Stochastic Processes

MATH4240 (2020/21 Term 2)

Instructor: Renjun DUAN

Course Web Page :
http://www.math.cuhk.edu.hk/course/2021/math4240

Outline I

Outline II

Chapter 0: Review on probability

§0.1: Probability

Perform an experiment:

An outcome: a particular **state** ω

Sample space: the set of all outcomes, Ω

An **event**: a subset of Ω, $e.g.$, $A \subseteq \Omega$

Examples:

1. Toss a coin.

$$\omega_1 = H, \omega_2 = T$$
$$\Omega = \{H, T\} = \{\omega_1, \omega_2\}$$
$$\text{all possible events: } A = \emptyset, \Omega, \{H\}, \{T\}$$

2. Toss 3 coins.

$$\omega_1 = (H, H, H), \omega_2 = ..., \cdots, \omega_8 = ...$$
$$\Omega = \{(H, H, H), (H, H, T), (H, T, H), (T, H, H), (H, T, T), (T, H, T), (T, T, H), (T, T, T)\}$$

Want an event $A \stackrel{\text{def}}{=}$ "exactly 2 heads occur". Then,

$$A = \{(H, H, T), (H, T, H), (T, H, H)\}$$

Probability measure P: a function that assigns real values in $[0, 1]$ to events, satisfying

(i) $P(\Omega) = 1$

(ii) $0 \leq P(A) \leq 1, \forall A$

(iii) $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i), \forall \{A_i\}_{i=1}^{n}$ which is disjoint (n finite or infinite)

Probability space $(\Omega, \mathcal{F}, P)$:

(i) $\mathcal{F}$ is an **event space**, i.e. a collection of events one is interested in, satisfying

(a) $\Omega \in \mathcal{F}$

(b) If $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$

(c) If $A_i \in \mathcal{F}, i = 1, 2, \ldots,$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

$\mathcal{F}$ is a σ-**field** over Ω in measure theoretical term.

(ii) $P : \mathcal{F} \to [0, 1]$ is a probability measure.

Examples:

1. Given Ω, the largest σ-field is the set of all subsets of Ω.

2. Given Ω, the smallest σ-field is $\mathcal{F} = \{\Phi, \Omega\}$.

Conditional probability: A, B are two events, the probability that B happens given that A occurs is

$$\boxed{P(B|A) = \frac{P(A \cap B)}{P(A)}}$$

Note:

- A, B are **independent** if

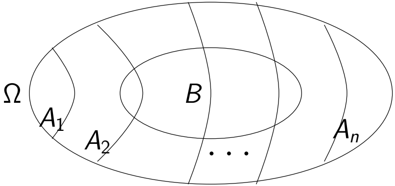
$$P(B|A) = P(B), \text{i.e. } P(A \cap B) = P(A)P(B).$$

- Let A be a fixed event,

$$P_A(\cdot) \overset{\text{def}}{=} P(\cdot|A)$$

  is called the **conditional probability measure**.

Theorem. Let $\Omega = \bigcup_{i=1}^{n} A_i$ where $A_1, \ldots, A_n$ are disjoint events. Then, for any event B



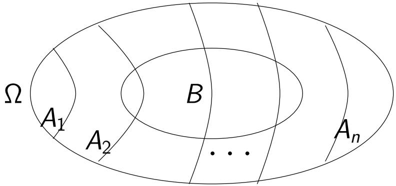
(i) $P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$

(ii) $P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$

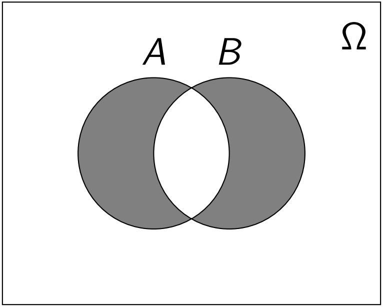
(Bayes' formula)

Note:

In many practical applications, we are given $P(B|A_i)$ and $P(A_i)$, and we want to find $P(A_i|B)$, i.e. to find the probability of the "causes" $A_i (i = 1, 2, \ldots, n)$ subject to the outcome B.



Example: $P(B) = P(B|A)P(A) + P(B|A^C)P(A^C)$
(B is caused by either A or A^c)



Proof: $B = (B \cap A) \cup (B \cap A^C)$

$$
\begin{aligned}
P(B) &= P(B \cap A) + P(B \cap A^C) \\
&= P(B|A)P(A) + P(B|A^C)P(A^C). \quad \square
\end{aligned}
$$

One more example:

Suppose that we have 3 cards that are identical in form, except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground.

If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?

Sol.: Denote

$\mathrm{RR} \overset{\mathrm{def}}{=}$ the event that the chosen card is red-red
$\mathrm{BB} \overset{\mathrm{def}}{=}$ the event that the chosen card is black-black
$\mathrm{RB} \overset{\mathrm{def}}{=}$ the event that the chosen card is red-black
$\mathrm{R} \overset{\mathrm{def}}{=}$ the event that the upper side of the chosen card is red

Then

$$\begin{aligned} P(\mathrm{RB}|\mathrm{R}) &= \frac{P(\mathrm{RB} \cap \mathrm{R})}{P(\mathrm{R})} \\ &= \frac{P(\mathrm{R}|\mathrm{RB})P(\mathrm{RB})}{P(\mathrm{R}|\mathrm{RR})P(\mathrm{RR}) + P(\mathrm{R}|\mathrm{BB})P(\mathrm{BB}) + P(\mathrm{R}|\mathrm{RB})P(\mathrm{RB})} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}} \\ &= \frac{1}{3}. \end{aligned}$$

□

§0.2 Random variables and distributions

Example: Toss a coin n-times.

$$\Omega = \{\omega = (\omega_1, \omega_2, \ldots, \omega_n) : \omega_i = H \text{ or } T\}$$

$\sharp$ of $\Omega = 2^n$

$$P(\{\omega\}) = \frac{1}{2^n}$$

Let X denote the number of heads,

then X takes values in $\{0, 1, 2, \ldots, n\}$,

Let $k = 0, 1, \ldots, n,$ then $X = k$ means

the event that we get k number of heads,

$$P(X = k) = \frac{\binom{n}{k}}{2^n}.$$

Random variable: A random variable (r.v.) X on $(\Omega, \mathcal{F}, P)$ is to assign an outcome with a real number

$$X : \Omega \to \mathbb{R}$$
$$\Omega \ni \omega \mapsto X(\omega) \in \mathbb{R}$$

Note: Let $R_X =$ the set of all possible values of X on Ω, then R_X is either "discrete" or "continuous":
Case 1: R_x is a discrete set. In this case, X is called a discrete r.v.
Case 2: R_x is an interval of $\mathbb{R}$ or itself. In this case, X is called a continuous r.v.

Discrete random variable: Assume

$$X(\Omega) = \{k\}_{k=0}^{N} \qquad (N \text{ finite or infinite})$$

Then the values

$$p_k = P(X = k), (k = 0, 1, \ldots, N)$$

is called the **probability density function** (p.d.f.).

Note: $\{X = k\} \stackrel{\text{def}}{=} \{\omega \in \Omega : X(\omega) = k\} \in \mathcal{F}$

Examples (Important!)

1. Binomial distribution

We perform n independent trials. At each trial, the prob of **success** is p,
and the prob of **failure** is $1 - p$.

Let X denote the *number of successes in n trials*. X has the p.d.f.

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, 0 \le k \le n.$$

$$\stackrel{\text{def}}{=} B(n, p)$$

2. Poisson distribution:

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \ldots$$

For instance, X counts the number of arrivals in a unit time with rate of arrivals given by $\lambda > 0$.

Theorem: For each $k = 0, 1, \cdots$

$$\lim_{n\to\infty, np=\lambda} \binom{n}{k} p^k (1-p)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$$

Note: Therefore, the Poisson distribution can be used to approximate the Binomial distribution when p is small and n is large compared to k.

3. Geometric distribution:

$$P(X = k) = p(1 - p)^{k-1}, k = 1, 2, \ldots$$

is the prob that the first occurrence of success requires k independent trials, each with success probability p.

X denotes the number of trials for the first success.

Continuous random variable:

If

$$P(a \le X \le b) = \int_a^b f(x)dx,$$

then f is called a **density function** of X.

Note:
the event "$a \le X \le b$" $\stackrel{\text{def}}{=} \{\omega \in \Omega : a \le X(\omega) \le b\}$

Examples (Important!)

1. Uniform distribution:

$$f(t) = \begin{cases} \frac{1}{b-a}, & a \leq t \leq b \\ 0, & \text{otherwise} \end{cases}$$

2. Exponential distribution:

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

3. Normal distribution:

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \stackrel{\text{def}}{=} N(\mu, \sigma^2)$$

$N(0, 1)$: standard normal density.

Exercise: Assume X, Y are two independent continuous (or discrete) r.v. with densities f, g (or $(p_k), (q_k)$).

Find the density function for the random variable $Z = X + Y$.

§0.3 Expectation and variance

The **expectation** (or **mean**) of X:

$$\mu = E(X) = \sum_k k p_k \text{ or } \int_{-\infty}^{\infty} t f(t) dt$$

The 2^{nd} **moment** of X:

$$E(X^2) = \sum_k k^2 p_k, \text{ or } \int_{-\infty}^{\infty} t^2 f(t) dt$$

The **variance** of X:

$$\sigma^2 \stackrel{\text{def}}{=} Var(X) = E(X-\mu)^2 = E(X^2) - \mu^2$$

(a measurement of how spread the distribution is)

Conditional expectation:

Discrete case: Suppose (X, Y) has a joint density

$$p(x_i, y_i) \stackrel{\text{def}}{=} P(X = x_i, Y = y_i)$$

$$E(Y|X = x_i) = \sum_j y_j P(Y = y_j | X = x_i)$$

$$= \sum_j y_j \frac{p(x_i, y_j)}{p(x_i)}, \;\; p(x_i) = \sum_j p(x_i, y_j)$$

Note:

a. $P(Y = y_j | X = x_i)$ is the conditioned density function of Y given $X = x_i$.

b. $E(Y|X = x_i)$ is a function of x_i, and thus regarded as a r.v. on the σ-field generated by X, denoted by $E(Y|X)$.

Continuous case: Let $f(x,y)$ be such that

$$P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\, dudv.$$

Then,

$$E(Y|X=x) = \int_{R_Y} y \frac{f(x,y)}{f(x)} dy,$$

$$f(x) = \int_{R_Y} f(x,y) dy.$$

Here $E(Y|X)$ can be understood to be a r.v. on the σ-field generated by X.

§0.4 Sequence of random variables

Repeat a random experiment independently. We obtain a sequence of random variables which are **independent and identically distributed** (i.i.d)

$$\{X_n\}_{n=0}^{\infty}.$$

Two basic theorems are:

- Law of Large Number
- Central Limit Theorem

(Ref: Ross p.389)

However, in many cases $\{X_n\}_{n=0}^{\infty}$ **may not be independent**. There exists dependence in a certain way.

In general, we call

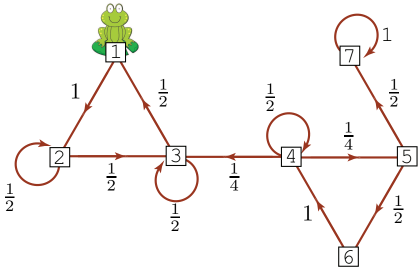
- $\{X_n\}_{n=0}^{\infty}$ a discrete time stochastic process, and
- $\{X_t\}_{t\geq 0}$ a continuous time stochastic process.

We will mainly consider the

"**Markov**" processes

(to be defined) in the discrete time and continuous time.

Example 1.1. A frog hops about on 7 lily pads. The numbers next to arrows show the probabilities with which, at the next jump, he jumps to a neighbouring lily pad (and when out-going probabilities sum to less than 1 he stays where he is with the remaining probability).



$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

There are 7 'states' (lily pads). In the matrix P the element P_{57} $(= 1/2)$ is the prob that, when starting in state 5, the next jump takes the frog t state 7.

Some questions we may want to know:

1. Starting in state 1, what is the prob that we are still in state 1 after 3 steps? after 5 steps? or after 1000 steps?
2. Starting in state 4, what is the prob that we ever reach state 7?
3. Starting in state 4, how long on average does it take to reach either 3 or 7?
4. Starting in state 2, what is the long-run proportion of time spent in state 3?

We can answer those by the end of this course

——End of Chapter 0——

Chapter 1: Markov Chain

§1.1: Definition & Examples

Example:

- Consider the weather (0=Sunny, 1=Rainy, 2=Cloundy) of days in Hong Kong.
- Let X_0 be a r.v. describing the weather of the 0th day, then
$$X_0 = 0,\ 1,\ \text{or}\ 2,$$
i.e. X_0 takes values in
$$S := \{0, 1, 2\}.$$
- Similarly, for $n \geq 0$ let X_n be a r.v. describing the weather of the nth day, then $X_n = 0,\ 1,\ \text{or}\ 2,$ i.e. X_n takes values in the same state space S.
- In the end we get a chain $\{X_n\}_{n\geq 0}$.

Definitions:

- Let S be a finite or countably infinite set of integers.

  For instance, $S = \{0, 1, 2, \ldots, N\}$, or $S = \{0, 1, 2, \ldots\}$, or $S = \{\ldots, -1, 0, 1, \ldots\}$.

  We call each element of S a **state** and S the **state space**.

- Let $\{X_n\}_{n=0}^{\infty}$ be a sequence of r.v. taking values in S, defined on a *common* probability space $(\Omega, \mathcal{F}, P)$.

Notation for the future:

• For random variables, we use

$$X, Y, Z, \cdots$$

• For states (which are values of random variables), we use

$$x, y, z, \cdots \in S$$

or

$$x_i, y_i, z_i, \cdots \in S,$$

or

$$i, j, k, \cdots \in S.$$

Def: $\{X_n\}_{n=0}^{\infty}$ is a **Markov chain** if

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n)$$
$$= P(X_{n+1} = x_{n+1} | X_n = x_n). \qquad (*)$$

Note:

- $(*)$ is called the **Markov property** which says that given the *present* state, the *past* states have no influence on the *future*!

- $P(X_{n+1} = y | X_n = x)$ is called the **transition probability**. If it is independent of n, we denote

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

which is the transition probability from state x to state y. In such case, $\{X_n\}_{n=0}^{\infty}$ is called a *time-homogeneous* Markov chain.

It is clear that

(i) $P(x, y) \geq 0$.

(ii) $\sum_{y \in S} P(x, y) = 1$.

Proof:

(i) $P(x, y) = P(X_{n+1} = y | X_n = x) \geq 0$.

(ii) $\sum_{y \in S} P(x, y) = \sum_{y \in S} P(X_{n+1} = y | X_n = x) = 1$.

e.g. for $S = \{0, 1, 2, \ldots, N\}$ (N finite or ∞), we may express all the transition probabilities

$$P(x, y), \quad x, y \in S$$

as a *matrix form*:

$$P = [P(x, y)] \text{ (or } [P(i, j)])$$
$$= \begin{bmatrix} P(0,0) & P(0,1) & \cdots & P(0,N) \\ P(1,0) & P(1,1) & \cdots & P(1,N) \\ \vdots & \vdots & \ddots & \vdots \\ P(N,0) & P(N,1) & \cdots & P(N,N) \end{bmatrix}$$

which is called the **Markov matrix** (or transition matrix) (Note: each row vector is a probability vector).

Example 1. Toss a possibly biased coin repeatedly with prob p for H and $1 - p$ for T.

Q.: Set up the model as a Markov chain.

Example 2. Consider a machine that at the start of the day is broken down or in operation. Assume

(i) if it is broken down, the prob that it will be repaired and in operation on the next day is p, $(0 < p < 1)$.

(ii) if it is in operation, the prob that it will be broken down on the next day is q, $(0 < q < 1)$.

Q: Set up the model as a Markov chain. Further,

(a) Find the transition prob.

(b) Find the prob that the machine is broken down on the n^{th} day.

(c) In the long term, what is the prob that the machine is broken down on a day.

Example 3 (Random walk):

Let $\{\xi_i\}_{i=1}^{\infty}$ be i.i.d. r.v. taking values in

$$S = \{\cdots, -1, 0, 1, \cdots\}$$

and having a pdf f, i.e. for each i

$$P(\xi_i = k) = f(k), \quad k = 0, \pm 1, \pm 2, \cdots$$

Let $X_n = X_0 + \xi_1 + \cdots + \xi_n$, where X_0 is the initial position independent of $\{\xi_i\}_{i=1}^{\infty}$. Then,

$$\begin{aligned} P(x, y) &= P(X_{n+1} = y | X_n = x) \\ &= P(\xi_{n+1} = y - x | X_n = x) \\ &= P(\xi_{n+1} = y - x) \\ &= f(y - x). \end{aligned}$$

A simple random walk:

Consider a move to left or right with prob $p, 1-p$ resp, i.e. $\xi_i = +1$ or -1 with prob $p, 1-p$ resp.

How does the chain behave as $n \to \infty$?

Example 4 (Gambler's ruin chain)

A gambler starts out with a certain amount and bets against the house.

(i) Each time he wins or loses \$1 with prob p and $q = 1 - p$ resp.

(ii) If he reaches \$0, he is ruined and his amount remain \$0. (he quits playing)

Q.: Set up the model as a Markov chain.

Let X_n denote the amount he has at the n-th stage.
$S = \{0, 1, 2, \cdots\}$.

• For $x = 0$,
$P(0,0) = 1$,
$P(0,y) = 0,\ y = 1, 2, \cdots$

<u>Def:</u> A state $a \in S$ is absorbing if $P(a,a) = 1$,
i.e. $P(a,y) = 0, \forall\, y \neq a$.

$\therefore$ 0 is an absorbing state.

• For $x > 0$,

$$P(x,y) = \begin{cases} p & y = x+1 \\ 1-p & y = x-1 \\ 0 & \textit{otherwise} \end{cases}$$

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

A modification of the model: Add a rule

(iii) If he reaches \$N, he quits playing.

Then,

$S = \{0, 1, \cdots, N\}$.

0 and N are absorbing,

$$P(x,y) = \begin{cases} p & y = x+1 \\ 1-p & y = x-1 \\ 0 & \textit{otherwise} \end{cases} \quad \text{for } 1 \le x \le N-1.$$

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

Alternative view to the above modified "gambler's ruin chain": Two gamblers start a series of \$1 bets against each other.

(i) The total amount is $\$N$.

(ii) $p =$ prob of the 1^{st} gambler wining
$q = 1 - p =$ prob of the 2^{nd} gambler winning.

(iii) The game is over when one of them losses all.

$X_n \stackrel{\text{def}}{=}$ \$ of the 1^{st} gambler at the n^{th} stage

Q:

- What is the expected value?
- Wo has higher prob of winning?
- How long does the game last?

Remark: The more general form of the chains in examples 3 & 4:

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

which corresponds to the "**birth** & **death**" chain. Here

$$p_x, q_x, r_x \geq 0,$$
$$p_x + q_x + r_x = 1.$$

Example 5 (Queueing chain)

Consider a check out counter at a supermarket.

(i) Let ξ_n denote the number of arrivals in the n^{th} period (say, one minute). Then $\{\xi_n\}_{n=1}^{\infty}$ is i.i.d. r.v. having pdf f (usually Poisson distribution).

(ii) Suppose that if there are any customers waiting for service at the beginning of any given period, then exactly one customer will be served during that period.

Q.: Set up the model as a Markov chain.

