{P}^{2n+1} \to \begin{pmatrix} 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \end{pmatrix}$$

Example 4.2.4. Let $\{X_n, n = 0, 1, 2\}$ be a Markov chain with $S = \{0, 1, 2, \dots\}$ and

$$\mathbb{P} = \begin{pmatrix} r_0 & p_0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & \cdots \\ \vdots & & & & \ddots \end{pmatrix}$$

Clearly, we have

$$d = \begin{cases} 1 & \text{if } r_i > 0 \text{for at least one } i \\ 2 & \text{else} \end{cases}$$

Example 4.2.5. Let $\{X_n, n = 0, 1, 2, \dots\}$ be a Markov chain with finite number of states.

Assume that the chain is irreducible and each column of $\mathbb{P}$ add up to 1. Find the stationary distribution of the chain.

5 Continuous Time Markov Chain

In this section, we will still be looking at at most countable state space. However, we introduce a new concept:

Definition 5.1 (Waiting time)**.** *Starting from state x, you wait at x for τ_x until the next "jump" occurs. In discrete time Markov chain, τ_x was fixed but in continuous time, τ_x is a random variable.*

In order to satisfy the Markov property, waiting time must follow the exponential distribution, the only random variable with memoryless property.

5.1 Exponential distribution

Definition 5.2 (Exponential distribution)**.** *A random variable X that follows an exponential distribution has the following properties:*

1. $X > 0$
2. $f(x) = \lambda e^{-\lambda x}$
3. $E[x] = 1/\lambda$

We observe that if λ is large, the waiting time becomes shorter. If λ goes to infinity, waiting time will go to zero. On the other hand, if λ goes to 0, waiting time will go to infinity. These are two boundary cases.

Example 5.1.1. Let X_1, X_2 be two independent exponential random variables with respective parameters: λ_1, λ_2. Set $X = \min(X_1, X_2)$. Then, X is exponential with parameter $\lambda = \lambda_1 + \lambda_2$

Proof. Notice that

$$\begin{aligned} P(X \leq x) &= 1 - P(X > x) \\ &= 1 - P(\min(X_1, X_2) > x) \\ &= 1 - P(X_1 > x, X_2 > x) \\ &= 1 - P(X_1 > x)P(X_2 > x) \end{aligned}$$

Since CDF of an exponential distribution is $F(x) = 1 - e^{-\lambda x}$, we get

$$\begin{aligned} 1 - P(X_1 > x)P(X_2 > x) &= 1 - (1 - F_{\lambda_1}(x))(1 - F_{\lambda_2}(x)) \\ &= 1 - e^{-\lambda_1 x} e^{-\lambda_2 x} \\ &= 1 - e^{-(\lambda_1 + \lambda_2)x} \end{aligned}$$

Therefore, X follows an exponential distribution with parameter $\lambda = \lambda_1 + \lambda_2$. $\square$

Theorem 5.1. *Let $X_1, X_2, \ldots, X_n$ be independent exponential random variables. Let $X = \min(X_1, X_2, \ldots, X_n)$. Then, X is a exponential random variable with parameter $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.*

Recall that the Markov property is defined as the following:

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_n = x_n | X_{n-1} = x_{n-1})$$

How does this apply in the continuous case?

Definition 5.3. *For any $n \geq 1$, let $x_0, x_1, \dots, x_n \in S$. Then, the Markov property is defined as the following:*

$$P(X_{\tau_n} = x_n | X_0 = x_0, \dots, X_{\tau_{n-1}} = x_{n-1}) = P(X_{\tau_n} = x_n | X_{\tau_{n-1}} = x_{n-1})$$

We now introduce a new notation. For any $0 \leq s \leq t$, and $x, y \in S$, we have

$$P(s, t, x, y) = P_{xy}(s, t) = P(X_t = y | X_s = x)$$

Definition 5.4. *Let $T = [0, \infty)$ and S be a finite or countable set. The stochastic process is given by $\{X_t, t \in T\} = \{X_t, t \geq 0\}$. Then,*

1. *The distribution of X_0 is called the* initial distribution *of the process, denoted by π.*