- $n = 0$:

$X_0 \overset{\textbf{def}}{=}$ the number of persons on the line initially.

- $n \geq 1$:

$X_n \overset{\textbf{def}}{=}$ the number of persons on the line present at the end of the n^{th} period.

- Then,

$$X_{n+1} = \begin{cases} 0 + \xi_{n+1} & \textit{if} \quad X_n = 0 \\ X_n + \xi_{n+1} - 1 & \textit{if} \quad X_n \geq 1 \end{cases}$$

- For $x = 0$,

$$\begin{aligned} P(0, y) &= P(X_{n+1} = y | X_n = 0) \\ &= P(\xi_{n+1} = y | X_n = 0) \\ &= P(\xi_{n+1} = y) \\ &= f(y). \end{aligned}$$

- For $x \geq 1$,

$$\begin{aligned} P(x, y) &= P(X_{n+1} = y | X_n = x) \\ &= P(\xi_{n+1} = y - x + 1 | X_n = x) \\ &= P(\xi_{n+1} = y - x + 1) \\ &= f(y - x + 1). \end{aligned}$$

For instance, f is Poisson:

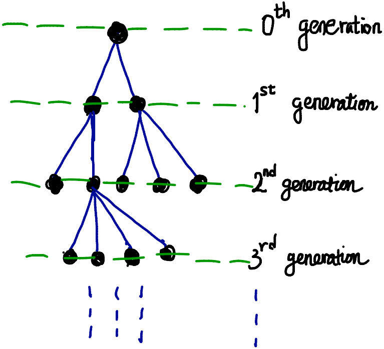
$$f(k) = P(\xi_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \cdots$$

Then

$$P = e^{-\lambda}\begin{bmatrix} 1 & \lambda & \frac{\lambda^2}{2!} & \frac{\lambda^3}{3!} & \cdots \\ 1 & \lambda & \frac{\lambda^2}{2!} & \frac{\lambda^3}{3!} & \cdots \\ 0 & 1 & \lambda & \frac{\lambda^2}{2!} & \cdots \\ \cdots & \ddots & \ddots & \ddots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

Example 6 (Branching chain, population growth)

Each individual generates ξ offspring in the next generation independently.



$X_n \stackrel{\text{def}}{=}$ the total NO in the n^{th} generation.

$$P(x, y) = P(\xi_1 + \xi_2 + \cdots + \xi_x = y)$$

Question concerns the extinction or growth of the population!

§1.2 Computations with transition probabilities

Setup:

- $\{X_n\}_{n=0}^{\infty}$: a time-homogeneous Markov chain
- $S = \{0, 1, 2, \ldots, N\}$: state space
  (N: finite or ∞)
- $P = [P(x, y)] = [P(X_{n+1} = y | X_n = x)]$:
  transition prob matrix

Question 1: Given pdf of X_0, can one compute pdf of X_n for any $n \geq 1$?

Let the pdf of X_0 be

$$\pi_k^{(0)} \stackrel{\text{def}}{=} P(X_0 = k), \quad k = 0, 1, \cdots, N,$$

or equivalently we write in the prob row-vector form

$$\pi^{(0)} = [\pi_0^{(0)}, \pi_1^{(0)}, \cdots, \pi_N^{(0)}].$$

- $n = 1$: $P(X_1 = k), k \in S$? or $\pi^{(1)} = [\pi_0^{(1)}, \cdots, \pi_N^{(1)}]$?

$$P(X_1 = k) = \sum_{i \in S} P(X_1 = k, X_0 = i)$$
$$= \sum_{i \in S} P(X_1 = k | X_0 = i) P(X_0 = i)$$
$$= [P(X_0 = 0), P(X_0 = 1), \ldots, P(X_0 = N)] \begin{bmatrix} P(0, k) \\ P(1, k) \\ \vdots \\ P(N, k) \end{bmatrix}$$

Write them for $k = 0, 1, \cdots, N$ in matrix:

$$[P(X_1 = 0), P(X_1 = 1), \ldots, P(X_1 = N)]$$
$$= [P(X_0 = 0), P(X_0 = 1), \ldots, P(X_0 = N)] \begin{bmatrix} P(0,0) & P(0,1) & \cdots & P(0,N) \\ P(1,0) & P(1,1) & \cdots & P(1,N) \\ \vdots & \vdots & \ddots & \vdots \\ P(N,0) & P(N,1) & \cdots & P(N,N) \end{bmatrix}$$

i.e.

$$\boxed{\pi^{(1)} = \pi^{(0)} P}$$

• In general, for $n \geq 1$, setting the pdf of X_n as a probability row-vector in the form

$$\pi^{(n)} = [P(X_n = 0), P(X_n = 1), \cdots, P(X_n = N)],$$

Then,

$$\boxed{\pi^{(n)} = \pi^{(n-1)} P}.$$

- Then, by iteration,

$$\begin{aligned}\pi^{(n)} &= \pi^{(n-1)}P \\ &= \pi^{(n-2)}P \cdot P \\ &= \cdots \\ &= \pi^{(0)} \underbrace{P \cdot P \cdot \ldots \cdot P}_{\text{n terms}} \\ &= \pi^{(0)}P^n\end{aligned}$$

where

$$P^n = \underbrace{P \cdot P \cdot \ldots \cdot P}_{\text{product of n terms}}$$

Theorem: $\pi^{(n)} = \pi^{(0)}P^n$, $n = 1, 2, \cdots$

Remark: How to compute the matrix product

$$P^n := \underbrace{P \cdot P \cdot \cdots \cdot P}_{n \text{ terms}}, \quad n = 2, 3, \cdots$$

Indeed, for $x, y \in S$,

$$P^2(x,y) = \sum_{x_1 \in S} P(x,x_1)P(x_1,y)$$

$$P^3(x,y) = \sum_{x_1} \sum_{x_2} P(x,x_1)P(x_1,x_2)P(x_2,y)$$

$$\cdots$$

$$P^n(x,y) = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_{n-1}} P(x,x_1)P(x_1,x_2) \cdots P(x_{n-1},y).$$

Proof: Left for an exercise. Argument: use induction in n and the formula $P^n = P^{n-1} \cdot P$.

Proposition:

(i) $P(X_n = y) = \sum_x \pi^{(0)}(x) P^n(x, y).$

(ii) $P(X_n = y | X_0 = x) = P^n(x, y).$

Proof: (i) is a direct consequence of the formula $\pi^{(n)} = \pi^{(0)} P^n$.
To show (ii),

$$\begin{aligned}
& P(X_n = y | X_0 = x) \\
& = P(X_n = y, X_{n-1} \in S, \cdots, X_1 \in S | X_0 = x) \\
& = \sum_{x_1 \in S} \cdots \sum_{x_{n-1} \in S} P(X_n = y, X_{n-1} = x_{n-1}, \cdots, X_1 = x_1 | X_0 = x) \text{(tutorial)} \\
& = \sum_{x_1, \cdots, x_{n-1} \in S} \frac{P(X_n = y, X_{n-1} = x_{n-1}, \cdots, X_1 = x_1, X_0 = x)}{P(X_0 = x)} \\
& = \sum_{x_1, \cdots, x_{n-1} \in S} \frac{P(X_0 = x) P(x_0, x_1) \cdots P(x_{n-1}, y)}{P(X_0 = x)} \text{(proof later)} \\
& = \sum_{x_1, \cdots, x_{n-1} \in S} P(x, x_1) \cdots P(x_{n-1}, y) = P^n(x, y). \quad \square
\end{aligned}$$

Claim:

$$P(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n)$$
$$= P(X_0 = x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n).$$

Proof of claim:

$$P(\underbrace{X_0 = x_0, X_1 = x_1, \cdots}_{A}, \underbrace{X_n = x_n}_{B})$$
$$= P(X_n = x_n | X_0 = x_0, \cdots, X_{n-1} = x_{n-1})$$
$$\cdot P(X_0 = x_0, \cdots, X_{n-1} = x_{n-1})$$
$$= P(X_n = x_n | X_{n-1} = x_{n-1}) P(X_0 = x_0, X_1 = x_1, \cdots, X_{n-1} = x_{n-1})$$
$$= P(x_{n-1}, x_n) P(X_0 = x_0, X_1 = x_1, \cdots, X_{n-1} = x_{n-1})$$
$$= \cdots$$
$$= P(x_{n-1}, x_n) P(x_{n-2}, x_{n-1}) \cdots P(x_0, x_1) P(X_0 = x_0). \quad \square$$

Remark: (ii) also immediately implies (i). In fact,

$$P(X_n = y) \overset{\text{by def}}{=} \sum_x P(X_n = y | X_0 = x) P(X_0 = x)$$

$$\overset{\text{by (ii)}}{=} \sum_x P^n(x, y) \pi^{(0)}(x).$$

Definition:

$$P^m(x,y), (m = 0, 1, \cdots)$$

is called the ***m*-step transition function**, which gives the prob of going from state x to state y in m steps. Here we set

$$P^0(x,y) = \delta_{xy} = \begin{cases} 1, & \textit{if} \quad x = y \\ 0, & \textit{otherwise}. \end{cases}$$

Correspondingly, P^m is called the ***m*-step transition matrix**.

Proposition:

$$P(X_{n+m} = y | X_n = x) = P^m(x, y).$$

Proof

$$\begin{aligned}
&P(X_{n+m} = y | X_n = x) \\
&= P(X_{n+m} = y, X_{n+m-1} \in S, \cdots, X_{n+1} \in S | X_n = x) \\
&= \sum_{x_{n+m-1}} \cdots \sum_{x_{n+1}} P(X_{n+m} = y, X_{n+m-1} = x_{n+m-1}, \cdots, X_{n+1} = x_{n+1} | X_n = x) \\
&= \sum_{x_{n+1}, \cdots, x_{n+m-1}} P(X_{n+m} = y, \cdots, X_{n+1} = x_{n+1} | X_0 \in S, \cdots, X_{n-1} \in S, X_n = x) \ (*) \\
&= \sum_{x_{n+1}, \cdots, x_{n+m-1}} P(x, x_{n+1}) \cdots P(x_{n+m-1}, y) \quad \text{(see below)} \\
&= P^m(x, y). \quad \text{(by def of } P^m\text{)}
\end{aligned}$$

Each term in the sum $(*)$ is equal to

$$\begin{aligned}
&P(X_{n+m} = y, \cdots, X_{n+1} = x_{n+1} | X_0 = x_0, \cdots, X_{n-1} = x_{n-1}, X_n = x) \\
&= P(X_{n+m} = y, X_{n+m-1}, \cdots, X_0 = x_0) / P(X_n = x, X_{n-1} = x_{n-1}, \cdots, X_0 = x_0) \\
&= \frac{P(X_0 = x_0) P(x_0, x_1) \cdots P(x, x_{n+1}) P(x_{n+1}, x_{n+2}) \cdots P(X_{n+m-1}, y)}{P(X_0 = x_0) P(x_0, x_1) \cdots P(x_{n-1}, x)} \\
&= P(x, x_{n+1}) \cdots P(x_{n+m-1}, y). \quad \square
\end{aligned}$$

Remark: To compute

$$\begin{aligned}
&P(X_n = y | X_0 = x) \\
&= P(X_{1+n} = y | X_1 = x) \\
&= P(X_{2+n} = y | X_2 = x) \\
&= \cdots \\
&= P(X_{m+n} = y | X_m = x), \quad m = 0, 1, 2, \cdots
\end{aligned}$$

is equivalent to compute $P^n(x, y)$, that is to find P^n.

For n large, one can reduce P to a diagonal matrix (if possible)

$$P = QDQ^{-1}$$

where $D = \begin{bmatrix} \lambda_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_N \end{bmatrix}$, and Q is the matrix for change basis consisting of eigenvectors. Then

$$P^n = \left[QDQ^{-1}\right]^n = QD^nQ^{-1} = Q \begin{bmatrix} \lambda_0^n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_N^n \end{bmatrix} Q^{-1}.$$

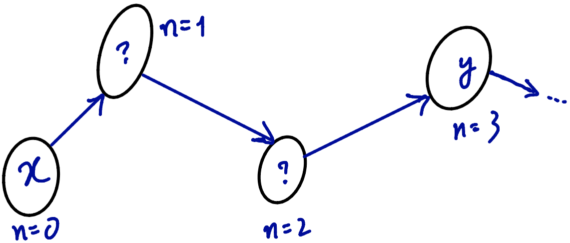
Hence P^n can be calculated in such situation.

Exercise: Do this for the two-state Markov matrix.

Question 2: How to compute the conditional prob that the chain visits y in **finite** time given that it starts from x?

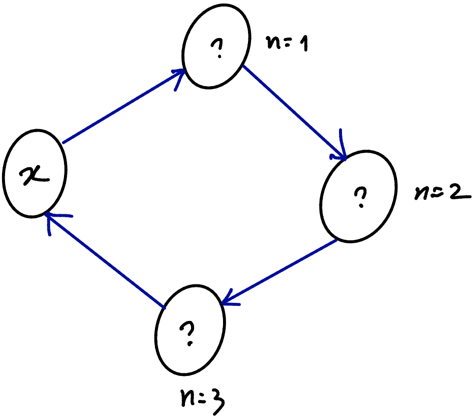
We set it as ρ_{xy}, then

$$\rho_{xy} = P(\exists\, n \geq 1 \text{ such that } X_n = y | X_0 = x).$$



We are interested in a state x such that

$$\rho_{xx} = 1, \text{ or } \rho_{xx} < 1.$$



Def (Hitting Time):
Let $A \subseteq S$. The **hitting time** T_A of A is defined by

$$T_A \stackrel{\text{def}}{=} min\{n \geq 1 : X_n \in A\}.$$

Rks:

- $T_A =$ the first positive time the chain hits A.
  T_A is a r.v. Range of $T_A = \{1, 2, 3, \cdots\} \cup \{\infty\}$.
  Convention: $T_A = \infty$ if $X_n \notin A$ for all $n \geq 1$.
  For $m = 1, 2, \cdots$
  $$\{T_A = m\} = \{X_1 \notin A, \cdots, X_{m-1} \notin A, X_m \in A\}.$$

- Convention:
  $$T_y \stackrel{\text{def}}{=} T_{\{y\}} = min\{n \geq 1 : X_n = y\}, \quad y \in S,$$
  i.e. the first positive time the chain visits y.

A convenient notion:

$$P_x(\cdot) \stackrel{\text{def}}{=} P(\;\cdot\;|X_0 = x)$$

i.e. the probabilities of various events defined in terms of the Markov chain starting at $x \in S$.

For instance,

$$P_x(A) = P(A|X_0 = x)$$

is the prob of A given that the chain starts at x.

Prop (i) $P_x(T_y = 1) = P(x, y)$.

Proof:

$\because \{T_y = 1\} = \{X_1 = y\}$

$\therefore$

$$\begin{aligned} & P_x(T_y = 1) \\ & = P(X_1 = y | X_0 = x) \\ & = P(x, y). \quad \square \end{aligned}$$

Prop (ii)

$$P_x(T_y = n+1) = \sum_{z \neq y} P(x,z) P_z(T_y = n), n \geq 1.$$

Proof: Note

$$\{T_y = n+1\} = \bigcup_{z: z \neq y} \{X_1 = z, X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y\}.$$

$\therefore$

$$\begin{aligned}
& P_x(T_y = n+1) \\
&= \sum_{z \neq y} P_x(X_1 = z, X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y) \\
&= \sum_{z \neq y} P_x(X_1 = z) P_x(X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y | X_1 = z) \\
&= \sum_{z \neq y} \underbrace{P(X_1 = z | X_0 = x)}_{= P(x,z)} \underbrace{P(X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y | X_0 = x, X_1 = z)}_{= P_z(T_y = n) \text{ why?}}
\end{aligned}$$

Recall an **Exercise:**

$$P(X_{n+1} \in B_1, \cdots, X_{n+m} \in B_m | X_0 \in A_0, \cdots, X_{n-1} \in A_{n-1}, X_n = x)$$
$$= P_x(X_1 \in B_1, X_2 \in B_2, \cdots, X_m \in B_m).$$

See the tutorial for the proof.

Rk: It is essentially due to the Markovian property (i.e., given "the present" state, "the past" has no influence on "the future"). So, the prob on the LHS is understood to be the prob in the situation when the chain initially starts at x.

$$\therefore \ P(X_2 \neq y, \cdots, X_n \neq y, X_{n+1} = y | X_0 = x, X_1 = z)$$
$$= P_z(X_1 \neq y, X_2 \neq y, \cdots, X_{n-1} \neq y, X_n = y)$$
$$= P_z(T_y = n). \quad \square$$

Prop (iii)

$$P^n(x,y) = \sum_{m=1}^{n} P_x(T_y = m)P^{n-m}(y,y).$$

Proof: Note: $P^n(x,y) = P(X_n = y | X_0 = x) = P_x(X_n = y)$

$\{X_n = y\} = \bigcup_{m=1}^{n} \{T_y = m, X_n = y\}$ (disjoint union)

$$\begin{aligned}
\therefore P^n(x,y) &= P_x(X_n = y) \\
&= \sum_{m=1}^{n} P_x(T_y = m, X_n = y) \\
&= \sum_{m=1}^{n} P_x(T_y = m) P_x(X_n = y | T_y = m) \\
&= \sum_{m=1}^{n} P_x(T_y = m) P(X_n = y | X_0 = x, X_1 \neq y, \cdots, X_{m-1} \neq y, X_m = y) \\
&= \sum_{m=1}^{n} P_x(T_y = m) P^{n-m}(y,y). \quad \square
\end{aligned}$$

Sum:

Proposition:

(i) $P_x(T_y = 1) = P(x, y)$.

(ii) $P_x(T_y = n+1) = \sum_{z \neq y} P(x, z) P_z(T_y = n), \quad n \geq 1$.

(iii) $P^n(x, y) = \sum_{m=1}^{n} P_x(T_y = m) P^{n-m}(y, y)$.

Corollary: If $a \in S$ is absorbing , i.e. $P(a,a) = 1$, then for any $n \geq 1$, $P^n(x,a) = P_x(T_a \leq n)$.

Proof:

$$P^n(x,a) = \sum_{m=1}^{n} P_x(T_a = m) \underbrace{P^{n-m}(a,a)}_{=1 \text{ (to be shown later)}}$$

$$= \sum_{m=1}^{n} P_x(T_a = m)$$

$$= P_x(\cup_{m=1}^{n}\{T_a = m\})$$

$$= P_x(T_a \leq n). \quad \square$$

It remains to show: For any $n \geq 0$, $P^n(a, a) = 1$.

Indeed:

- $n = 0, 1$ is obvious.
- $n \geq 2$:

$$\begin{aligned} P^n(a, a) &= \sum_{x_1, \cdots, x_{n-1}} P(a, x_1)P(x_1, x_2) \cdots P(x_{n-1}, a) \\ &= \sum_{x_2, \cdots, x_{n-1}} P(a, x_2) \cdots P(x_{n-1}, a) \\ &= \cdots \\ &= \sum_{x_{n-1}} P(a, x_{n-1})P(x_{n-1}, a) \\ &= P(a, a) \\ &= 1. \quad \square \end{aligned}$$

Recall: $\rho_{xy} = P_x(T_y < \infty)$ is the prob that the chain starting at x will visit y at some positive time.

In particular,

$$\rho_{yy} = P_y(T_y < \infty)$$

is the prob that the chain starting at y will ever return to y.

Def.:

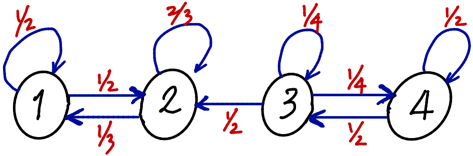
- A state y is called **recurrent** if $\rho_{yy} = 1$, and **transient** if $\rho_{yy} < 1$.
- A chain is called a recurrent (transient) chain if **all** states are recurrent (transient).

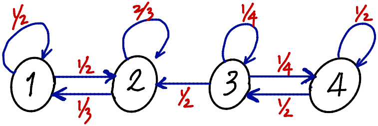
Rk: An absorbing state is recurrent.

Example:

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

Q: Find the matrix $[\rho_{xy}]$ from $P = [P(x, y)]$.





Observe:

(i) $0 = \rho_{13} = \rho_{14}, \quad 0 = \rho_{23} = \rho_{24}.$

(ii) $1 = \rho_{11} = \rho_{22}, \qquad \therefore$ 1, 2 are recurrent.

(iii) $\rho_{33} < 1, \rho_{44} < 1, \qquad \therefore$ 3, 4 are transient.

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

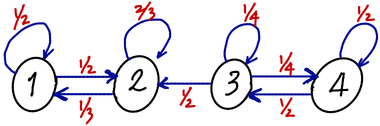
Recalling $\rho_{xy} = P_x(T_y < \infty)$, we have

$$\rho_{xy} = P(x, y) + \sum_{z:z \neq y} P(x, z)\rho_{zy}$$

(Exercise)

Argument: Start at x.

- If $T_y = 1$, i.e. visit y at $n = 1$, prob is $P(x, y)$.
- If it does not visit y at $n = 1$, then it will first visit $z(z \neq y)$ and then start from such z to visit y at some positive time.



$$\begin{cases} \rho_{33} = \cancelto{0}{0 \cdot \rho_{13}} + \cancelto{0}{\frac{1}{2} \cdot \rho_{23}} + \frac{1}{4} + \frac{1}{4} \cdot \rho_{43} \\ \rho_{43} = \cancelto{0}{0 \cdot \rho_{13}} + \cancelto{0}{0 \cdot \rho_{23}} + \frac{1}{2} + \frac{1}{2} \cdot \rho_{43} \end{cases}$$

$$\therefore \rho_{43} = 1, \quad \rho_{33} = \frac{1}{2}$$

Similarly,

$$\rho_{34} = \cancelto{0}{0 \cdot \rho_{14}} + \cancelto{0}{\frac{1}{2} \cdot \rho_{24}} + \frac{1}{4} \cdot \rho_{34} + \frac{1}{4}, \quad \therefore \rho_{34} = \frac{1}{3},$$

$$\rho_{44} = \cancelto{0}{0 \cdot \rho_{14}} + \cancelto{0}{0 \cdot \rho_{24}} + \frac{1}{2} \cdot \rho_{34} + \frac{1}{2} \quad \therefore \rho_{44} = \frac{2}{3}.$$

$$
[\rho_{ij}] = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}
\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\
\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & \frac{1}{2} & \frac{1}{3} \\ 1 & 1 & 1 & \frac{2}{3} \end{bmatrix} \end{array}.
$$

Note: There is a matrix argument for finding $[\rho_{xy}]$. See *Lawler p.23-27*.

Question 3. Times of visit to a state.

$\{X_n\}_{n=0}^{\infty}$: a time-homogeneous Markov chain

$S = \{0, \cdots, N\}$ (N: finite or ∞): state space

$X_0 = x \in S$

$$\boxed{N(y) \overset{\text{def}}{=} \text{no of times that } X_n (n \geq 1) \text{ visits } y.}$$

Note:

- $N(y) = \sum_{n=1}^{\infty} 1_y(X_n)$, where $1_y(X_n) = \begin{cases} 1, & X_n = y \\ 0, & X_n \neq y \end{cases}$
- $N(y) \in \{0, 1, 2, 3, \cdots\} \cup \{\infty\}$.

$$\{N(y) = 0\} = \text{"}y \text{ is not visited"}$$
$$\{N(y) = k\} = \text{"}y \text{ is visited exactly } k \text{ times"}$$
$$\{N(y) = \infty\} = \text{"}y \text{ is visited infinitely times"}$$

Some **Facts**:

- $$\underbrace{\{N(y) \geq 1\}}_{\text{"}y\text{ is visited at least one time"}} = \underbrace{\{T_y < \infty\}}_{\text{"}y\text{ is visited at a positive finite time"}}.$$

$$\therefore \boxed{P_x(N(y) \geq 1) = P_x(T_y < \infty) = \rho_{xy}}.$$

- $\{N(y) = 0\} = \{N(y) \geq 1\}^c$.

$$\therefore \boxed{P_x(N(y) = 0) = 1 - \rho_{xy}}.$$

Claim: For $m \geq 1$, $P_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1}$.

Case $m = 2$. To show: $P_x(N(y) \geq 2) = \rho_{xy}\rho_{yy}$.
Note:

$$\{N(y) \geq 2\} = \cup_{k\geq 1} \cup_{n \geq 1} \{\text{chain starting at } x \text{ first visits } y \text{ at } k \geq 1 \\ \text{and next visit } y \text{ again after } n \text{ units of time}\}.$$

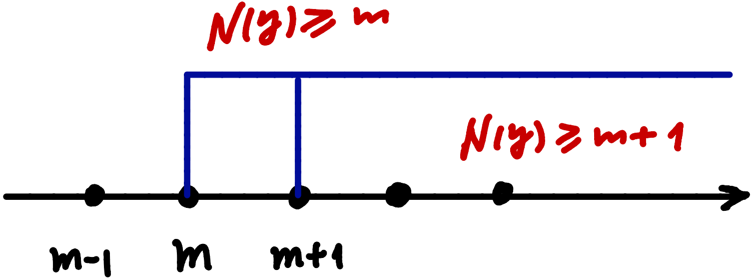
For each $k \geq 1$ and $n \geq 1$, prob $= P_x(T_y = k)P_y(T_y = n)$.
Therefore,

$$\begin{aligned} P_x(N(y) \geq 2) &= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} P_x(T_y = k)P_y(T_y = n) \\ &= \sum_{n=1}^{\infty} P_x(T_y < \infty)P_y(T_y = n) \\ &= \rho_{xy}P_y(T_y < \infty) \\ &= \rho_{xy}\rho_{yy}. \end{aligned}$$

Use the same idea to show $P_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1}$ for $m \geq 2$.

A further **fact:**

$$\{N(y) = m\} = \{N(y) \geq m\} \backslash \{N(y) \geq m+1\}$$



$$\therefore P_x(N(y) = m) = \rho_{xy}\rho_{yy}^{m-1} - \rho_{xy}\rho_{yy}^{(m+1)-1}$$
$$= \rho_{xy}\rho_{yy}^{m-1}(1 - \rho_{yy}).$$

Sum:

Proposition:

(i) $P_x(N(y) \geq 1) = P_x(T_y < \infty) = \rho_{xy}$,

$$P_x(N(y) = 0) = 1 - \rho_{xy}.$$

(ii) For $m \geq 1$,

$$P_x(N(y) \geq m) = \rho_{xy}\rho_{yy}^{m-1},$$

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

Proposition: $E_x(N(y)) = \sum_{n=1}^{\infty} P^n(x,y)$.

l.h.s.=the expected no of visit to y from x.

Warning: The value can be ∞!

Proof:

$$
\begin{aligned}
E_x(N(y)) &= E_x\Big(\sum_{n=1}^{\infty} 1_y(X_n)\Big) \\
&= \sum_{n=1}^{\infty} E_x(1_y(X_n)) \\
&= \sum_{n=1}^{\infty} P_x(X_n = y) \\
&= \sum_{n=1}^{\infty} P(X_n = y | X_0 = x) = \sum_{n=1}^{\infty} P^n(x,y). \quad \square
\end{aligned}
$$

Theorem (i): y is transient iff $P_y(N(y)=\infty)=0$.

Proof: Note

$$\begin{aligned} P_x(N(y)=\infty) &= \lim_{m\to\infty} P_x(N(y)\geqslant m) \\ &= \lim_{m\to\infty} \rho_{xy}\rho_{yy}^{m-1} \\ &= \begin{cases} 0 & \text{if } \rho_{yy}<1 \\ \rho_{xy} & \text{if } \rho_{yy}=1 \end{cases} \qquad (*) \end{aligned}$$

$\therefore$ y transient

$\iff \rho_{yy}<1$

$\overset{(*)}{\iff} P_y(N(y)=\infty)=0.$ □

Theorem (ii): If y is transient then

$$E_x(N(y)) = \frac{\rho_{xy}}{1-\rho_{yy}} < \infty, \quad x \in S.$$

Proof: For a transient state y,

$$\begin{aligned} E_x(N(y)) &= \sum_{m=0}^{\infty} m P_x(N(y)=m) \\ &= \sum_{m=1}^{\infty} m \rho_{xy} \rho_{yy}^{m-1}(1-\rho_{yy}) \quad (\rho_{yy} < 1) \\ &= \rho_{xy}(1-\rho_{yy}) \cdot \frac{1}{(1-\rho_{yy})^2} \\ &= \frac{\rho_{xy}}{1-\rho_{yy}} < \infty. \quad \square \end{aligned}$$

Theorem (iii):
y is recurrent,
iff $P_y(N(y) = \infty) = 1$,
iff $E_y(N(y)) = \infty$.

Proof: y recurrent

$$\iff \rho_{yy} = 1 \overset{(*)}{\iff} P_y(N(y) = \infty) = 1$$
$$\overset{(**)}{\iff} E_y(N(y)) = \infty.$$

To show $(**)$:

"$\Longrightarrow$": $\because P_y(N(y) = \infty) = 1$
$\therefore E_y(N(y)) = \infty$.

"$\Longleftarrow$": If $E_y(N(y)) = \infty$ then y must be recurrent by Theorem (ii). □

Remark: If y is recurrent, then for $x \in S$,

$$E_x(N(y)) = \begin{cases} 0 & \text{if } \rho_{xy} = 0 \\ \infty & \text{if } \rho_{xy} > 0. \end{cases}$$

WHY? It is heuristically obvious.

Left for an exercise.

Corollary: If S is **finite**, then the chain must have at least one recurrent state.

Proof: Otherwise, all states are transient. Then, for any $x \& y$,

$$\sum_{n=1}^{\infty} P^n(x,y) = E_x(N(y)) = \frac{\rho_{xy}}{1-\rho_{yy}} < \infty.$$

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

$$\begin{aligned} 0 &= \sum_{y\in S} \lim_{n\to\infty} P^n(x,y) \\ &= \lim_{n\to\infty} \sum_{y\in S} P^n(x,y) \qquad (S : \text{finite}) \\ &= \lim_{n\to\infty} P_x(X_n \in S) = \lim_{n\to\infty} 1 = 1. \qquad \square \end{aligned}$$

Question 4. Decomposition of state space.

Def: x **leads to** y (denoted by $x \to y$) if

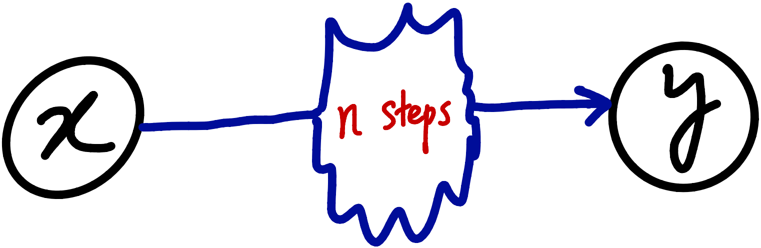
$$\rho_{xy} > 0.$$

Fact 1: $x \to y$ (i.e. $\rho_{xy} > 0$) **iff**

$$P^n(x, y) > 0 \quad \text{for some } n \geqslant 1.$$

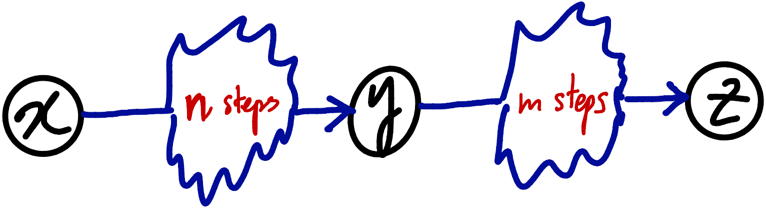
Proof: Note:

- $\rho_{xy} = P_x(T_y < \infty) = P_x(\{\exists\, m \geq 1\, s.t.\ X_m = y\})$.
- $P^n(x, y) = P(X_n = y | X_0 = x) = P_x(X_n = y)$.


x
n steps
y

Fact 2: $\left.\begin{array}{l} x \to y \\ y \to z \end{array}\right\} \Longrightarrow x \to z.$

Proof: Note

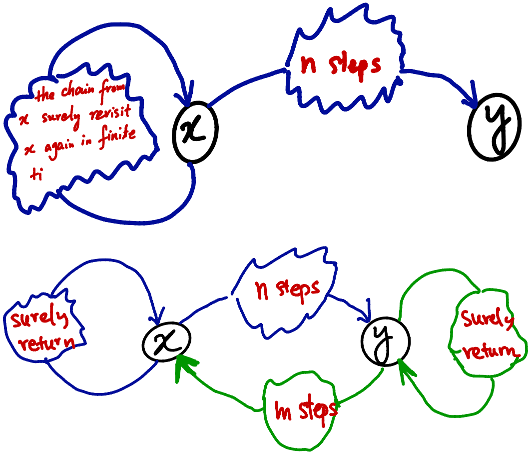


$$P^{n+m}(x,z) = \sum_{i \in S} P^n(x,i)P^m(i,z) \geqslant P^n(x,y)P^m(y,z) > 0.$$

Fact 3:

$$\left.\begin{array}{l} x \text{ recurrent } (\rho_{xx} = 1) \\ x \to y \end{array}\right\} \Longrightarrow \left\{\begin{array}{ll} \text{(i)} & y \to x \\ \text{(ii)} & y \text{ recurrent} \\ \text{(iii)} & \rho_{yx} = \rho_{xy} = 1 \end{array}\right.$$

Proof (Heuristic):



Def.:

(i) $C \subseteq S$ is **closed** if

$$\rho_{xy} = 0, \quad \forall\, x \in C,\ \forall\, y \notin C,$$

i.e. no state in C leads to any state out C.

(ii) A closed set C is **irreducible** if

$$x \to y \,(\text{i.e. } \rho_{xy} > 0), \quad \forall\, x \in C, \forall\, y \in C,$$

namely, any two in C can communicate with each other.

(iii) $\{X_n\}_{n=0}^{\infty}$ is an **irreducible MC** if its state space S is irreducible.

Remark (a): One can claim that

$$
\begin{aligned}
C \text{ is closed, i.e. } \rho_{xy} = 0, \forall x \in C, \forall y \notin C \quad & (1) \\
\iff P^n(x,y) = 0, \forall x \in C, \forall y \notin C, \forall n \geqslant 1 \quad & (2) \\
\iff P(x,y) = 0, \forall n \in C, \forall y \notin C. \quad & (3)
\end{aligned}
$$

- Direct to see: (1) $\iff$ (2) $\implies$ (3).
- To show (3) $\implies$ (2): For $x \in C$ & $y \notin C$,

$$
\begin{aligned}
P^2(x,y) &= \sum_{x_1 \in S} P(x,x_1) P(x_1,y) \\
&= \sum_{x_1 \in C} P(x,x_1) \underbrace{P(x_1,y)}_{0} + \sum_{x_1 \notin C} \underbrace{P(x,x_1)}_{0} P(x_1,y) \\
&= 0.
\end{aligned}
$$

Induction $\implies P^n(x,y) = 0, \ \forall n \geqslant 1.$

Remark (b): If

$$C \text{ is closed, } x \in C,\ P(x, y) > 0$$

then

$$y \in C.$$

Remark (c): If $C \subset S$ is closed, then

$$\{X_n\}_{n=0}^{\infty}$$

can also be regarded as a Markov Chain with the state space C.

Theorem: If C is an irreducible closed set, then **either**

all states in C are recurrent

or

all states in C are transient.

In particular, if C is a **finite** irreducible closed set, then all states in C must be recurrent.

Proof: Two cases in general:

(i) C does NOT contain any recurrent state. In the case, all states in C are transient.

(ii) C contains at least one recurrent state. As C is irreducible, all states in C are recurrent.

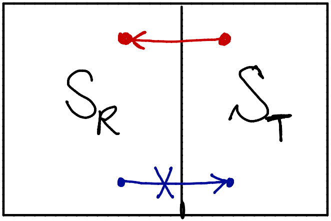
The particular case follows from the fact that any finite closed set must contain at least one recurrent state. □

Set

$$S_R = \{\text{recurrent states}\},$$
$$S_T = \{\text{transient states}\}.$$

Then,

$$S = S_R \cup S_T.$$



$$\therefore\ S_R \text{ is closed!}$$

A further question: Is S_R irreducible? namely, can any two recurrent states communicate to each other?

Observe: Assume $S_R \neq \phi$, for instance, $\exists\, x_0 \in S_R$. Define

$$C_{x_0} = \{x \in S_R : x_0 \to x\}.$$

Then, C_{x_0} must be **closed & irreducible**.

Proof:

(1) "C_{x_0} closed" $\iff$ "If $x \in C_{x_0}$ & $x \to y \in S$ then $y \in C_{x_0}$"
(Indeed, $y \in S_R, \therefore x_0 \to x \to y \in S_R$)

(2) "C_{x_0} irreducible" $\iff$ " If $x, y \in C_{x_0}$ then $x \to y$".

Indeed, $\left.\begin{matrix} x_0 \to x \in S_R \\ x_0 \to y \in S_R \end{matrix}\right\} \Longrightarrow x \to x_0 \to y.$

□

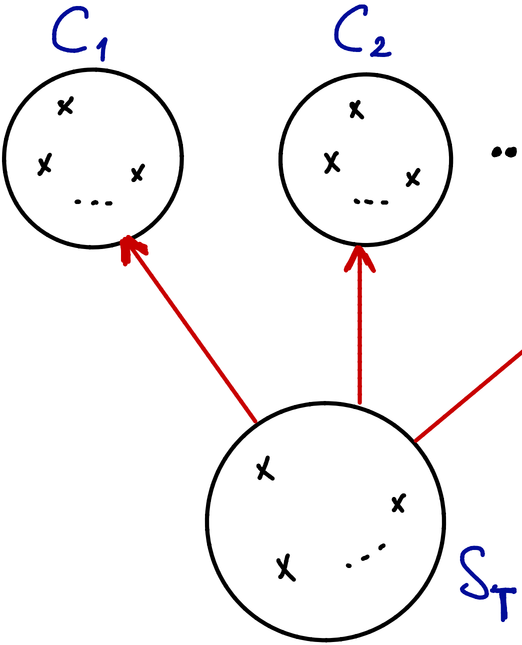
Theorem: Assume $S_R \neq \phi$. Then

$$S_R = \bigcup_{i=1}^{k} C_i \quad (k\text{: finite or infinite}),$$

where C_i, $1 \leqslant i \leqslant k$ are disjoint irreducible closed sets of recurrent states.

Proof: It suffices to show: If $C_1 \& C_2$ are two irreducible & closed sets, then either $C_1 = C_2$ or $C_1 \cap C_2 = \phi$.