2. *The process $\{X_t, t \geq 0\}$ is* continuous time Markov chain *if for any $n \geq 1, x_0, x_1, \dots, x_n \in S, 0 \leq t_1 < \cdots < t_n$,*

$$P(X_{t_n} = x_n | X_{t_0} = x_0, \dots, X_{t_{n-1}} = x_{n-1}) = P(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1})$$

3. *The Markov chain $\{X_t, t > 0\}$ is* time homogeneous *if for any $t, s \in T$,*

$$P(X_{t+s} = y | X_s = x) = P(X_t = y | x_0 = x)$$

4. *For time-homogeneous Markov chain, $P_{xy}(t) = P(X_t = y | X_0 = x)$ is called the* transition function *of the chain.*

5. *A state x is absorbing if $P(X_t = x | x_0 = x) = 1$ for all t.*

6. $\delta_{xy} = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{else} \end{cases}$

7. *$Q = (Q_{xy})$ is a transition probbability matrix such that $Q_{xx} = 0$.*

8. *For each $x \in S$, let $q_x \geq 0$. Set* $q_{xy} = \begin{cases} -q_x & \text{if } x = y \\ q_x Q_{xy} & \text{if } x \neq y \end{cases}$

Theorem 5.2. *Let $\{X_t, t > 0\}$ be a homogeneous continuous time Markov chain with waiting time distribution $\{\exp(q_x), x \in S\}$ and transition probability Q. Then, we have*

1. *Chaphen-Kolmogorov Equation:*

$$P_{xy}(t + s) = \sum_{z \in S} P_{xz}(t) P_{zy}(s)$$

2. Backward Equation:

$$P'_{xy}(t) = \sum_{z\in S} q_{xz}P_{zy}(t) \iff \mathbb{P}'(t) = A\mathbb{P},$$

where $A = (q_{xy})$.

Proof. For any $x, y \in S$ and $t, s \in T$, we have

$$\begin{aligned}
P_{xy}(t+s) &= P(X_{t+s} = y|X_0 = x)\\
&= P(X_{t+s} = y, X_t \in S|X_0 = x)\\
&= \frac{P(X_{t+s} = y, X_t \in S, X_0 = x)}{P(X_0 = x)}\\
&= \frac{\sum_{z\in S} P(X_{t+s} = y, X_t = z, X_0 = x)}{P(X_0 = x)}\\
&= \sum_{z\in S} \frac{P(X_{t+s} = y, X_t = z, X_0 = x)}{P(X_t = z, X_0 = x)}\,\frac{P(X_t = z, X_0 = x)}{P(X_0 = x)}\\
&= \sum_{z\in S} P(X_{t+s} = y|X_0 = x, X_t = z)P(X_t = z|x_0 = x)\\
&= \sum_{z\in S} P(X_{t+s} = y|X_t = z)P(X_t = z|x_0 = x)\\
&= \sum_{z\in S} P_{zy}(s)P_{xz}(t)
\end{aligned}$$

This proves the first statement.

Now, we prove the second statement:

$$\begin{aligned}
P_{xy}(t) &= P(X_t = y|X_0 = x)\\
&= \underbrace{P(\tau_1 > t, X_t = y|X_0 = x)}_{\text{no jump occurs}} + P(\tau_1 \le t, X_t = y|X_0 = x)\\
&= \delta_{xy}\underbrace{P(\tau_1 > t)}_{e^{-q_x t}} + P(\tau_1 \le t, X_{\tau_1} \ne x, X_t = y|X_0 = x)\\
&= \delta_{xy}e^{-q_x t} + \sum_{z\ne x} P(\tau_1 \le t, X_{\tau_1} \ne x, X_t = z|X_0 = x)\\
&= \delta_{xy}e^{-q_x t} + \sum_{z\ne x}\int_0^t q_x e^{-q_x s} Q_{xz}P_{zy}(t-s)ds\\
&= \delta_{xy}e^{-q_x t} + \int_0^t q_x e^{-q_x s}\sum_{z\ne x} Q_{xz}P_{zy}(t-s)ds\\
&= \delta_{xy}e^{-q_x t} + \int_0^t q_x e^{-q_x (t-u)}\sum_{z\ne x} Q_{xz}P_{zy}(u)du\\
&= \delta_{xy}e^{-q_x t} + e^{-q_x t}\int_0^t q_x e^{q_x u}\sum_{z\ne x} Q_{xz}P_{zy}(u)du
\end{aligned}$$

Then, we get

$$
\begin{aligned}
P'_{xy}(t) &= -q_x e^{-q_x t}\left[\delta_{xy} + \int_0^t q_x e^{q_x u}\sum_{z\neq x} Q_{xz}P_{zy}(u)du\right] + q_x \sum_{z\neq x} Q_{xz}P_{zy}(t) \\
&= -q_x P_{xy}(t) + \sum_{z\neq x} q_x Q_{xz} P_{zy}(t) \\
&= q_{xx}P_{xy}(t) + \sum_{z\neq x} q_{xz}P_{zy} \\
&= \sum_z q_{xz}P_{zy}(t)
\end{aligned}
$$

Now, this completes the proof. □

So the matrix A determines the existence of Markov chain and is analogous to transition matrix in discrete time Markov chain. Notice that Q provies the *jumping mechanism* and q_x provides the *waiting mechanism*. So the matrix A, which is a combination of q_x and Q, is fundamental in continuous Markov chains.

What happens if we let $t \to 0$? We get

$$P'_{xy}(0) = \sum_{z\in S} q_{xz}P_{zy}(0)$$

We can consider two cases here:

1. $y = x$. $P'_{xx}(0) = q_{xx} = -q_x$.
2. $y \neq x$. $P'_{xy}(0) = q_{xy}$.

Now, notice that $A = q_{xy}$ is not a transition matrix. We can first look at some of its properties:

- $|q_{xy}| \in [0, \infty]$
- $q_{xy} > 0$ for $y \neq x$, $q_{xy} \leq 0$ for $y = x$
- $\sum_y q_{xy} = 0$

Let's take a look at an example:

Example 5.1.2. Consider the following matrix:

$$A = \begin{pmatrix} -100 & 60 & 40 \\ 10 & -20 & 10 \\ 1 & 1 & -2 \end{pmatrix}$$

If we start from state 1, we are very likely to move to other states (notice that the magnitude of A_{11} is large). If we start from state 3, we are not as likely to move to other states. This directly translates to waiting time.

Also, we notice that $A_{21} = A_{23}$ and $A_{31} = A_{32}$. This implies that you are equally likely to move to any other states if you start from either state 2 or 3. On the other hand, $A_{12} > A_{13}$. So if you start from state 1, you are more likely to move to state 2 than 3.

Let's go back to Champhen-Kolmogorov equation:

$$P_{xy}(t+h) = \sum_{z \in S} P_{xz}(t) P_{zy}(h)$$

We can then divide both sides by h

$$\frac{P_{xy}(t+h) - P_{xy}(t)}{h} = \frac{\sum_{z \in S} P_{xz}(t) P_{zy}(h) - P_{xy}(t)}{h}$$

As we let $h \to 0$, we get the *forward equation*.

$$P'_{xy} = \sum_{z} P_{xz}(t) q_{zy}$$

Example 5.1.3 (Poisson process)**.** Consider

$$A = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & \cdots \\ \vdots & & & & & \ddots \end{pmatrix}$$

Find the distribution of X_t.

First, notice that

$$\begin{aligned} P(X_t = 0) &= P_0(t) \\ &= P(\tau_1 > t) \\ &= 1 - P(\tau_1 < t) \\ &= 1 - (1 - e^{-\lambda t}) = e^{-\lambda t} \end{aligned}$$

Then, by using the forward equation, we can compute $P_0(X_t = 1) = P'_{01}(t)$:

$$\begin{aligned} P'_{01}(t) &= \sum_{z \in S} P_{0z}(t) q_{z1} \\ &= P_{00}(t) q_{01} + P_{01}(t) q_{11} \\ &= \lambda P_{00}(t) - \lambda P_{01}(t) \end{aligned}$$

Solving the differential equation, we get

$$P_{01}(t) = \lambda t e^{-\lambda t}.$$

By induction, we find that

$$\begin{aligned} e^{\lambda t} P_{0n}(t) &= \int_0^t \lambda e^{\lambda s} P_{0(n-1)}(s) ds \\ &= \int_0^n \lambda e^{\lambda s} \frac{(\lambda s)^{n-1}}{n-1} e^{-\lambda s} ds \\ &= \frac{(\lambda t)^n}{n!} \end{aligned}$$

Therefore, we get

$$P_{0n}(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

Example 5.1.4. Consider a Markov chain with state space $S = \{0, 1\}$ where

$$A = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$$

Find $P_{xy}(t)$ for $x, y \in S$.