Assuming $C_1 \cap C_2 \neq \phi$, we need to show $C_1 = C_2$. In fact, let $y \in C_1$ be arbitrary, we want: $y \in C_2$
($\therefore C_1 \subseteq C_2 \subseteq C_1$).Indeed, $\exists x \in C_1 \cap C_2$, then $C_2 \ni x \to y$.
$\therefore y \in C_2$.



C1
C2
Ck
ST

Corollary: If C is an irreducible & closed set, then

$$\text{either } C \subseteq S_R \text{ or } C \subseteq S_T.$$

In particular, if C is a finite, irreducible & closed set, then

$$C \subseteq S_R.$$

In terms of the (*disjoint*) decomposition

$$S = S_R \cup S_T = \left(\cup_{i=1}^{k} C_i\right) \cup S_T,$$

we may rewrite P as the canonical form:

$$P = \begin{array}{c||c|c|c|c|c} & \boxed{C_1} & \boxed{C_2} & \cdots & \boxed{C_k} & \boxed{S_T} \\ \hline \boxed{C_1} & \boxed{*} & 0 & \cdots & 0 & 0 \\ \hline \boxed{C_2} & 0 & \boxed{*} & \ddots & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \boxed{*} & 0 & 0 \\ \hline \boxed{C_k} & 0 & 0 & 0 & \boxed{*} & 0 \\ \hline \boxed{S_T} & \boxed{*} & \boxed{*} & \boxed{*} & \boxed{*} & \boxed{*} \end{array},$$

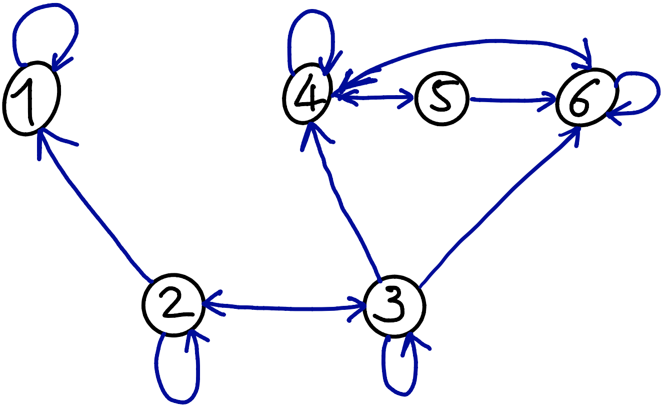
where $\boxed{*}$ denotes the sub-matrix with possible $\neq 0$ entries.

Example:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix}$$

Q.: Determine $S = S_R \cup S_T = \left(\cup_{i=1}^{k} C_i\right) \cup S_T$.

- $1 \to 1 \;\therefore\; C_1 = \{1\}$.
- $4 \to 5 \to 6 \to 4$ (irreducible), and $4, 5, 6$ do not lead to any other state (closed). $\therefore\; C_2 = \{4, 5, 6\}$.
- $2 \to 1$, $3 \to 4$, $\therefore\; S_T = \{2, 3\}$.



We then reformulate P in the canonical form of

$$P = \begin{array}{c} \\ a=1 \\ b=4 \\ c=5 \\ d=6 \\ e=2 \\ f=3 \end{array} \begin{array}{c} \begin{array}{cccccc} a=1 & b=4 & c=5 & d=6 & e=2 & f=3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{5} & \frac{2}{5} \end{bmatrix} \end{array}.$$

Final Issue: Assume that C is an irreducible & closed set of recurrent states. Then,

$$\boxed{T_c \stackrel{\text{def}}{=} \min\{n \geqslant 1 : X_n \in C\}}$$

denotes the **hitting time** of C.

We can also consider

$$\boxed{\rho_C(x) \stackrel{\text{def}}{=} P_x(T_C < \infty)}$$

is the prob that the chain starting at x hits C in finite time (or is absorbed by the set C).

NOTE: Once the chain hits C, it remains in C forever. (Why?)

$\therefore \rho_C(\cdot)$ is called the **absorption prob**.

It is clear to see:

$$\rho_C(x) = \begin{cases} 1 & \text{if } x \in C, \\ 0 & \text{if } x \text{ is recurrent but } \notin C. \end{cases}$$

Q.: How to compute $\rho_C(x)$, $x \in S_T$?

Indeed, assume S_T **is finite**, then for $x \in S_T$,

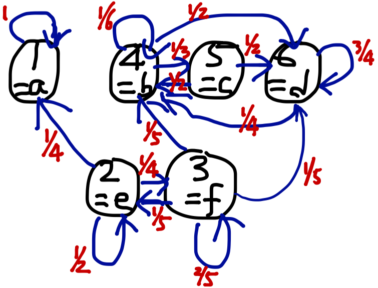
$$\boxed{\rho_C(x) = \sum_{y \in C} P(x,y) + \sum_{y \in S_T} P(x,y)\rho_C(y).} \qquad (*)$$

Assume $d_T \overset{\text{def}}{=} \#$ of S_T is finite
$\#$ of unknowns $= d_T$: $\rho_C(x)$, $x \in S_T$
$\#$ of equations $= d_T$
$\therefore$ it is possible to find out $\rho_C(x)$, $x \in S_T$ by solving the linear system of d_T equations.

Theorem. Let S_T be finite. Then $(*)$ admits a unique solution.

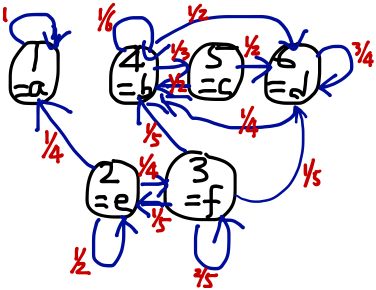
Proof. Omitted.

Example: Find $\underbrace{\rho_{C_2}(e)}_{\overset{\text{def}}{=}x}$, $\underbrace{\rho_{C_2}(f)}_{\overset{\text{def}}{=}y}$?



$$\begin{cases} x = \rho_{C_2}(e) = \underbrace{[0+0+0]}_{\sum_{j\in C_2=\{b,c,d\}} P(e,j)} + \underbrace{[\frac{1}{2}x + \frac{1}{4}y]}_{\sum_{j\in S_T=\{e,f\}} P(e,j)\rho_{C_2}(j)} \\ y = \rho_{C_2}(f) = \underbrace{[\frac{1}{5}+0+\frac{1}{5}]}_{\sum_{j\in C_2=\{b,c,d\}} P(f,j)} + \underbrace{[\frac{1}{5}x + \frac{2}{5}y]}_{\sum_{j\in S_T=\{e,f\}} P(f,j)\rho_{C_2}(j)} \end{cases} \quad \therefore\ x = \frac{2}{5}, y = \frac{4}{5}.$$

Similarly, let $\rho_{C_1}(e) = x$ and $\rho_{C_1}(f) = y$,



then

$$\begin{cases} x = [\frac{1}{4}] + [\frac{1}{2}x + \frac{1}{4}y] \\ y = [0] + [\frac{1}{5}x + \frac{2}{5}y] \end{cases} \Longrightarrow x = \frac{3}{5}, \quad y = \frac{1}{5}.$$

Remark (i): $\sum_i \rho_{C_i}(x) \equiv 1$, $x \in S_T$ (**finite**).

Indeed,

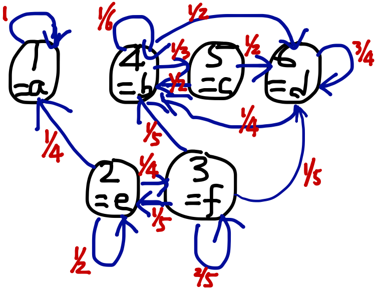
$$\sum_i \rho_{C_i}(x) = \sum_i P_x(T_{C_i} < \infty) = P_x(T_{S_R} < \infty) = 1.$$

Heuristically, it is obvious:

- We totally have finite transient states.
- Each transient state is visited only finite times.
- Surely the chain from x hits a recurrent state in finite time, so the prob $= 1$.

Remark (ii): $\rho_{xy} = \rho_C(x)$, $x \in S_T$, $y \in C$.

Apply it to the previous example:



$$\frac{2}{5} = \rho_{C_2=\{b,c,d\}}(e) = \rho_{eb} = \rho_{ec} = \rho_{ed},$$
$$\frac{4}{5} = \rho_{C_2=\{b,c,d\}}(f) = \rho_{fb} = \rho_{fc} = \rho_{fd}.$$

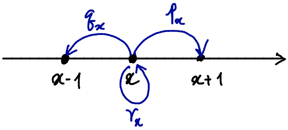
§1.3 More examples

Examples 1: Birth & Death Chain.

- Setting:

$\{X_n\}_{n=0}^{\infty}$, $S = \{0, 1, \cdots, d\}$ (d : finite or ∞)

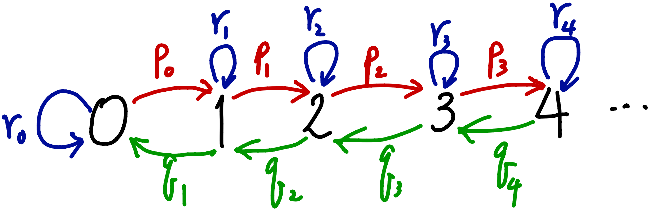
$$P(x,y) = \begin{cases} q_x & \text{if } y = x-1 \\ r_x & \text{if } y = x \\ p_x & \text{if } y = x+1 \\ 0 & \text{otherwise} \end{cases} \quad \text{where } q_x + \gamma_x + \rho_x = 1.$$



$q_0 = 0$; $p_d = 0$, if d is finite.

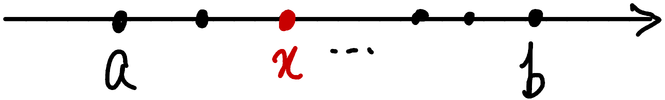
Note: the transition probs are functions of states!

$$
P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ d-1 \\ d \end{array}
\begin{array}{c}
\begin{array}{ccccccc} 0 & 1 & 2 & 3 & \cdots & d-1 & d \end{array} \\
\begin{bmatrix}
r_0 & p_0 & & & & & \\
q_1 & r_1 & p_1 & & & & \\
& q_2 & r_2 & p_2 & & & \\
& & \ddots & \ddots & \ddots & & \\
& & & q_{d-1} & r_{d-1} & p_{d-1} \\
& & & & q_d & r_d
\end{bmatrix}
\end{array}
$$



A general question: Given $a, b \in S$ with $a < b$, compute

$$u(x) \stackrel{\text{def}}{=} P_x(T_a < T_b), \quad a < x < b,$$
$$v(x) \stackrel{\text{def}}{=} P_x(T_a > T_b), \quad a < x < b.$$



$\{T_a < T_b\}$ = Before the chain hits b, it hits a, (i.e., the chain hits a earlier than b)

$\{T_a > T_b\}$ = Before the chain hits a, it hits b, (i.e., the chain hits b earlier than a)

Claim:

(i) $u(a) = 1$, $u(b) = 0$.

(ii) $u(x) = q_x u(x-1) + r_x u(x) + p_x u(x+1)$ for $a < x < b$.

(iii) $u(x) = \dfrac{\sum_{y=x}^{b-1} \gamma_y}{\sum_{y=a}^{b-1} \gamma_y}$ for $a < x < b$.

(iv) $v(x) = 1 - u(x) = \dfrac{\sum_{y=a}^{x-1} \gamma_y}{\sum_{y=a}^{b-1} \gamma_y}$ for $a < x < b$,

where γ_x are defined by

$$\gamma_x \overset{\text{def}}{=} \begin{cases} 1 & \text{if } x = 0 \\ \frac{q_1 \cdots q_x}{p_1 \cdots p_x} & \text{if } 1 \le x \le d-1. \end{cases}$$

Proof:

(i) is obvious.

(ii) follows by

$$
\begin{aligned}
P_x(A) =& P_x(A, X_1 = x-1) + P_x(A, X_1 = x) \\
& + P_x(A, X_1 = x+1) \\
=& P_x(X_1 = x-1) P_x(A | X_1 = x-1) \\
& + P_x(X_1 = x) P_x(A | X_1 = x) \\
& + P_x(X_1 = x+1) P_x(A | X_1 = x+1) \\
=& P(X_1 = x-1 | X_0 = x) P(A | X_0 = x, X_1 = x-1) \\
& + P(X_1 = x | X_0 = x) P(A | X_0 = x, X_1 = x) \\
& + P(X_1 = x+1 | X_0 = x) P(A | X_0 = x, X_1 = x+1) \\
=& q_x P_{x-1}(A) + r_x P_x(A) + p_x P_{x+1}(A).
\end{aligned}
$$

Proof of (iii):

$$u(x) = q_x u(x-1) + (1-p_x-q_x)u(x) + p_x u(x+1)$$
$$(p_x+q_x)u(x) = q_x u(x-1) + p_x u(x+1)$$

$$\begin{aligned} u(x+1)-u(x) &= \frac{q_x}{p_x}[u(x)-u(x-1)] \quad (a<x<b) \\ &= \frac{q_x \cdot q_{x-1}}{p_x \cdot p_{x-1}}[u(x-1)-u(x-2)] \\ &= \cdots \\ &= \left(\frac{q_x}{p_x}\right)\left(\frac{q_{x-1}}{p_{x-1}}\right)\cdots\left(\frac{q_{a+1}}{p_{a+1}}\right)[u(a+1)-u(a)] \\ &= \frac{\gamma_x}{\gamma_a}[u(a+1)-u(a)]. \end{aligned}$$

Note: $\sum_{x=1}^{b-1}(\cdot) \Rightarrow \underbrace{u(b)-u(a)}_{=-1} = \frac{\sum_{x=a}^{b-1}\gamma_x}{\gamma_a}[u(a+1)-u(a)]$

$$\therefore u(x+1)-u(x) = -\frac{\gamma_x}{\sum_{x=a}^{b-1}\gamma_x} \quad (a \leqslant x < b)$$

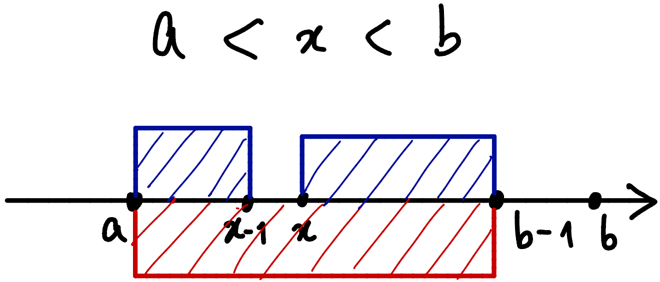
Further, change x to y, $\sum_{y=x}^{b-1} \Rightarrow$

$$\underbrace{u(b)}_{=0} - u(x) = -\frac{\sum_{y=x}^{b-1} \gamma_y}{\sum_{y=a}^{b-1} \gamma_y}, \quad \therefore u(x) = \sum_{y=x}^{b-1} \gamma_y \Big/ \sum_{y=a}^{b-1} \gamma_y .$$

Reminder:

- $u(x) \stackrel{\text{def}}{=} P_x(T_a < T_b), \quad a < x < b.$
- $\gamma_x \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } x = 0 \\ \frac{q_1 \cdots q_x}{p_1 \cdots p_x} & \text{if } 1 \le x \le d-1. \end{cases}$

Sum:



$$P_x(\underbrace{T_a < T_b}_{\text{“Death faster”}}) = \sum_{y=x}^{b-1} \gamma_y \Big/ \sum_{y=a}^{b-1} \gamma_y,$$

$$P_x(\underbrace{T_a > T_b}_{\text{“Birth faster”}}) = \sum_{y=a}^{x-1} \gamma_y \Big/ \sum_{y=a}^{b-1} \gamma_y.$$

e.g.: Set:

- A gambler bets \$1 each time.
- The prob of winning or losing each bet is $9/19$ and $10/19$, resp.
- The gambler will quit as soon as his net winning is \$25 or his net loss is \$10.

Q.:

(i) Find the prob he quits and wins.

(ii) Find his expected loss.

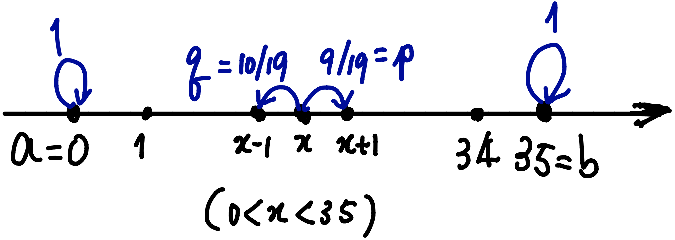
Sol.: Let

$$X_n \overset{\text{def}}{=} \text{the capital of the gambler at time } n = 0, 1, 2, \cdots$$

For simplicity, we choose

$$X_0 = 10, \quad S = \{0, 1, \cdots, 35\}.$$

$\{X_n\}_{n=0}^{\infty}$ forms a birth & death chain on S with



$$\gamma_y = \left(\frac{q}{p}\right)^y = \left(\frac{10}{9}\right)^y, 0 \leqslant y \leqslant 34.$$

(i) **Find the prob he quits and wins:** To find

$$P_{10}(\underbrace{T_{35} < T_0}_{\text{"Birth faster"}}) = \frac{\sum_{y=0}^{9} \gamma_y}{\sum_{y=0}^{34} \gamma_y} = \frac{\sum_{y=0}^{9} (\frac{10}{9})^y}{\sum_{y=0}^{34} (\frac{10}{9})^y} = \frac{(\frac{10}{9})^{10} - 1}{(\frac{10}{9})^{35} - 1}$$

$$= 0.047.$$

(ii) **Find his expected loss:**

gain (+25)	loss (−10)
0.047	1 − 0.047

The expected loss is

$$(1 - 0.047)(-10) + (0.047)(25) = -8.36. \quad \square$$

- We are further interested in the below situation:

Assume that $S = \{0, 1, 2, \cdots\}$ is **infinite**, and the birth & death chain is **irreducible**, namely,

$$p_x > 0, \forall\, x \geqslant 0, \quad \text{and} \quad q_x > 0, \forall\, x \geqslant 1.$$

Q.: When such chain is recurrent or transient? (NOT obvious for an irreducible chain with infinite states!)

Proposition: The chain is recurrent **iff**

$$\sum_{k=0}^{\infty} \gamma_k = \infty.$$

Pf.: Since the chain is irreducible, we only need to consider one state, namely, 0. Observe that

$$\rho_{00} = P_0(T_0 < \infty) = r_0 + p_0 P_1(T_0 < \infty), \qquad (*)$$

where

$$\begin{aligned}\rho_{10} = P_1(T_0 < \infty) &= \lim_{n\to\infty} P_1(T_0 < T_n) \\ &= \lim_{n\to\infty} \left[1 - \frac{1}{\sum_{k=0}^{n-1} \gamma_k}\right]. \quad (**)\end{aligned}$$

Therefore,

$$\begin{aligned}&0 \text{ is recurrent, i.e. } \rho_{00} = 1 \\ &\overset{(*)\&r_0+p_0=1}{\iff} \rho_{10} = P_1(T_0 < \infty) = 1 \\ &\overset{(**)}{\iff} \sum_{k=0}^{\infty} \gamma_k = \infty.\end{aligned}$$

□

Remark: For instance, let

$$p_x \equiv p > 0,\ q_x \equiv q > 0,\ 0 < \rho + q \leqslant 1.$$

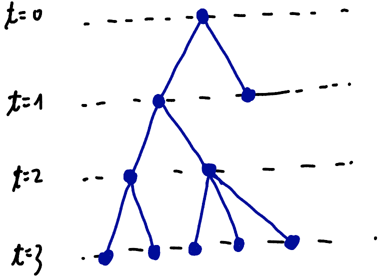
Then,

$$\boxed{\sum_{k=0}^{\infty} \gamma_k = \sum_{k=0}^{\infty} \left(\frac{q}{p}\right)^k.}$$

- If $p > q$, then $\sum_{k=0}^{\infty} \gamma_k$ is finite. The chain is transient .
- If $p = q$ or $p < q$, then $\sum_{k=0}^{\infty} \gamma_k = \infty$. The chain recurrent.