In this example, we can use the backward equation. First, observe that

$$\begin{aligned} P'_{00}(t) &= \sum_{z \in S} q_{0z} P_{z0}(t) \\ &= q_{00} P_{00}(t) + q_{01} P_{10}(t) \\ &= -\lambda P_{00}(t) + \lambda P_{10}(t) \end{aligned}$$

Likewise, we have

$$\begin{aligned} P'_{10}(t) &= \sum_{z \in S} q_{1z} P_{z0}(x) \\ &= q_{10} P_{00}(t) + q_{11} P_{10}(t) \\ &= \mu P_{00}(t) - \mu P_{10}(t) \end{aligned}$$

Combining the two equations, we get

$$P'_{00}(t) - P'_{10}(t) = -(\lambda + \mu)(P_{00}(t) - P_{10}(t)).$$

Solving the differential equation, we get

$$P_{00}(t) - P_{10}(t) = e^{-(\lambda+\mu)t}$$

Then, we have

$$\begin{aligned} P'_{00}(t) &= -\lambda P_{00}(t) + \lambda (P_{00}(t) - e^{-(\lambda+\mu)t}) \\ &= -\lambda e^{-(\lambda+\mu)t} \end{aligned}$$

Integrating,

$$\begin{aligned}
P_{00}(t) - P_{00}(0) &= -\lambda \int_0^t e^{-(\lambda+\mu)s} ds \\
&= -\frac{\lambda}{\lambda+\mu}(1 - e^{-(\lambda+\mu)t}) \\
\implies P_{00}(t) - 1 &= -\frac{\lambda}{\lambda+\mu}(1 - e^{-(\lambda+\mu)t}) \\
\implies P_{00}(t) &= \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}
\end{aligned}$$

Lastly, we have

$$P_{10}(t) = \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

Example 5.1.5 (Birth-death continuous time)**.** We can write a general infinitesimal matrix for continuous time birth-death Markov chain as follows:

$$A = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \\ \vdots & & & & \ddots \end{pmatrix}$$

Example 5.1.6. Consider

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & & & & \ddots \end{pmatrix}$$

Then, we get $P_{00}(t) = 1, P_{55}(t) = 1, P_{17}(t) = 0$.

Example 5.1.7 (Branching process)**.** Each individual will wait an exponential time with parameter $\lambda > 0$ independently. At the end of the waiting time, the individual will produce 2 offsprings with probability p or no offsprings with probability $1-p$. Let X_t be the total number of individuals at time t.

It is clear that $S = \{0, 1, 2, \dots\}$. First, we know that $q_{00} = 0$ because if you have no population, it will stay at 0 forever and no offsprings will not be produced. Then, it directly follows that $q_{0i} = 0$ for all $i > 0$.

Now, we look at q_{11}. Since we have only one individual whose waiting time has parameter λ, we have $q_{11} = -\lambda$. Using the probabilities given, we get $q_{10} = (1-p)\lambda$ and $q_{12} = p\lambda$.

Notice that each individual has an independent waiting time distribution. So two individuals cannot have identical waiting time so $p(\tau_i = \tau_j) = 0$ for all $i \neq j$ where τ_i is the waiting time of an individual. So you can either go to state $x+1$ (produce 2 offsprings) or $x-1$ (produce 0 offsprings) given that you're at state x. Now, we want to compute q_{xx}. Notice that the jump will happen when the first person out of x individuals produces offsprings. So we want the minimum time of n waiting times. By theorem 5.1, we know that the minimum waiting time is an exponential distribution with λx. So we get $q_{xx} = -\lambda x$. Then, it directly follows that $q_{x\,x+1} = \lambda x p$ and $q_{x\,x-1} = \lambda x(1-p)$.

Example 5.1.8 (Infinite server Queueing model)**.** Customers arrive for service according to a poisson process with parameter λ. Each customer will be served after arrival. The serving then follows exponential with parameter μ. All services are independent.

Then, we want to look at the transition matrix, Q. Notice that a customer arrives at a rate λ and any of the customers can be served at a per customer rate μ. Then, clearly, the probability of going from x to $x+1$ is $\lambda/(\lambda + x\mu)$ and the probability of going from x to $x-1$ is $x\mu/(\lambda + x\mu)$, where x is the number of customers.

So we can write hte infinitesimal matrix as follows:

$$A = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots \\ \mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & \cdots \\ & & & \\ & & & \end{pmatrix}$$

Example 5.1.9 (Pure birth Markov Chain)**.** For pure bith Markov Chain, we have $\mu_x = 0$ for all x. Then, the forward equation is given by

$$\begin{aligned} P'_{xy}(t) &= \sum_{z \in S} P_{xz}(t) q_{zy} \\ &= P_{x(y-1)}(t) q_{(y-1)y} + P_{xy}(t) q_y y \\ &= P_{x(y-1)}(t) \lambda_{y-1} - \lambda_y P_{xy}(t) \end{aligned}$$

Therefore, we get

$$P'_{xy}(t) = -\lambda_y P_{xy}(t) + \lambda_{y-1} P_{x(y-1)}(t)$$

Now, we want to solve this system of equation.

If $y < x$, then $P_{xy}(t) = 0$.

If $y = x$, then we have $\lambda_{x-1} P_{x(x-1)}(t)$. Thus, we have

$$\begin{aligned} & P'_{xx}(t) = -\lambda_x P_{xx}(t) \\ \implies & P'_{xx}(t) = e^{-\lambda_x t}, P_{xx}(0) = 1 \end{aligned}$$

If $y = x + 1$, we get

$$\begin{aligned} P_{x(x+1)}(t) &= -\lambda_{x+1} P_{x(x+1)}(t) + \lambda_x P_{xx}(t) \\ &= -\lambda_{x+1} P_{x(x+1)}(t) + \lambda_x e^{-\lambda_x t}, \end{aligned}$$

which yields

$$P'_{x(x+1)} + \lambda_{x+1} P_{x(x+1)} = \lambda_x e^{-\lambda_x t}.$$

Integrating, we get

$$P_{x(x+1)}(t) = \begin{cases} \lambda_x t e^{-\lambda_{x+1} t}, & \lambda_{x+1} = \lambda_x \\ \frac{\lambda_x}{\lambda_{x+1}-\lambda_x} e^{-\lambda_{x+1} t} \left(e^{(\lambda_{x+1]}-\lambda_x)t} - 1 \right), & \lambda_{x+1} \neq \lambda_x \end{cases}$$

Lastly, if $y > x+1$, we get

$$\begin{aligned} P'_{xy}(t) &= -\lambda_y P_{xy}(t) + \lambda_{y-1} P_{x(y-1)}(t) \\ e^{\lambda_y t}\left(P'_{xy}(t) + \lambda_y P_{xy}(t)\right) &= \lambda_{y-1} P_{x(y-1)}(t) e^{\lambda_y t} \\ \left(e^{\lambda_y t} P_{xy}(t)\right)' &= \lambda_{y-1} P_{x(y-1)}(t) e^{\lambda_y t} \\ \implies P_{xy}(t) &= \lambda_{y-1} \int_0^t e^{-\lambda_y (t-s)} P_{x(y-1)}(s) ds. \end{aligned}$$

This is the general pure birth process.

Example 5.1.10 (Pure birth)**.** Pure birth happens when $\mu_x > 0$ and $\lambda_x = 0$. This relates to a biological process called Coalescent Markov Chain.

Example 5.1.11 (Yule Process)**.** Yule process is a linear growth Markov chain were $\lambda_x = ax$, where $a > 0$. Essentially, it is a special type of Pure Birth process. So let's find the solution.

When $y < x$, we get $P_{xy}(t) = 0$.

When $y = x$, we get $P_{xx}(t) = e^{-axt}$.