Example 2. Branching chain.

Each particle generates ξ particles independently in the next generation.



$X_n \stackrel{\text{def}}{=}$ the total no of particles in the n^{th} generation

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

$$P(x,y) = P(\xi_1 + \xi_2 + \cdots + \xi_x = y), \quad x \geqslant 1.$$

Q.: Determine

$$\rho \stackrel{\text{def}}{=} \text{the prob that the descendants of a given particle eventually become } \textbf{extinct}.$$

We call ρ to be the **extinction prob** of the chain. Then,

$$\rho = \rho_{10} = P_1(T_0 < \infty).$$

1^{st} **Obervation:** Suppose ξ has the pdf

$$p_k = P(\xi = k),\ k = 0, 1, 2, \cdots$$

Then,

$$P(1, k) = P(\xi_1 = k) = p_k,\ k = 0, 1, 2, \cdots$$

From this we see:

- If $p_0 = 0$, then each individual cannot change to zero, so population never extinct, i.e. $\rho = 0$.
- If $p_0 = 1$, then it extincts for sure, i.e., $\rho = 1$.

To avoid two trivial cases, we always assume

$$0 < p_0 < 1.$$

2^{nd} **Obervation:** Assuming there are x particles, the prob for them to extinct is

$$\rho_{x0} = \rho^x.$$

(**Pf.:** Use **independence**!)

3[rd] **Obervation:** Let

$$\mu \overset{\text{def}}{=} E(\xi) = \sum_{k=0}^{\infty} kp_k = \sum_{k=1}^{\infty} kp_k.$$

Then, $E(X_{n+1}|X_n = k) = E(\xi_1 + \cdots + \xi_k) = k\mu$,

$$\begin{aligned} E(X_n) &= \sum_{k=0}^{\infty} E(X_n|X_{n-1} = k)P(X_{n-1} = k) \\ &= \sum_{k=0}^{\infty}(k\mu)P(x_{n-1} = k) \\ &= \mu E(X_{n-1}) \\ &= \cdots \\ &= \mu^n E(X_0). \end{aligned}$$

Claim: If $\mu < 1$, then population will **extinct for sure**, i.e., $\rho = 1$.

Proof:

$$
\begin{aligned}
P_1(T_0 > n) &\leqslant P_1(X_n \geqslant 1) \quad (\because \{T_0 > n\} \subseteq \{X_n \geqslant 1\}) \\
&= \sum_{k=1}^{\infty} P_1(X_n = k) \leqslant \sum_{k=1}^{\infty} k P_1(X_n = k) \\
&= \sum_{k=0}^{\infty} k P_1(X_n = k) \\
&= E(X_n) = \mu^n E(X_0) \xrightarrow{n \to \infty} 0 \ (\because \mu < 1)
\end{aligned}
$$

Therefore

$$
\underbrace{\rho}_{\text{extinction prob}} = \rho_{10} = P_1(T_0 < \infty) = \lim_{n \to \infty} P_1(T_0 \leqslant n)
$$

$$
= \lim_{n \to \infty} [1 - P_1(T_0 > n)] = 1. \quad \square
$$

What about $\mu \geqslant 1$?

$$\begin{aligned}\rho = \rho_{10} = P_1(T_0 < \infty) \\ &= P(1,0) + \sum_{k=1}^{\infty} P(1,k)\rho_{k0} \\ &= p_0 + \sum_{k=1}^{\infty} p_k \rho^k = \sum_{k=0}^{\infty} p_k \rho^k,\end{aligned}$$

i.e., ρ solves the equation $\boxed{t = \Phi(t)}$ with

$$\boxed{\Phi(t) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} p_k t^k,}$$

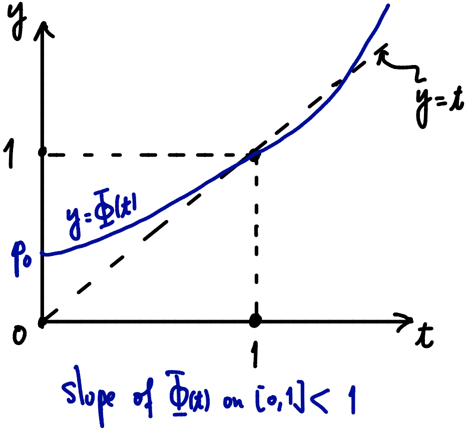
which is called the **moment generating function** of the pdf $(p_k)_{k\geq 0}$ of ξ.

Observe:

- $\Phi'(t) \geqslant 0$, $\Phi''(t) \geqslant 0$, ($\therefore \Phi(t) \uparrow$ & concave upward).
- $\Phi(0) = p_0 \in (0,1)$, $\Phi(1) = \sum_{k=0}^{\infty} p_k = 1$.
- $\Phi'(1) = \sum_{k=1}^{\infty} k p_k = E(\xi) = \mu$.

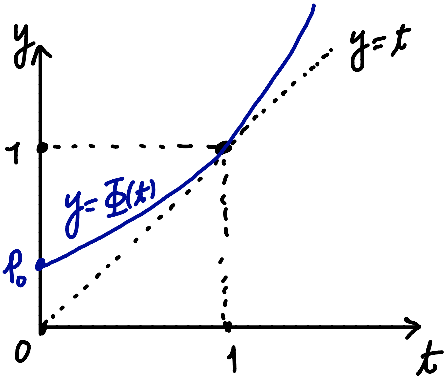
Then, we have **three cases:**

Case (i): $\mu < 1$.



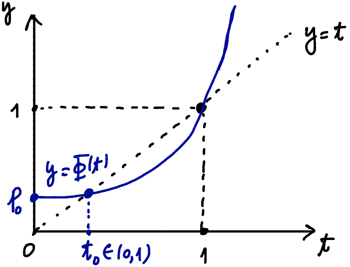
$\therefore\ \rho = 1$ (**extinct for sure**, as proved before)

Case (ii): $\mu = 1$.



$\therefore\ \rho = 1$ (**extinct for sure!**)

Case (iii): $\mu > 1$.



$\Phi(t) = t$ at $t = t_0 \in (0, 1)$ or $t = 1$.

Claim: In this case, $P_1(T_0 \leqslant n) \leqslant t_0$ for all $n = 1, 2, \cdots$ (**proved later**).

$$\therefore \rho = \rho_{10} = P_1(T_0 < \infty)$$
$$= \lim_{n\to\infty} P_1(T_0 \leqslant n) \leqslant t_0$$

$\therefore \rho = t_0$ is the **only** solution.

Proof of Claim: Use induction. Set

$$a_n \stackrel{\text{def}}{=} P_1(T_0 \leqslant n).$$

$n = 0$: $a_0 = P_1(T_0 \leqslant 0) = 0 < t_0$.

Assuming $a_n \leqslant t_0 (n \geqslant 0)$, consider

$$\begin{aligned} a_{n+1} &= P_1(T_0 \leqslant n+1) \\ &= \underbrace{P(1,0)}_{=p_0} + \sum_{k=1}^{\infty} \underbrace{P(1,k)}_{p_k} \underbrace{P_k(T_0 \leqslant n)}_{=[P_1(T_0 \leqslant n)]^k = a_n^k} \\ &= \sum_{k=0}^{\infty} p_k a_n^k \\ &= \Phi(a_n) \leqslant \Phi(t_0) = t_0 \quad (\Phi \text{ is nondecreasing}). \qquad \square \end{aligned}$$

e.g.: Every man has 3 kids with prob $1/2$ being boy and $1/2$ being girl. Find the prob that the male live eventually extinct.

Sol.: $p_0 = P(\xi = 0) = \frac{1}{8}$, $p_1 = P(\xi = 1) = \frac{3}{8}$, $p_2 = P(\xi = 2) = \frac{3}{8}$, $p_3 = P(\xi = 3) = \frac{1}{8}$.

$$E(\xi) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2} > 1$$

$$\Phi(t) = \frac{1}{8} + \frac{3}{8}t + \frac{3}{8}t^2 + \frac{1}{8}t^3$$

$$\text{let } \Phi(t) = t, \text{i.e. } t = \frac{1}{8} + \frac{3}{8}t + \frac{3}{8}t^2 + \frac{1}{8}t^3$$

Solutions: $t = 1,\ \sqrt{5} - 2$. Then

$$\rho = \sqrt{5} - 2$$

is the extinct prob. □

Example 3. Queuing chain.

Setting:

- In a queue, let ξ_n denote the no of arrivals in the n-th unit time. $\{\xi_n\}_{n=1}^{\infty}$ are i.i.d.r.v. with pdf:

$$f(k) = p_k, \quad k = 0, 1, 2, \cdots$$

- The service of a customer is exactly one in a unit time.

Let X_n denote the no of customers in the queue.

$$P(x, y) = f(\underbrace{y - (x-1)}_{\text{no of arrivals}}), \quad x \geqslant 1,$$

$$P(0, y) = f(y).$$

Note: $P(1, y) = P(0, y)$.

Q.: Assuming that the chain is irreducible, check if the chain is **recurrent** or **transient**, i.e. letting

$$\rho = \rho_{00} = P_0(T_0 < \infty),$$

decide

$$\text{if } \rho = 1 \text{ or } \rho < 1.$$

Note. If

$$p_0 > 0 \ \& \ p_0 + p_1 < 1,$$

then the chain is irreducible. (Ex. 37 on Page 46).

Let

$$\Phi(t) \stackrel{\text{def}}{=} p_0 + p_1 t + p_2 t^2 + \cdots$$
$$= \sum_{k=0}^{\infty} p_k t^k$$
$$= \sum_{k=0}^{\infty} f(k) t^k$$

be the moment generating function of f.

Claim: $\rho = \rho_{00}$ solves $\Phi(t) = t$.

Pf.:

- Note

$$\rho_{00} = P(0,0) + \sum_{k=1}^{\infty} P(0,k)\rho_{k0},$$

$$\rho_{10} = P(1,0) + \sum_{k=1}^{\infty} P(1,k)\rho_{k0},$$

$$P(1,k) = P(0,k), \ \forall k \geq 0.$$

Therefore,

$$\rho_{10} = \rho_{00} = \rho.$$

- **To show: $\rho_{x,x-1} = \rho_{10} = \rho$ for all $x > 1$.**

In fact, we observe that
for the chain starting at $x > 1$ ($\therefore\ x - 1 \geq 1$),
the event $T_{x-1} = n$ means

$$n = \min\{m > 0 : x + (\xi_1 - 1) + \cdots (\xi_m - 1) = x - 1\},$$

i.e.

$$n = \min\{m > 0 : 1 + (\xi_1 - 1) + \cdots (\xi_m - 1) = 0\}.$$

Therefore, $P_x(T_{x-1} = n) = P_1(T_0 = n)$, $\forall n \geq 1$.

$$\therefore \rho_{x,x-1} = \rho_{10} = \rho.$$

- **To show:**

$$\rho_{x,0} = \rho_{x,x-1} \cdot \rho_{x-1,0}, \ \ \forall\, x \geqslant 2. \qquad (*)$$

(Ex. 39, P46). If so, then

$$\rho_{x,0} = \rho\rho_{x-1,0} = \cdots = \rho^x,$$

(also true for $x = 1$), and hence

$$\begin{aligned}\rho = \rho_{00} &= P(0,0) + \sum_{k=1}^{\infty} P(0,k)\rho_{k0} \\ &= p_0 + \sum_{k=1}^{\infty} p_k \rho^k \\ &= \Phi(\rho).\end{aligned}$$

Proof of $(*)$: Let $x \geq 2$. Note that for $m \geq 2$,

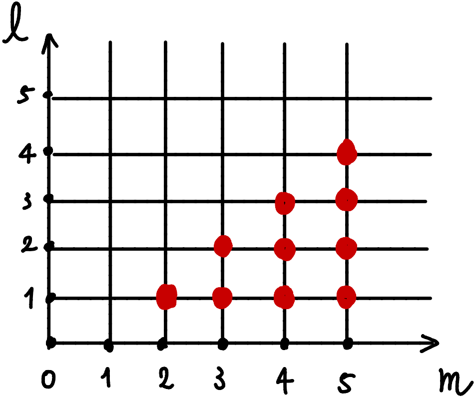
$$P_x(T_0 = m) = \sum_{\ell=1}^{m-1} P_x(T_{x-1} = \ell) P_{x-1}(T_0 = m - \ell).$$

Then,

$$\begin{aligned}
\rho_{x,0} &= P_x(T_0 < \infty) = \sum_{m=1}^{\infty} P_x(T_0 = m) \\
&= \sum_{m=2}^{\infty} P_x(T_0 = m) \quad \text{(Note: } P_x(T_0 = 1) = 0 \text{ for } x \geq 2\text{)} \\
&= \sum_{m=2}^{\infty} \sum_{\ell=1}^{m-1} P_x(T_{x-1} = \ell) P_{x-1}(T_0 = m - \ell) \\
&= \sum_{\ell=1}^{\infty} \sum_{m=\ell+1}^{\infty} P_x(T_{x-1} = \ell) P_{x-1}(T_0 = m - \ell) \text{ (see later)} \\
&= \sum_{\ell=1}^{\infty} P_x(T_{x-1} = \ell) \rho_{x-1,0} \\
&= \rho_{x,x-1} \rho_{x-1,0}. \quad \square
\end{aligned}$$

Note:

$$\sum_{m=2}^{\infty}\sum_{\ell=1}^{m-1} = \sum_{\ell=1}^{\infty}\sum_{m=\ell+1}^{\infty} .$$



Sum: Let $\mu = E(\xi)$. Then

- If $\mu \leqslant 1$, then $\Phi(\rho) = \rho$ has the only solution $\rho = 1$. The chain is recurrent.
- If $\mu > 1$, then $\Phi(\rho) = \rho$ has two solutions 1 and $t_0 \in (0, 1)$. As in the previous example, one has to take $\rho = t_0$. The chain is transient.

$\square$

Chapter 2:
Stationary Distribution

§2.1 Stationary Distribution

Motivation: Recall the two state MC with

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}, \quad 0 < p, q < 1.$$

We have shown (Chapter 1):

$$\lim_{n\to\infty} P(X_n = 0) = q/(p+q) \stackrel{\text{def}}{=} a,$$
$$\lim_{n\to\infty} P(X_n = 1) = p/(p+q) = 1 - a.$$

Denote $\pi = [a, 1-a]$ (limit distribution), i.e.

$$\pi = \lim_{n\to\infty} \underbrace{[P(X_n = 0), P(X_n = 1)]}_{\text{pdf of } X_n} = \lim_{n\to\infty} \pi_0 P^n, \qquad (*)$$

where $\pi_0 = [P(X_0 = 0), P(X_0 = 1)]$ is the initial distribution.

Note: Here π is independent of π_0.

We discuss two issues related to π:

- It is direct to verify

$$\pi P = \pi,$$

i.e. $$[\frac{q}{p+q}, \frac{p}{p+q}] \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} = [\frac{q}{p+q}, \frac{p}{p+q}].$$

Hence by induction,

$$\pi P^n = \pi, \quad n = 1, 2, \cdots$$

It means that if the chain starts with X_0 with pdf π, then at any time $n = 1, 2, \cdots$, X_n has the same distribution as π.

Note: $(*)$ also directly implies

$$\pi = \lim_{n\to\infty} (\pi_0 P^{n-1}) P = \pi P.$$

- One can also show:

$$\lim_{n\to\infty} P^n = \begin{bmatrix} a & 1-a \\ a & 1-a \end{bmatrix} = \begin{bmatrix} \pi \\ \pi \end{bmatrix}.$$

Two ways:

(i) Diasonalize P. See Tutorial or Exercise.

(ii) Find

$$\lim_{n\to\infty} P^n(x,y) = \lim_{n\to\infty} P(X_n = y | X_0 = x)$$
$$= \lim_{n\to\infty} P_x(X_n = y).$$

As proved before, for $x = 0$ or 1

$\lim_{n\to\infty} P_x(X_n = 0) = a$ i.e. the 1^{st} column is a,

$\lim_{n\to\infty} P_x(X_n = 1) = 1 - a$ i.e. the 2^{nd} column is $1 - a$.

Observe: The fact that

$$\pi P = \pi = 1 \cdot \pi$$

means that π is the **left 1-eigenvector** of P.

Thus, we may also find the limit distribution π directly by solving

$$[u, v] \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} = [u, v],$$

i.e.

$$[u, v] \begin{bmatrix} -p & p \\ q & -q \end{bmatrix} = [0, 0], i.e., \begin{bmatrix} -p & q \\ p & -q \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

$$(\because\ u \geqslant 0, v \geqslant 0, u+v = 1\ \therefore\ u = \frac{q}{p+q}, v = \frac{p}{p+q})$$

General Situaion:

If

$$\pi = \lim_{n\to\infty} \pi_0 P^n \quad \text{exists}$$

for some initial distribution π_0,

then π satisfies

$$\pi = \left(\lim_{n\to\infty} \pi_0 P^{n-1}\right) P = \pi P,$$

i.e.,

$$\pi = \pi P,$$

or equivalently,

$$\pi(y) = \sum_{x\in S} \pi(x) P(x, y), \quad \forall\, y \in S.$$

Definition: We say that a probability row-vector π is a **stationary distribution** for P **if**

$$\pi = \pi P,$$

i.e. the pdf π is a left 1-eigenvector of P.

Two basic questions:

(i) **Existence ($\exists$):** Does every P have a SD?

(ii) **Uniqueness (!):** Is the SD unique?

Two notes:

(1) **If** $\pi = \pi P$ has a unique solut'n **then** the limit $\lim_{n\to\infty} \pi_0 P^n$ (if it exists) is independent of π_0.

(2) If $\lim_{n\to\infty} P^n = \begin{bmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{bmatrix}$ then for any initial distri π_0,

$$\lim_{n\to\infty} \pi_0 P^n = \pi,$$

i.e. the limit exists and is independent of π_0.

This also suggests a way of finding the SD of P.

Proposition: Let P be a Markov matrix with **finite** state space S. Assume:

(i) The left 1-eigenvector (which must exist) can be chosen to have all nonegative entries;

(ii) 1 is a simple eigenvalue;

(iii) other eigenvalues $|\lambda_i| < 1$.

Then P has a unique SD π, i.e. $\pi P = \pi$, and

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{bmatrix}.$$

Pf.: (Sketch only) (See Lawler P11-15)

$$P = QDQ^{-1}, \quad D = \left[\begin{array}{c|c} 1 & O \\ \hline O & M \end{array}\right], \quad M^n \to 0$$

Q: columns are right eigenvectors; 1st row is $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$

Q^{-1}: rows are left eigenvectors; 1st row is a prob vector, denoted by π

$$\therefore \lim_{n\to\infty} P^n = \lim_{n\to\infty} QD^nQ^{-1} = Q \left[\begin{array}{c|c} 1 & O \\ \hline O & O \end{array}\right] Q^{-1} = \begin{bmatrix} \pi \\ \vdots \\ \pi \end{bmatrix}$$

□

Remarks:

(1) **If** 1 is NOT a simple eigenvalue, **then** π may not be unique: e.g.

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}$$

$$\left.\begin{array}{l} \pi_1 \text{ is the SD of } P_1 \\ \pi_2 \text{ is the SD of } P_2 \end{array}\right\} \Rightarrow [\lambda\pi_1, (1-\lambda)\pi_2] \text{ is the SD of } P$$

(2) **Without** (iii), the limit $\lim_{n\to\infty} P^n$ may not exist (but π still may exist): e.g.

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \pi = [\frac{1}{2}, \frac{1}{2}],$$

BUT eigenvalues: ± 1.

(3) Two further **Facts** for **finite** S (without proof):

Fact a. If for some $n \geq 1$, P^n has all entries strictly positive, then three conditions are satisfied, therefore, P has a unique SD π, and

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi \\ \vdots \\ \pi \end{bmatrix}.$$

Fact b. If P is irreducible, then P still has a unique SD.

(But $\lim_{n\to\infty} P^n$ may not exist, e.g., $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$)

Computation Technique for **finite** S:

Case 1: P is irreducible.

$$\pi P = \pi, \text{i.e., } \; P^T \pi^T = \pi^T, \text{i.e. } (P^T - I)\pi^T = 0.$$

$$P^T - I \xrightarrow{\text{row operation}} \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & * \end{bmatrix},$$

Upper diagonal form.

Fact b above assures that the solution exists uniquely. (**Note:** Find π as a prob row-vector)

Case 2. P is reducible.

For instance, let $S = C_1 \cup C_2 \cup S_T$.
Reordering S accordingly, write

$$P = \begin{bmatrix} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ S_1 & S_2 & Q \end{bmatrix}, \quad P^n = \begin{bmatrix} P_1^n & 0 & 0 \\ 0 & P_2^n & 0 \\ S_{1n} & S_{2n} & Q^n \end{bmatrix}$$

$$i = 1, 2 : \lim_{n \to \infty} P_i^n = \begin{bmatrix} \pi_i \\ \vdots \\ \pi_i \end{bmatrix}, \pi_i : \text{SD of } P_i,$$

$$\lim_{n \to \infty} Q^n = 0,$$

(Chap1: For $y \in S_T$, $\lim_{n \to \infty} P^n(x, y) = 0$, $\forall x \in S$).

In fact, all eigenvalues of Q have moduli < 1.

$$\therefore \lim_{n\to\infty} P^n = \begin{bmatrix} \begin{bmatrix}\pi_1\\ \vdots \\ \pi_1\end{bmatrix} & 0 & 0 \\ 0 & \begin{bmatrix}\pi_2\\ \vdots \\ \pi_2\end{bmatrix} & 0 \\ A_1 & A_2 & \mathbf{0} \end{bmatrix},$$

$A_1 = \lim_{n\to\infty} S_{1n}$, $A_2 = \lim_{n\to\infty} S_{2n}$:

$A_1(x,y) =$ prob from $x \in S_T$ to $y \in C_1$ in the long run,

$A_2(x,y) =$ prob from $x \in S_T$ to $y \in C_2$ in the long run.

Q.: How to find A_1, A_2?

Solution: Assume

$$S_T = \{x_1, x_2, \cdots, x_\ell\}.$$

First find

$$\rho_{C_i}(x), \quad x \in S_T,\ i = 1, 2,$$

(**absorption prob** of C_i, i.e. prob to enter C_i).

Then, distribute according to π_i, e.g.

$$A_1 = \begin{bmatrix} \rho_{C_1}(x_1)\pi_1 \\ \rho_{C_1}(x_2)\pi_1 \\ \vdots \\ \rho_{C_1}(x_\ell)\pi_1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} \rho_{C_2}(x_1)\pi_2 \\ \rho_{C_2}(x_2)\pi_2 \\ \vdots \\ \rho_{C_2}(x_\ell)\pi_2 \end{bmatrix}. \quad \square$$

Example 1. (Gambler's ruin chain) Let

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

Show that

$$\lim_{n\to\infty} P^n = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{3}{4} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & 0 & 0 & \frac{3}{4} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Solution: From P, one can check that

$$C_1 = \{0\},\ C_2 = \{4\},\ S_T = \{1, 2, 3\},$$

and

$$S = C_1 \cup C_2 \cup S_T.$$

After reordering,

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

Set $P = \left[\begin{array}{cc|c} 1 & 0 & \multirow{2}{*}{$\mathbf{0}$} \\ 0 & 1 & \\ \hline \multicolumn{2}{c|}{S} & Q \end{array}\right]$. Then, $P^n = \left[\begin{array}{cc|c} 1 & 0 & \multirow{2}{*}{$\mathbf{0}$} \\ 0 & 1 & \\ \hline \multicolumn{2}{c|}{S_n} & Q^n \end{array}\right]$, and

$$\lim_{n\to\infty} S_n = A, \qquad \lim_{n\to\infty} Q_n = 0,$$

$$\lim_{n\to\infty} P^n = \left[\begin{array}{cc|c} 1 & 0 & \multirow{2}{*}{$\mathbf{0}$} \\ 0 & 1 & \\ \hline \multicolumn{2}{c|}{A} & \mathbf{0} \end{array}\right].$$

Need to find $A = A_{3\times 2}$:

Note that for $i \in S_T = \{1, 2, 3\}$, $j \in C_1 \cup C_2 = \{0, 4\}$,

$$\begin{aligned} A(i,j) &= \text{prob that the chain starting at } i \text{ eventually visits } j \\ &= P_i(T_j < \infty) = \rho_{ij}, \end{aligned}$$

$$\rho_{ij} = P(i,j) + \sum_{k\in S_T} P(i,k)\rho_{kj}.$$

Put in matrix form

$$A = S + QA \qquad \therefore \boxed{A = (I - Q)^{-1}S.}$$

Here,

$$S_{3\times 2} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \; Q_{3\times 3} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{bmatrix}, \; \therefore (I-Q)^{-1} = \begin{bmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ 1 & 2 & 1 \\ \frac{1}{2} & 1 & \frac{3}{2} \end{bmatrix}.$$

$$\therefore A = (I - Q)^{-1}S = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}.$$

	0	4	1	2	3
0	1	0	0	0	0
4	0	1	0	0	0
1	$\frac{3}{4}$	$\frac{1}{4}$	0	0	0
2	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
3	$\frac{1}{4}$	$\frac{3}{4}$	0	0	0

$\xrightarrow{\text{reorder}}$

	0	1	2	3	4
0	1	0	0	0	0
1	$\frac{3}{4}$	0	0	0	$\frac{1}{4}$
2	$\frac{1}{2}$	0	0	0	$\frac{1}{2}$
3	$\frac{1}{4}$	0	0	0	$\frac{3}{4}$
4	0	0	0	0	1

$= \lim_{n\to\infty} P^n$.

□

Remark: Such computation also gives us a way to find

$$\rho_{10} = \frac{3}{4},\ \rho_{20} = \frac{1}{2},\ \rho_{30} = \frac{1}{4},$$
$$\rho_{14} = \frac{1}{4},\ \rho_{24} = \frac{1}{2},\ \rho_{34} = \frac{3}{4}.$$

Exercise: (Tutorial)

Modify the above to the MC with

$$P = \left[\begin{array}{cc|c|ccc} \frac{1}{3} & \frac{2}{3} & \multirow{2}{*}{$\mathbf{0}$} & & \multirow{2}{*}{$\mathbf{0}$} & \\ \frac{1}{2} & \frac{1}{2} & & & & \\ \hline \multicolumn{2}{c|}{\mathbf{0}} & 1 & \multicolumn{3}{c}{\mathbf{0}} \\ \hline \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{array}\right]$$

and find $\lim_{n\to\infty} P^n$.

Example 2. Consider the random walk on

$$S = \{0, 1, 2, \cdots\} \quad \text{(no longer finite!)}$$

with

$$P = \begin{bmatrix} q & p & & & \\ q & 0 & p & & \\ & q & 0 & p & \\ & & \ddots & \ddots & \ddots \end{bmatrix}, \quad p, q > 0, p + q = 1.$$

Q.: Find the SD.

Note:

- This is an irreducible BD chain.
- The chain is recurrent iff $\sum_{k=0}^{\infty} (\frac{q}{p})^k = \infty$, iff $q \geq p$.

Solution: Let π be the SD. Set

$$x_k = \pi(k), \quad k = 0, 1, \cdots$$

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

$$[x_0, x_1, \cdots] = [x_0, x_1, \cdots] \begin{bmatrix} q & p & & & \\ q & 0 & p & & \\ & q & 0 & p & \\ & & \ddots & \ddots & \ddots \end{bmatrix}$$

we get

$$\begin{cases} x_0 = qx_0 + qx_1, \quad \text{i.e. } px_0 = qx_1, \\ k \geq 1 : x_k = qx_{k+1} + px_{k-1}. \end{cases}$$

$$\therefore qx_{k+1} - px_k = qx_k - px_{k-1} = \cdots = qx_1 - px_0 = 0$$
$$\therefore x_k = (\frac{p}{q})x_{k-1} = \cdots = (\frac{p}{q})^k x_0,\ k = 0, 1, 2, \cdots$$

(i) $p < q$ (recurrent):

$$1 = \sum_{k=0}^{\infty} x_k = \sum_{k=0}^{\infty} (\frac{p}{q})^k x_0 = \frac{x_0}{1 - \frac{p}{q}} \qquad (0 < \frac{p}{q} < 1)$$

$$\therefore x_0 = \frac{q-p}{q} > 0$$

SD: $\pi = \frac{q-p}{q}[1, \frac{p}{q}, (\frac{p}{q})^2, \cdots].$

(ii) $p = q$ (recurrent): $\sum_{k=0}^{\infty} (\frac{p}{q})^k = \infty$. π does not exist.

(iii) $p > q$ (transient): π does NOT exist. □

Exercise: Modify it to the general irreducible birth & death chain on $S = \{0, 1, \cdots\}$ with

$$P = \begin{bmatrix} r_0 & p_0 & & & \\ q_1 & r_1 & p_1 & & \\ & q_2 & r_2 & p_2 & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \quad \text{row sum} = 1,$$

$$\text{all } p_i > 0, \text{ all } q_i > 0.$$

Q.: Find the SD π.

Example 3. Queueing model:

- In a telephone exchange, ξ_n denotes no of new calls coming in starting at time $n \geqslant 1$. $\{\xi_n\}_{n=1}^{\infty}$ is i.i.d. and has a Poisson distribution with rate $\lambda > 0$:

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \cdots$$

- Suppose that each call has prob $q \stackrel{\text{def}}{=} 1 - p$ to finish in one unit time.

$$X_n \stackrel{\text{def}}{=} \text{no of calls in progress at time } n.$$

Q.: Find the transition prob and the SD.

Solution: To find

$$P(x,y) = P(X_{n+1} = y | X_n = x),$$

we consider

$$X_{n+1} = \xi_{n+1} + Y_{n+1}$$

with $Y_{n+1} \stackrel{\text{def}}{=}$ no of calls at time n that remain at time $n+1$.

Fact:

$$P(Y_{n+1} = z | X_n = x) = \binom{x}{z} p^z (1-p)^{x-z},$$

$$0 \leqslant z \leqslant k.$$

Note: $p =$**non-finish** prob, $q = 1 - p =$**finish** prob.

$$\therefore P(x,y) = P(X_{n+1} = y | X_n = x)$$

$$= \sum_{z=0}^{x\wedge y} P(X_{n+1} = y, Y_{n+1} = z | X_n = x)$$

$$= \sum_{z=0}^{x\wedge y} P(\xi_{n+1} = y - z, Y_{n+1} = z | X_n = x)$$

$$= \sum_{z=0}^{x\wedge y} P(\xi_{n+1}=y-z)P(Y_{n+1}=z|X_n=x)$$

$$= \sum_{z=0}^{x\wedge y} e^{-\lambda} \frac{\lambda^{y-z}}{(y-z)!} \binom{x}{z} p^z (1-p)^{x-z}.$$

To find SD, we will verify that if X_0 is Poisson then X_n $(n \geqslant 1)$ satisfy the same Poisson distribution.

Lemma 1. If X_n is Poisson with rate t, then Y_{n+1} is Poisson with rate pt.

Pf.:

$$\begin{aligned}
P(Y_{n+1}=y) &= \textstyle\sum_{x=y}^{\infty} P(Y_{n+1}=y, X_n=x) \\
&= \sum_{x=y}^{\infty} P(X_n=x)P(Y_{n+1}=y|X_n=x) \\
&= \sum_{x=y}^{\infty} e^{-t}\frac{t^x}{x!}\binom{x}{y}p^y(1-p)^{x-y} \\
&= \frac{(pt)^y e^{-t}}{y!}\sum_{x=y}^{\infty}\frac{[t(1-p)]^{x-y}}{(x-y)!} \\
&= \frac{(pt)^y e^{-t}}{y!}e^{t(1-p)} \\
&= e^{-pt}\frac{(pt)^y}{y!}, \quad y=0,1,2,\cdots \quad \square
\end{aligned}$$

Lemma 2. If X, Y are independent Poisson with rates t_1 and t_2 resp, then $Z = X + Y$ is Poisson with rate $t_1 + t_2$.

Pf.:

$$\begin{aligned}
P(Z = z) &= P(X + Y = z) \\
&= \textstyle\sum_{x=0}^{z} P(X + Y = z, X = x) \\
&= \textstyle\sum_{x=0}^{z} P(X = x, Y = z - x) \\
&= \textstyle\sum_{x=0}^{z} P(X = x)P(Y = z - x) \\
&= \sum_{x=0}^{z} e^{-t_1}\frac{t_1^x}{x!} e^{-t_2}\frac{t_2^{z-x}}{(z-x)!} \\
&= \frac{e^{-(t_1+t_2)}}{z!} \sum_{x=0}^{z} \binom{z}{x} t_1^x t_2^{z-x} \\
&= \frac{e^{-(t_1+t_2)}}{z!}(t_1 + t_2)^z, \quad z = 0, 1, \cdots \quad \square
\end{aligned}$$

Two lemmas above give:

- Assume X_0 is Poisson with rate t (**TBD**).
- $X_1 = \xi_1 + Y_1$ is Poisson with rate

$$\lambda + pt = t. \quad (\therefore t \overset{\text{def}}{=} \frac{\lambda}{1-p} = \frac{\lambda}{q})$$

- $X_2 = \xi_2 + Y_2$ is Poisson with rate $\lambda + pt = t$.
- $\cdots$
- $X_n = \xi_n + Y_n$ is Poisson with rate $\lambda + pt = t$.

$\therefore$ **The chain has a SD** (Poisson, rate$= \lambda/q$):

$$\pi(x) = e^{-\lambda/q}\frac{(\lambda/q)^x}{x!}, \quad x = 0, 1, \cdots . \quad \square$$

Exercise: Check the textbook (Page 55-56) to

(i) Derive an explicit formular $P^n(x, y)$.

(ii) Show directly that

$$\lim_{n\to\infty} P^n(x, y) = \pi(y), \quad \forall\, x, y \geqslant 0.$$

(Hence, π that we have found is the **unique** SD)

Sketch: • The key is to find P^n:

$$\begin{aligned}
X_0 &: \quad t \\
X_1 &: \quad \lambda + tp \\
X_2 &: \quad \lambda + (\lambda + pt)p = tp^2 + \lambda(1 + p^2) \\
X_3 &: \quad \lambda + [tp^2 + \lambda(1 + p^2)]p = tp^3 + \lambda(1 + p^2 + p^3) \\
\dots &\quad \dots \\
X_n &: \quad tp^n + \lambda(1 + p + \cdots + p^n) \\
&\quad = tp^n + \lambda \frac{1 - p^n}{1 - p} := t_n
\end{aligned}$$

then

$$\sum_{x=0}^{\infty} e^{-t} \frac{t^x}{x!} P^n(x, y) = P_x(X_n = y) = e^{-t_n} \frac{t_n^y}{y!}.$$

Rewrite it as

$$\sum_{x=0}^{\infty} \frac{P^n(x,y)}{x!} t^x = e^{-\lambda \frac{1-p^n}{1-p}} e^{t(1-p^n)} \frac{\left[tp^n + \lambda \frac{1-p^n}{1-p}\right]^y}{y!},$$

Apply Taylor expansion and binormial expansion on the right, do the product, and compare coefficient of t^x for each x, then

$$P^n(x,y) = e^{-\lambda \frac{1-p^n}{1-p}} \sum_{z=0}^{\min(x,y)} \binom{x}{z} p^{nz} (1-p^n)^{x-z} \frac{\left[\lambda \frac{1-p^n}{1-p}\right]^{y-z}}{(y-z)!}.$$

Let $n \to \infty$, note $p^n \to 0$ as $0 \leq p < 1$, in $\sum$, except for the term of $z = 0$, all other terms tend to zero, then

$$\lim_{n\to\infty} P^n(x,y) = e^{-\frac{\lambda}{1-p}} \frac{\left(\frac{\lambda}{1-p}\right)^y}{y!} = \pi(y). \quad \square$$

§2.2 Average number of visits

Given $\{X_n\}_{n=0}^{\infty}$, S (finite or infinite),

$$N_n(y) \stackrel{\text{def}}{=} \text{no of visits to } y \ \underline{\text{in } n\text{-steps}},$$

$$\text{i.e. during times } m = 1, 2, \cdots, n.$$

We are interested in the limits of

$$\frac{N_n(y)}{n}, \quad \frac{E_x(N_n(y))}{n} \quad \text{as} \quad n \to \infty.$$

Note:

- $\frac{N_n(y)}{n}$: **proportion** of the first n units of time that the chain visits y, or **average no** of visits to y per unit time.
- $\frac{E_x(N_n(y))}{n}$: **expected proportion** for a chain starting at x, or **frequency** that the chain visits y from x.

It is direct to see

$$N_n(y) = \sum_{m=1}^{n} 1_y(X_m),$$

$$E_x(N_n(y)) = \sum_{m=1}^{n} P^m(x, y).$$

Case: y is transient.

$$\text{Recall: } P_x(N(y) < \infty) = 1.$$

$$\lim_{n\to\infty} N_n(y) = N(y) < \infty \text{ with prob 1,}$$

$$\lim_{n\to\infty} E_x(N_n(y)) = E_x(N(y)) = \frac{\rho_{xy}}{1 - \rho_{yy}} < \infty.$$

So,

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = 0 \quad \text{with prob 1,}$$

$$\lim_{n\to\infty} \frac{E_x(N_n(y))}{n} = 0.$$

Hence, we only consider y **as a recurrent state**.

Let y be **recurrent**. Denote

$$m_y \stackrel{\text{def}}{=} E_y(T_y):$$ the mean return time to y for a chain starting at y.

Recall

$$T_y \stackrel{\text{def}}{=} \min\{n \geqslant 1 : X_n = y\}.$$

Theorem: Suppose

$$\{X_n\}_{n=0}^{\infty} \text{ is } \textbf{irreducible} \text{ and } \textbf{recurrent}.$$

Then

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{1}{m_y} \quad \text{with prob 1,}$$

$$\lim_{n\to\infty} \frac{E_x(N_n(y))}{n} = \frac{1}{m_y}, \quad \forall\, x \in S.$$

Remarks:

(1) Heuristically, the limit is the **frequency** and m_y is the **waiting time**. They are reciprocal to each other.