When $y = x+1$, we get

$$\begin{aligned} P_{x(x+1)}(t) &= \frac{\lambda_x}{\lambda_{x+1} - \lambda_x} \left(e^{-\lambda_x t} - e^{-\lambda_{x+1} t} \right) \\ &= x \left(e^{-at} \right)^x \left(1 - e^{-at} \right). \end{aligned}$$

When $y = x+2$, we get

$$\begin{aligned} P_{x(x+2)}(t) &= \lambda_{x+1} \int_0^t e^{\lambda_{x+2}(t-s)} P_{x(x+1)}(s) ds \\ &= \lambda_{x+1} \int_0^t e^{\lambda_{x+2}(t-s)} x \left(e^{-as} \right)^x \left(1 - e^{-as} \right) ds \\ &= \binom{x+1}{2} \left(e^{-at} \right)^x \left(1 - e^{at} \right)^2 \end{aligned}$$

By induction, we get

$$P_{xy}(t) = \binom{y-1}{y-x} \left(e^{-at} \right)^x \left(1 - e^{-at} \right)^{y-x}.$$

This is negative binomial!

5.2 Recurrent and transient states

Definition 5.5 (Hitting time)**.** *Let $\{x_T, t \geq 0\}$ be continuous time Markov Chain with S. For any $y \in S$, let τ_1 denote the waiting time before the first jump, and set*

$$T_y = \inf\{t \geq \tau_1 : X_t = y\}.$$

From now on, we denote $\rho_{xy} = P(T_y < \infty | X_0 = x)$.

Definition 5.6 (Transient and recurrent states)**.** *A state x is transient iff $\rho_{xx} < 1$. x is recurrent iff $\rho_{xx} = 1$.*

Theorem 5.3. *Let $\tau_1, \tau_2, \dots$ be the jump times of $\{X_t, t \geq 0\}$. Set $y_n = X_{\tau_n}$. Then, $\{y_n, n = 0, 1, \dots\}$ is a discrete time Markov Chain with one-step transition probability matrix $\mathbb{P} = Q$, where*

$$Q_{xy} = \begin{cases} 0, & \text{if } y = x \\ \frac{q_{xy}}{q_x}, & \text{if } y \neq x \end{cases}$$

$\{y_n, n \in \mathbb{N}\}$ is called the embedded Markov Chain *of $\{X_t, t \geq 0\}$.*

Theorem 5.4. *A state is recurrent or transient under $\{X_t, t \geq 0\}$ iff x is recurrent or transient under $\{y_n, n \in \mathbb{N}\}$.*

Remark. If S is finite, then $S = C_R \cup C_T$ and at least one recurrent class exists.

Definition 5.7 (Stationary Distribution)**.** *$\pi(x)$ is a stationary distribution of a Continuous Time Markov Chain iff $\pi\mathbb{P}(t) = \pi$ for every t.*

Equivalently, by differentiating, we obtain the following expression:

$$\pi\mathbb{P}'(0) = 0.$$

Theorem 5.5. *A distribution π is stationary iff $\pi A = 0$ or $\sum\limits_{x \in S} \pi(x) q_{xy} = 0$ for all $y \in S$.*

Example 5.2.1. Consider a continuous time Markov Chain, $\{X_t, t > 0\}$ with $S = \{1, 2\}$. Let

$$A = \begin{pmatrix} -1 & 1 \\ -10 & 10 \end{pmatrix}$$

Then, the transition matrix can be found by solving the following system of linear equations:

$$(\pi(1), \pi(2)) \begin{pmatrix} -1 & 1 \\ 10 & -10 \end{pmatrix} = 0$$

So we get

$$\pi = (10/11, 1/11)$$

Definition 5.8 (Mean return time)**.** $m_x = \sum_x (T_x)$.

Definition 5.9 (Positive recurrent)**.** *A state x is positive recurrent iff $m_x < \infty$.*

Remark. The positive recurrenct set under $\{X_t, t \geq 0\}$ is different from the positive recurrent set under $\{y_m, m \geq 0\}$.

Theorem 5.6. *Stationary distribution is concentrated on positive recurrent states only.*

Theorem 5.7. *Absorbing state is positive reccurent. If x is non-absorbing and positive recurrent, then the stationary distribution on the irreducible closed set containing x is*

$$\pi(x) = \frac{1}{q_x m_x}$$

5.3 Continuous time Birth-Death Markov Chain

Clearly, pure birth Markov Chain is transient. On the other hand, pure death is not actually recurrent. In this case, we have to write $S = \{0\} \cup \{1, 2, 3, \dots\}$, and $\{0\}$ is the only irreducible and recurrent set.