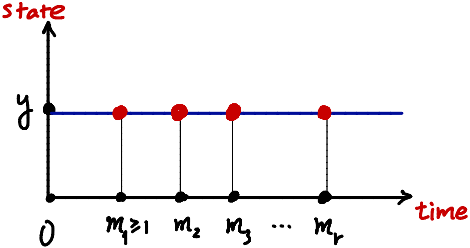
(2) If the chain is NOT irreducible, the statement of Theorem can be modified slightly; see the textbook Pages 58-59.

Pf.: Let the chain start from y. Introduce new r.v.:

$$T_y^r = \min\{n \geqslant 1 : N_n(y) = r\}, \quad r = 1, 2, \cdots$$

i.e. the min ptv-time of the r^{th} visit to y. **Note:**

- $N_n(y) = r$: By time n, the chain visits y for r times. (Warning: time 0 not counted).
- T_y^r: the min positive time up to which the chain visits y for exactly r times.



$$\mathbb{1}_y(X_i) = \begin{cases} 1 & i = m_1, m_2, \cdots, m_r \\ 0 & \text{otherwise} \end{cases}$$

Set

$$W^1 \stackrel{\text{def}}{=} T_y^1 = T_y \text{ (i.e. hitting time of } y)$$

$$W^r \stackrel{\text{def}}{=} T_y^r - T_y^{r-1}, \quad r = 2, 3, \cdots$$

(i.e., waiting time between the $(r-1)^{th}$ visit to y and the r^{th} visit to y)

Then

$$T_y^r = W_y^1 + \cdots + W_y^r, \quad r = 1, 2, \cdots$$

Note: $\{W_y^r\}_{r=1}^{\infty}$ is i.i.d.

(it is intuitively obvious due to the Markov property; see the textbook (page 59) for the rigorous proof)

Apply the **SLLN**, we have

$$\lim_{r\to\infty} \frac{T^r}{r} = \lim_{r\to\infty} \frac{W_y^1+\cdots+W_y^r}{r} = E_y(T_y) \text{ with prob } 1$$
$$= m_y.$$

Next, let $r = N_n(y)$, i.e. by time n, the chain visits y for r-times, and the $(r+1)^{th}$ visit to y will be after n, hence

$$T_y^r \leqslant n < T_y^{r+1},$$

so that

$$\frac{T_y^r}{r} \leqslant \frac{n}{N_n(y)} = \frac{n}{r} < \frac{T_y^{r+1}}{r} \to m_y \quad \text{as } r \to \infty.$$

This implies that $\lim_{n\to\infty} \frac{n}{N_n(y)} = m_y$ with prob 1. □

Moreover, we observe

$$\lim_{n\to\infty} E_x\left(\frac{N_n(y)}{n}\right) = E_x\left(\lim_{n\to\infty}\frac{N_n(y)}{n}\right) \quad \text{(why? DCT)}$$
$$= E_x\left(\frac{1}{m_y}\right)$$
$$= \frac{1}{m_y}. \quad \square$$

Added: **Theorem** (Dominated convergence theorem). Let (ξ_n) be a sequence of rv's and ξ be a rv s.t. for each $\omega \in \Omega$, $\xi_n(\omega) \to \xi(\omega)$ as $n \to \infty$, and there is a rv η such that $|\xi_n| \le \eta$ and $E(\eta) < \infty$. Then

$$E|\xi_n - \xi| \to 0 \text{ as } n \to \infty.$$

Particularly,

$$E(\xi_n) \to E(\xi) \text{ as } n \to \infty. \quad \square$$

Remark: The statement of the theorem can be slightly modified in case when the chain is not irreducible. Indeed, for a general MC, as long as y is recurrent,

$$\frac{N_n(y)}{n} = \frac{\sum_{m=1}^{n} 1_y(X_m)}{n} \to \frac{1_{\{T_y<\infty\}}}{m_y} \text{ as } n \to \infty \text{ with prob } 1,$$

$$E_x(\frac{N_n(y)}{n}) = \frac{\sum_{m=1}^{n} P^m(x,y)}{n} \to \frac{\rho_{xy}}{m_y} \text{ as } n \to \infty,$$

where $1_{\{T_y<\infty\}}$ is a rv meaning that $1_{\{T_y<\infty\}} = 1$ if $T_y < \infty$, and $1_{\{T_y<\infty\}} = 0$ if $T_y = \infty$.

§2.3 Waiting time & stationary distribution

Def.:

- A state x is called **positive recurrent if** it is recurrent and $m_x = E_x(T_x) < \infty$.
- x is called **null recurrent if** it is recurrent and $m_x = E_x(T_x) = \infty$.

Note:

- For a null recurrent sate x,

$$\lim_{n\to\infty} \frac{N_n(x)}{n} = 0 \text{ with prob } 1, \quad \lim_{n\to\infty} \frac{E_x(N_n(x))}{n} = 0.$$

- A positive recurrent state means it comes back in finite waiting time; a null recurrent means it comes back very rarely.

THREE Theorems and THREE Corollaries
are COMING soon.....

no worry

Theorem 1. If x is positive recurrent and $x \to y$, then y is also positive recurrent.

Pf.: $\because x \to y$
$\therefore P^{n_1}(x, y) > 0$ for some $n_1 \geqslant 1$
$\because x$ recurrent, $x \to y$
$\therefore y \to x$, then $P^{n_2}(y, x) > 0$ for some $n_2 \geqslant 1$.
Hence

$$P^{n_2+m+n_1}(y, y) \geqslant P^{n_2}(y, x)P^m(x, x)P^{n_1}(x, y).$$

Sum over $m = 1, 2, \cdots, n$, and divide by n:

$$\frac{E_y(N_{n_2+n+n_1}(y)) - E_y(N_{n_2+n_1}(y))}{n} \geqslant P^{n_2}(y, x)\frac{E_x(N_n(x))}{n}P^{n_1}(x, y).$$

Take limit $n \to \infty$:

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

$\therefore m_y < \infty$, i.e. y is positive recurrent. □

Theorem 2. An irreducible MC having a finite number of states must be positive recurrent.

Pf.: We know that all states are recurrent ($\because$ finite state + irreducible).

Assuming that the theorem is false, all states are null recurrent. Note

$$1 = \sum_{y \in S} P^m(x, y) \qquad \text{(row sum is 1).}$$

Sum over $m = 1, \cdots, n$ and divide by n:

$$1 = \sum_{y \in S} \frac{E_x(N_n(y))}{n}.$$

Take limit:

$$\begin{aligned} 1 &= \lim_{n\to\infty} \sum_{y\in S} \frac{E_x(N_n(y))}{n} \\ &= \sum_{y\in S} \lim_{n\to\infty} \frac{E_x(N_n(y))}{n} \quad \text{(}S\text{ is finite)} \\ &= \sum_{y\in S} 0 \\ &= 0, \ \text{contradiction!} \quad \square \end{aligned}$$

Theorem 3. An irreducible positive recurrent MC has a unique SD π given by

$$\pi(x) = \frac{1}{m_x}, \quad x \in S.$$

Pf.: Step 1. Uniqueness.

We first assume the SD exists, denoted by π, to show $\pi(x) = \frac{1}{m_x}$, $x \in S$. In fact,

$$\pi(x) = \sum_z \pi(z) P^m(z, x) \quad (i.e., \pi = \pi P^m, \forall\, m \geqslant 1)$$

Sum over $m = 1, \cdots, n$ and divide by n:

$$\pi(x) = \sum_z \pi(z) \frac{E_z(N_n(x))}{n}.$$

Take limit:

$$\begin{aligned}
\pi(x) &= \lim_{n\to\infty} \sum_z \pi(z) \frac{E_z(N_n(x))}{n} \\
&= \sum_z \pi(z) \lim_{n\to\infty} \frac{E_z(N_n(x))}{n} \ \big(\text{infinite sum need DCT}\big) \\
&= \sum_z \pi(z) \frac{1}{m_x} \\
&= \frac{1}{m_x}.
\end{aligned}$$

Therefore, the uniqueness follows.

Added: **Dominated Convergence Theorem:** Suppose

(i) $|a_n(k)| \leqslant M < \infty$, $\lim_{n\to\infty} a_n(k) = a(k)$.

(ii) $\sum_{k=1}^{\infty} p_k = 1$ (or just $< \infty$)

Then

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} a_n(k)p(k) = \sum_{k=1}^{\infty} a(k)p(k).$$

(e.g. $\lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{1}{n+k}$ exists or not?)

Pf.: Apply ϵ-N argument to

$$\underbrace{\sum_{k=1}^{N} a_n(k)p_k}_{\text{(I)}} + \underbrace{\sum_{k=N+1}^{\infty} a_n(k)p_k}_{\text{(II)}}$$

(II)$\leq \epsilon/2$ for a large N.

(I): can be close to $\sum_{k=1}^{N} a(k)p(k)$ as long as n is large! □

Step 2. **Existence.**

To show existence, it suffices to show

(i) $\sum_{x \in S} \frac{1}{m_x} = 1.$ (distribution)

(ii) $\sum_{x \in S} \frac{1}{m_x} P(x, y) = \frac{1}{m_y}, \forall\, y.$ (stationary)

Step 2.1 To show: (i) (ii) are two <u>inequalities</u> "$\leqslant$".

- Note: $\sum\limits_x P^m(z,x) = 1, \ \forall\, z$. Then

$$\xrightarrow{\frac{\sum_{m=1}^{n}(\cdots)}{n}}\ \sum_{x\in S} \frac{E_z(N_n(x))}{n} = 1 \text{ (if } S \text{ is infinite, why? Fubini!)}$$

(It will be direct if we take limit on n then "$=$" will follow, however, we cannot apply DCT here (why?) we need a slight modification)

$$\Rightarrow \sum_{x\in S_1} \frac{E_z(N_n(x))}{n} \leqslant 1, \forall\, S_1 \text{ finite}$$

$$\Rightarrow \sum_{x\in S_1} \frac{1}{m_x} \leqslant 1, \forall\, S_1 \text{ finite} \Rightarrow \sum_{x\in S} \frac{1}{m_x} \leqslant 1.$$

- Note

$$\sum_{x\in S} P^m(z,x)P(x,y) = P^{m+1}(z,y).$$

$\sum_{m=1}^{n}(\cdots)/n \Rightarrow$

$$\sum_{x\in S} \frac{E_z(N_n(x))}{n} P(x,y) = \frac{E_z(N_{n+1}(y))}{n} - \frac{P(z,y)}{n}.$$

$$\Rightarrow \sum_{x\in S_1} \frac{E_z(N_n(x))}{n} P(x,y) \leqslant \frac{E_z(N_{n+1}(y))}{n} - \frac{P(z,y)}{n}, \forall\, S_1 \text{ finite}$$

$$\Rightarrow \sum_{x\in S_1} \frac{1}{m_x} P(x,y) \leqslant \frac{1}{m_y}, \forall\, S_1 \text{ finite}$$

$$\Rightarrow \sum_{x\in S} \frac{1}{m_x} P(x,y) \leqslant \frac{1}{m_y}.$$

Step 2.2 To show: (i) (ii) are two equalities "=".

To show: (ii) $\sum_{x \in S} \frac{1}{m_x} P(x,y) = \frac{1}{m_y}, \ \forall\, y.$

Otherwise, $\exists\, y_0$ s.t.

$$\sum_{x \in S} \frac{1}{m_x} P(x, y_0) < \frac{1}{m_{y_0}}.$$

Then

$$\begin{aligned} 1 \geq \sum_{y \in S} \frac{1}{m_y} &> \sum_{y \in S} \left[\sum_{x \in S} \frac{1}{m_x} P(x,y) \right] \\ &= \sum_{x \in S} \frac{1}{m_x} \left[\sum_{y \in S} P(x,y) \right] \qquad \text{(Use Fubini)} \\ &= \sum_{x \in S} \frac{1}{m_x} \qquad \text{a contradiction!} \end{aligned}$$

To show (i): $\sum_{x\in S} \frac{1}{m_x} = 1$.

Note $\sum_{x\in S} \frac{1}{m_x} \leq 1$. Let c be such that $\sum_{x\in S} \frac{c}{m_x} = 1$. Then

$$\pi(x) = \frac{c}{m_x}, \quad x \in S$$

is a SD. Now, by uniqueness

$$\frac{c}{m_x} = \frac{1}{m_x}, \qquad \forall x \in S.$$

$\therefore c = 1$. So $\sum_{x\in S} \frac{1}{m_x} = 1$, i.e. (i) follows. □

Corollary 1. An irreducible MC with finite state space has a unique SD:

$$\pi(x) = \frac{1}{m_x}, \quad x \in S.$$

e.g.: P (finite Matrix). $\pi P = \pi$. We then solve

$$(P^T - I)\pi^T = 0$$

though the row operation. **Cor 1** says that

- the solution exists and is unique.
- it gives us a way to find $m_x = E_x(T_x)$:

$$m_x = \frac{1}{\pi(x)}, \quad x \in S.$$

$$(\because\ m_x < \infty \ \therefore\ \pi(x) > 0)$$

Corollary 2. Let the chain be irreducible, then the chain has a SD **iff** it is positive recurrent!

Pf.: "$\Leftarrow$": It's just the theorem.

"$\Rightarrow$": Otherwise, all states are either null recurrent or transient (why?), then in both cases,

$$\lim_{n\to\infty} \frac{\sum_{m=1}^{n} P^m(z,x)}{n} = 0, \quad \forall\, z, x \in S.$$

Let π be the SD. Take $x \in S$, then

$$\pi(x) = \sum_z \pi(z) P^m(z,x).$$

$$\sum_{m=1}^{n}(\cdots)/n \Rightarrow \pi(x) = \sum_z \pi(z) \frac{E_z(N_n(x))}{n}.$$

$$n \to \infty + \text{DCT} \Rightarrow \pi(x) = \sum_z \pi(z) \cdot 0 = 0.$$ Contradiction! □

e.g.: $P = \begin{bmatrix} q & p & & & \\ q & 0 & p & & \\ & q & 0 & p & \\ & & \ddots & \ddots & \ddots \end{bmatrix}$, $S = \{0, 1, 2, \cdots\}$ is infinite.

Assume: $p > 0$, $q > 0$, $p + q = 1$ (irreducible).

Recall:

- This chain is recurrent **iff** $q \geqslant p$.
- The chain has a SD **iff** $q > p$.

Then, the chain is positive recurrent **iff**

$$q > p.$$

(Once again, in this case, $E_x(T_x) = m_x = \frac{1}{\pi(x)}$.)

Exercise: Consider a general birth & death chain

$$P = \begin{bmatrix} r_0 & p_0 & & & \\ q_1 & r_1 & p_1 & & \\ & q_2 & r_2 & p_2 & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \qquad \text{Row sum} = 1$$

Assume it is irreducible.

Q.: Determine if it is either positive recurrent, null recurrent, or transient.

Corollary 3. Let C be an irreducible closed set of positive recurrent states. Then the MC has a unique SD π concentrated on C:

$$\pi(x) = \begin{cases} \frac{1}{m_x} & x \in C, \\ 0 & \text{Otherwise.} \end{cases}$$

Indeed, we can regard $\{X_n\}$ as a MC on C and obtain $\pi_C(x) = \frac{1}{m_x}$, $x \in C$. Define

$$\pi(x) = \begin{cases} \pi_C(x) & x \in C, \\ 0 & \text{Otherwise.} \end{cases}$$

Then it is direct to check that π is a SD on S.

e.g.: Let $S = (C_1 \cup \cdots) \cup S_T$ (finite or ∞), and

$$P = \begin{array}{c} \\ C_1 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cc} C_1 & \cdots \end{array} \\ \begin{bmatrix} P_1 & 0 \\ * & * \end{bmatrix} \end{array}, \quad C_1 \text{ positive recurrent.}$$

Regard $\{X_n\}_{n=0}^{\infty}$ as a MC on C_1. Then, by Thm 3, $\pi_{C_1}(x) = \frac{1}{m_x}$ $(x \in C_1)$ is the SD. Define

$$\pi(x) = \begin{cases} \pi_{C_1}(x) & \text{if } x \in C_1, \\ 0 & \text{Otherwise.} \end{cases}$$

We may write $\pi = [\pi_{C_1}, 0]$. Check:

$$\pi P = [\pi_{C_1}, 0] \begin{bmatrix} P_1 & 0 \\ * & * \end{bmatrix} = [\pi_{C_1} P_1, 0] = [\pi_{C_1}, 0] = \pi,$$

i.e. π is a SD of P.

Two further notes:

- If no C_i is positive recurrent (i.e. all states in S are either transient or null recurrent), then the chain has no SD.
- Let $P = \begin{array}{c} \\ C_1 \\ C_2 \\ S_T \end{array} \begin{array}{c} \begin{array}{ccc} C_1 & C_2 & S_T \end{array} \\ \begin{bmatrix} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ * & * & * \end{bmatrix} \end{array}$,

  C_i $(i = 1, 2)$: positive recurrent,
  π_i $(i = 1, 2)$: SD of P_i concentrated on C_i.
  Then

  $$\pi \overset{\text{def}}{=} \lambda \pi_1 + (1 - \lambda)\pi_2, \ 0 \leqslant \lambda \leqslant 1$$

  is also the SD of P.

§2.4 Periodicity

Recall: For an irreducible & positive recurrent MC,

$$\lim_{n\to\infty} \frac{\sum_{m=1}^{n} p^m(x,y)}{n} = \lim_{n\to\infty} \frac{E_x(N_n(y))}{n} = \frac{1}{m_y} = \pi(y), \ \forall\, y \in S,$$

i.e. $$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m = \begin{bmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{bmatrix}$$ exist

(S: finite or infinite)

Q.: How about $\lim_{n\to\infty} P^n$?

Example: $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, SD: $\pi = [\frac{1}{2}, \frac{1}{2}]$. Note:

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

$\therefore \lim_{n\to\infty} P^n$ does NOT exist.

BUT, both $\lim_{n\to\infty} P^{2n}$ and $\lim_{n\to\infty} P^{2n+1}$ exist!

The problem is on the "**periodicity**" of the chain.

Definition. The period d_x of a state x is the **greatest common divisor** (**g.c.d.**) of

$$\{n \geqslant 1 : P^n(x,x) > 0\}.$$

Remarks:

(i) $1 \leqslant d_x \leqslant \min\{n \geqslant 1, P^n(x,x) > 0\}$.

(ii) If $P(x,x) > 0$ then $d_x = 1$.

(iii) For Example above, $d_0 = 2 = d_1$. Indeed, note:

$$1 = P^2(0,0) = P^4(0,0) = \cdots = P^{2n}(0,0) = \cdots \,,$$
$$0 = P^1(0,0) = P^3(0,0) = \cdots = P^{2n+1}(0,0) = \cdots \,,$$

$$\therefore \text{ g.c.d.}\{n \geq 1 : P^n(0,0) > 0\} = \text{g.c.d.}\{2, 4, \cdots\} = 2.$$

Prop. For an irreducible MC, all d_x are equal.

Pf.: T ake $x, y \in S$.
$\because$ The chain is irreducible
$\therefore x \to y$ & $y \to x$,
i.e. $\exists n_1 \geq 1, n_2 \geq 1$ s.t. $P^{n_1}(x,y) > 0$, $P^{n_2}(y,x) > 0$
So

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

$\therefore d_x | n_1 + n_2$ $(*)$ (i.e., d_x is a divisor of $n_1 + n_2$)
Let $A_y \overset{\text{def}}{=} \{n \geqslant 1 : P^n(y,y) > 0\}$. Then, for $n \in A_y$,

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

$\therefore d_x | n_1 + n + n_2$ Note: $n = (n_1 + n + n_2) - (n_1 + n_2)$
Together with $(*) \Rightarrow d_x | n$, $\forall\, n \in A_y$.

$$\therefore d_x | d_y$$

The same argument gives $d_y | d_x$. $\therefore d_x = d_y$. □

Definition: Consider an **irreducible** MC.