Now, consider a case where $\lambda_x > 0, \mu_x > 0$. The chain is irreducible. To understand the behaviour, we have to study the embedded chain, $\{Y_n = X_{\tau_n}\}$, whose transition matrix is given by

$$\mathbb{P} = \mathbb{Q} = \begin{pmatrix} 0 & 1 & 0 & \dots \\ \frac{\mu_1}{\mu_1+\lambda_1} & 0 & \frac{\mu_2}{\mu_2+\lambda_2} & 0 \\ & & & \ddots \end{pmatrix}$$

Recall that a birth-death Markov Chain is transient iff $\sum_x \Gamma_x < \infty$. So in this case, the chain is transient iff

$$\begin{aligned} &\sum_{n=1}^{\infty} \frac{\frac{\mu_1}{(\lambda_1+\mu_1)} \frac{\mu_2}{(\lambda_2+\mu_2)} \frac{\mu_3}{(\lambda_3+\mu_3)} \cdots \frac{\mu_n}{(\lambda_n+\mu_n)}}{\frac{\lambda_1}{(\lambda_1+\mu_1)} \frac{\lambda_2}{(\lambda_2+\mu_2)} \frac{\lambda_3}{(\lambda_3+\mu_3)} \cdots \frac{\lambda_n}{(\lambda_n+\mu_n)}} \\ =&\sum_{n=1}^{\infty} \frac{\mu_1 \mu_2 \cdots \mu_n}{\lambda_1 \lambda_2 \cdots \lambda_n} < \infty \end{aligned}$$

By checking at this criteria, we can test whether the chain is recurrent or not.

Now, let's look at the stationary distribution. Recall that stationary distribution, π, is given by solving $\pi A = 0$. Then, we get the following set of linear equations:

$$\begin{aligned} -\lambda_0 \pi(0) + \mu_1 \pi(1) &= 0 \\ \lambda_0 \pi(0) - (\lambda_1 + \mu_1)\pi(1) + \mu_2 \pi(2) &= 0 \\ \lambda_1 \pi(1) - (\lambda_2 + \mu_2)\pi(2) + \mu_3 \pi(3) &= 0 \\ &\vdots \\ \lambda_{n-1} \pi(n1) - (\lambda_n + \mu_n)\pi(n) + \mu_{n+1} \pi(n+1) &= 0 \end{aligned}$$

Solving this recursively, we get

$$\pi(1) = \frac{\lambda_0}{\mu_1}\pi(0), \pi(2) = \frac{\lambda_0\lambda_1}{\mu_1\mu_2}\pi(0), \pi(3) = \frac{\lambda_0\lambda_1\lambda_2}{\mu_1\mu_2\mu_3}, \dots$$

In general, we have

$$\pi(n+1) = \frac{\lambda_0 \cdots \lambda_n}{\mu_0 \cdots \mu_{n+1}}\pi(0)$$

Then, we have

$$\sum_{n=0}^{\infty} \pi(n) = \pi(0)\left[1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0\lambda_1}{\mu_1\mu_2} + \cdots\right] = 1$$

So we can solve this iff the term within the bracket is finite. In other words, the chain is positive recurrent iff

$$\sum_{n=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} < \infty.$$

Example 5.3.1. Consider a continuous time Markov Chain with

$$A = \begin{pmatrix} -1 & 1 & 0 & \cdots & \\ 1 & -1 & 1 & 0 & \cdots \\ 0 & 1 & -2 & 1 & \cdots \\ & & & & \ddots \end{pmatrix}$$

Since $\sum_{n=1}^{\infty} \frac{\mu_1 \cdots \mu_n}{\lambda_1 \cdots \lambda_n} = \sum_{n=1}^{\infty} 1 = \infty$, the chain is recurrent. Likewise, we can check the second criteria to find out that this is null recurrent. Then, we get $\lim_{t\to\infty} P_{xy}(t) = 0$ for all state $x, y \in S$.

Example 5.3.2. Consider a continuous time birth-death Markov Chain with $\lambda_x = 1$ and $\mu_x = 1 + 1/x$. Then, we get a positive recurrent chain. We can prove this using telescoping.