- Note that all states have

  $$\textbf{the same} \text{ period } d \geqslant 1.$$

  The chain is called **periodic** with period $d \geqslant 1$.

- If $d = 1$, we say the chain is **aperiodic**.

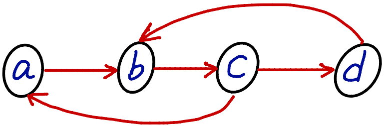
Remark: Consider an irreducible MC. If

$$P(x, x) > 0 \text{ for some } x \in S,$$

then the chain must be aperiodic. ($\because$ $d_x = 1 = d$)

Example 1.

$$P = \begin{bmatrix} 0 & \times & 0 & 0 \\ 0 & 0 & \times & 0 \\ \times & 0 & 0 & \times \\ 0 & \times & 0 & 0 \end{bmatrix}, \quad \times\text{: nonzero entries.}$$



It is obvious to see that the chain is irreducible, and

$$d_a = 3,$$

(**Note:** $d_a = 3$ means that the chain from a returns to a in $3m$ steps, i.e. $P^{3m}(a,a) > 0$, $\forall\, m \geqslant 1$.)

$\therefore$ **Period** $= 3.$

We may directly compute: P^n, $(n = 2, 3, 4, \cdots)$.
For $m = 1, 2, \cdots$,

$$P^{3m} = \begin{bmatrix} \times & 0 & 0 & \times \\ 0 & \times & 0 & 0 \\ 0 & 0 & \times & 0 \\ \times & 0 & 0 & \times \end{bmatrix}, \quad P^{3m+1} = \begin{bmatrix} 0 & \times & 0 & 0 \\ 0 & 0 & \times & 0 \\ \times & 0 & 0 & \times \\ 0 & \times & 0 & 0 \end{bmatrix}$$

$$P^{3m+2} = \begin{bmatrix} 0 & 0 & \times & 0 \\ \times & 0 & 0 & \times \\ 0 & \times & 0 & 0 \\ 0 & 0 & \times & 0 \end{bmatrix}.$$

Recall: $d_x = \text{g.c.d. } \{n \geqslant 1 : P^n(x, x) > 0\}$.

$\therefore$ **Period** $= 3$.

Example 2. Determine the **period** of an irreducible birth and death chain:

$$P = \begin{bmatrix} r_0 & p_0 & & & \\ q_1 & r_1 & p_1 & & \\ & q_2 & r_2 & p_2 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}, \quad \text{all } p_x > 0, q_x > 0.$$

- If **some** $r_x > 0$, then $P(x,x) = r_x > 0$, hence the chain is aperiodic.
- If **all** $r_x = 0$, then the chain can return to its initial state ONLY after an even number of steps.
  Then, for a given state $x \in S$, any integer $n \geq 1$ such that $P^n(x,x) > 0$ must be even.
  Then $d \geqslant 2$ must be even.
  Note $P^2(0,0) = P(0,1)P(1,0) = p_0 q_1 > 0$.
  $\therefore$ **Period** $= 2$.

Theorem. Let $\{X_n\}_{n=0}^{\infty}$ be irreducible and positive recurrent with SD π.

(i) **If** the chain is **aperiodic**, **then**

$$\lim_{n\to\infty} P^n(x,y) = \pi(y), \quad \forall\, x, y \in S.$$

(ii) **If** the chain is **periodic** with period $d \geqslant 2$, **then** for any $x, y \in S$, there exists

$$r \in \{0, 1, 2, \cdots, d-1\}$$

which may depend on x and y, s.t.

$$P^n(x,y) = \begin{cases} \xrightarrow[m\to\infty]{} d\pi(y) & \text{if } n = md + r, \\ = 0 & \text{if } n \neq md + r, \end{cases}$$

where $m \geqslant 0$ is an integer.

Pf.: Pages 75-80 in the textbook. □

Remark: Theorem tells that in case

$$\textbf{Period} = d \geq 2,$$

we are able to determine the limits of

$$P^{md}, \quad P^{md+1}, \cdots, P^{md+(d-1)} \quad (m \to \infty).$$

Precisely, for any given x, y,

$$P^{md}(x,y), \quad P^{md+1}(x,y), \cdots, P^{md+(d-1)}(x,y)$$

are zeros, **except that**

exactly one of them tends to $d\pi(y)$ as $m \to \infty$.

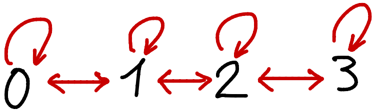
You have to figure out which one!

Example 3. Determine the long term behavior of P^n for given P.

$$\textbf{(a)}\ P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{array}.$$

Solution:

- Note:



$\therefore$ irreducible.

- $\exists\, x$ s.t. $P(x,x) > 0$. $\therefore$ Period $= 1$.
- Solving $\pi = \pi P$, we get the ! SD

$$\pi = [\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}].$$

- Hence, by the theorem,

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi \\ \pi \\ \pi \\ \pi \end{bmatrix} = \begin{bmatrix} \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{bmatrix}. \quad \square$$

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

Solution:

- Note:

$$0 \leftrightarrow 1 \leftrightarrow 2 \leftrightarrow 3$$

$\therefore$ irreducible.

- Period $= 2$. (By the previous example)
- Solving $\pi = \pi P$, we get the ! SD: $\pi = [\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}]$.

$$0 \leftrightarrow 1 \leftrightarrow 2 \leftrightarrow 3$$

By the theorem,

if $x - y$ is even, $\begin{cases} P^{2m+1}(x,y) = 0, & \forall\, m, \\ P^{2m}(x,y) \xrightarrow[m\to\infty]{} 2\pi(y). \end{cases}$

If $x - y$ is odd, $\begin{cases} P^{2m}(x,y) = 0, & \forall\, m, \\ P^{2m+1}(x,y) \xrightarrow[m\to\infty]{} 2\pi(y). \end{cases}$

Recall: $\pi = [\frac{1}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{8}]$.

$$\therefore \lim_{m\to\infty} P^{2m} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{3}{4} & 0 & \frac{1}{4} \end{bmatrix} \end{array}, \quad \lim_{m\to\infty} P^{2m} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \end{bmatrix} \end{array}.$$

Chapter 3:
Markov Jump Process

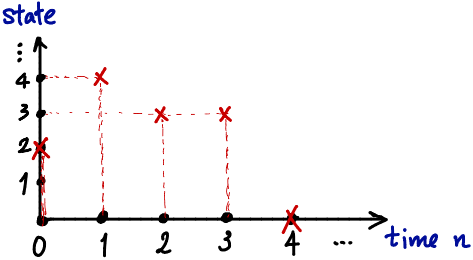
§3.1 Introduction

- **Jump process.**

Recall: a MC (**discrete-time** stochastic process with the Markovian property):

$$X(n) \in S, \quad n = 0, 1, 2, \cdots .$$

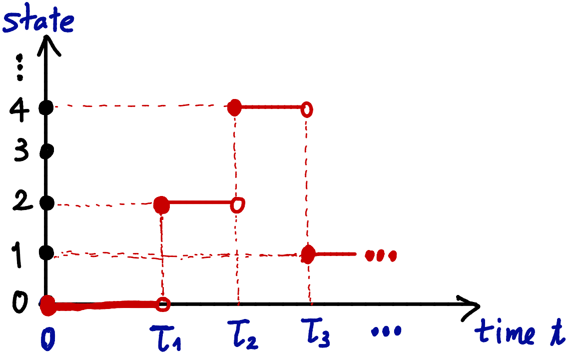
S: finite or countably infinite,
e.g. $S = \{0, 1, ..., N\}$ $(N \leq \infty)$.



Consider a **continuous-time** stochastic process:

$$X(t) \in S, \quad 0 \le t < \infty,$$

S: finite or countably infinite.



– $\tau_1, \tau_2, \cdots$: the waiting time to jump (random).

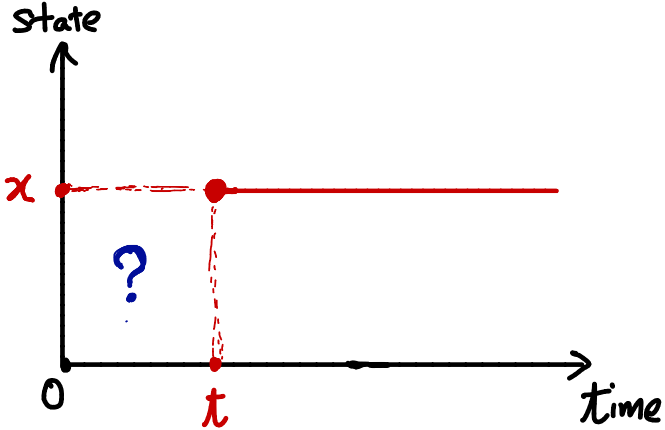
– $X(\tau_1), X(\tau_2), \cdots$: where to jump (random).

– Always assume: $\lim_{n\to\infty} \tau_n = \infty$ (No blow-up!)

- **Probability structure.**

Def.: $x \in S$ is **absorbing if**

"$X(t) = x$ for some $t \geqslant 0$" $\Rightarrow$ "$X(s) = x, \forall\, s \geqslant t$".



Given a **non-absorbing** state $X(0) = x \in S$, we need to know two things:

(i) $F_x(t)$, $t \geqslant 0$: the distribution of the waiting time τ_1. Note:

$$F_x(t) = P_x(\tau_1 \leq t).$$

(ii) Q_{xy}: the transition prob to jump from a state x to another state $y(\neq x)$:

$$Q_{xx} = 0, \quad \sum_{y \in S} Q_{xy} = 1.$$

(If x is **absorbing**, $Q_{xy} = \delta_{xy} = \begin{cases} 1, \text{ for } x = y, \\ 0, \text{ otherwise.} \end{cases}$)

For **non-absorbing** x, we assume:

$$P_x(\tau_1 \leqslant t, X(\tau_1) = y) = P_x(\tau_1 \leqslant t) Q_{xy},$$

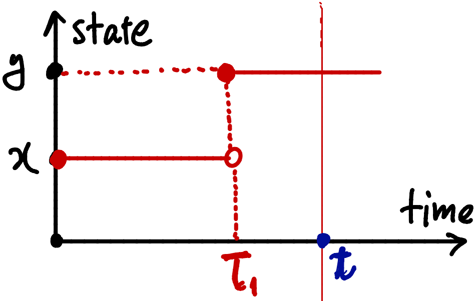
i.e.

τ_1 (**the waiting time to jump**)

and

$X(\tau_1)$ (**jump to where**)

are **independent**!



Similar to the MC (discrete-time), **our concern** is to determine the **transition function**:

$$\boxed{P_{xy}(t) \overset{\text{def}}{=} P(X(t) = y | X(0) = x) = P_x(X(t) = y),}$$

i.e., the prob that the process starting at x will be at y at time $t \geqslant 0$.

Note:

(i) $\sum_{y} P_{xy}(t) = 1$, $P_{xy}(0) = \delta_{xy}$.

(ii) If initial distribution is known, for instance, it is given by $\pi_0(x)$, $x \in S$, then

$$P(X(t) = y) = \sum_{x \in S} \pi_0(x) P_{xy}(t),$$

or $\pi_t = \pi_0 P(t)$ in matrix form.

- **Markov property:**

$$P(X(t) = y | X(s_1) = x_1, \cdots, X(s_n) = x_n, X(s) = x)$$
$$= P(X(t) = y | X(s) = x),$$
$$\forall\, 0 \leqslant s_1 \leqslant \cdots \leqslant s_n \leqslant s \leqslant t, \forall\, x_1, \cdots, x_n, x, y \in S.$$

Note:

- We always assume the process is **time-homogeneous**:

$$P(X(t) = y | X(s) = x) = P(X(t-s) = y | X(0) = x),$$
$$\forall\, 0 \leqslant s \leqslant t, \forall\, x, y \in S.$$

Therefore

$$P(X(t) = y | X(s) = x) = P_{xy}(t-s).$$

- A Markov jump process (**MJP**) $\stackrel{\text{def}}{=}$ a continuous-time jump process with the Markovian property.

Now, we always consider the **MJP**.

Q.: How to determine $F_x(t) = P_x(\tau_1 \leq t)$?

Recall that $F_x(t)$ is the distribution of τ_1 (the waiting time for a jump to occur!).

To show: τ_1 is an **exponential** rv with density:

$$f(t) = \lambda e^{-\lambda t}, \quad t \geq 0; \quad \lambda \stackrel{\text{def}}{=} \frac{1}{E(\tau_1)}.$$

Hence:

$$F_x(t) = P(\tau_1 \leq t) = \int_{-\infty}^{t} f(s)\, ds = 1 - e^{-\lambda t}, \; t \geq 0.$$

Def.: Let τ be a r.v. taking values in $[0, \infty)$. Then τ is said to be **memoryless if**

$$P(\tau > s + t | \tau > s) = P(\tau > t), \quad \forall\, s, t \geqslant 0,$$

(i.e., after waiting for time s, the prob for waiting for another time t has no memory that it already waits for time s.)

e.g. Model: Wait for an unreliable bus driver. Then, the waiting time is a memoryless r.v.:

> "If we have been waiting for s units of time then the prob we must wait t more units of time is the same as if we have not waited at all!"

Proposition. Let τ be a **memoryless** r.v. Then τ is an **exponential** r.v., and the density is given by

$$\lambda e^{-\lambda t}, \ t \geqslant 0; \quad \lambda = 1/E(\tau).$$

Pf.: Let $G(t) \stackrel{\text{def}}{=} P(\tau > t)$. As τ is memoryless,

$$\begin{aligned} G(t) = P(\tau > t) &= P(\tau > s + t | \tau > s) \\ &= \frac{P(\tau > s + t)}{P(\tau > s)} = \frac{G(s + t)}{G(s)}, \end{aligned}$$

i.e.

$$G(s + t) = G(s)G(t), \quad \forall \, s, t \geqslant 0.$$

Assuming G is differentiable,

$$\begin{aligned} G'(t) &= \lim_{h \to 0_+} \frac{G(t+h) - G(t)}{h} \\ &= \lim_{h \to 0_+} \frac{G(t)G(h) - G(t)}{h} \\ &= G(t) \lim_{h \to 0_+} \frac{G(h) - 1}{h} \\ &\stackrel{\text{def}}{=} G(t)\alpha. \end{aligned}$$

Note: $G(0) = P(\tau > 0) = 1$. $\therefore$ $G(t) = e^{\alpha t}$.

Note: $G(t) = P(\tau > t)$ is decreasing. $\therefore \alpha < 0$. Set $\alpha = -\lambda$ $(\lambda > 0)$. The density function is

$$f(t) = (1 - G(t))' = \lambda e^{-\lambda t}. \quad \square$$

Proposition. Let $X(t)$, $t \geqslant 0$ be a MJP. For a non-absorbing state $x \in S$, letting $X(0) = x$,

$$\tau_x \stackrel{\text{def}}{=} \inf\{t > 0 : X(t) \neq x\}. \quad \text{(first time to jump)}$$

Then, τ_x is a memoryless r.v.

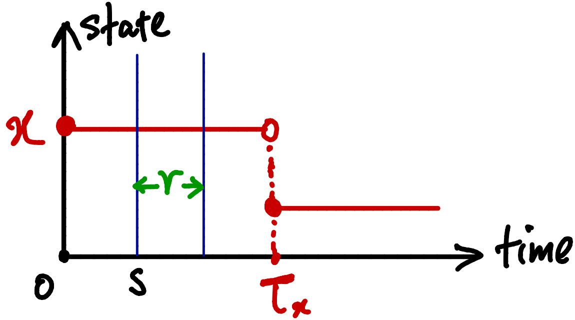
Pf.:

$$\begin{aligned}
&P(\tau_x > s + r | \tau_x > s) \\
&= P(X(t) = x, 0 \leqslant t \leqslant s + r | X(t) = x, 0 \leqslant t \leqslant s) \\
&= P(X(t) = x, s \leqslant t \leqslant s + r | X(t) = x, 0 \leqslant t \leqslant s) \\
&= P(X(t) = x, s \leqslant t \leqslant s + r | X(s) = x) \text{ (Markovian)} \\
&= P(X(t) = x, 0 \leqslant t \leqslant r | X(0) = x) \text{ (time-homog)} \\
&= P(\tau_x > r). \quad \square
\end{aligned}$$

Remarks:

- For a MJP, as τ_x is memoryless:

$$P(\tau_x > s + r | \tau_x > s) = P(\tau_x > r),$$



it looks like that the process starts from s.

- Set $q_x \stackrel{\text{def}}{=} 1/E(\tau_x)$. Then, τ_x has an exponential density given by $q_x e^{-q_x t}$ $(t \geq 0)$.

§3.2 Poisson process

We shall give the definition of **Poisson process** in terms of the **waiting time**.

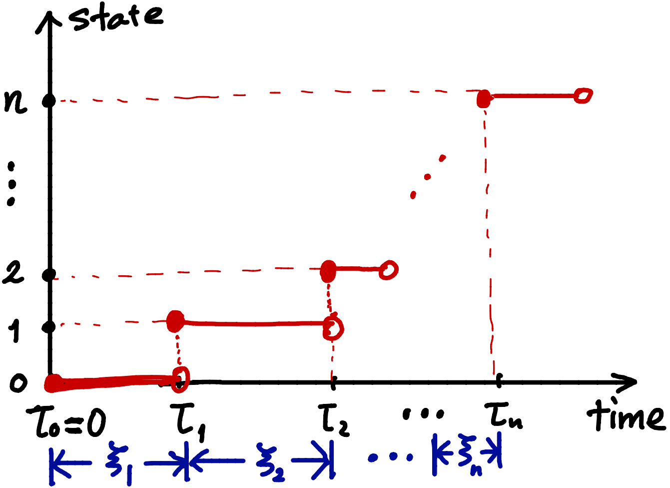
Setup:

• Let $\xi_n \sim \xi$, $n = 1, 2, \cdots$, be i.i.d. exp. r.v. with parameter λ:

$$P(\xi > t) = e^{-\lambda t}, \quad \lambda = 1/E(\xi).$$

• Define $\tau_0 = 0$, and

$$\tau_n \stackrel{\text{def}}{=} \xi_1 + \xi_2 + \cdots + \xi_n, \quad n = 1, 2, \cdots$$



For $n = 1, 2, \cdots,$

$\xi_n \sim \xi$: the waiting time for one arrival.

τ_n: the waiting time for the n^{th}-arrival.

For $t \geqslant 0$,

$$X(t) \stackrel{\text{def}}{=} \max\{n \geqslant 0, \tau_n \leqslant t\},$$

i.e., the no of arrival in $[0, t]$.

Then, we get a jump process:

$$X(t) \in \{0, 1, 2, \cdots\}, \quad t \geqslant 0.$$

Q.:

- **What's the density of $X(t)$? (Poisson with rate λt!)**
- **Is $X(t)$ a MJP? (YES!)**

Theorem. $X(t)$ is Poisson with $E(X(t)) = \lambda t$:

$$P(X(t) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2 \cdots .$$

Pf.: By definition,

$$\{X(t) = n\} = \{\tau_n \leqslant t < \tau_{n+1}\} = \{\tau_{n+1} > t\} \backslash \{\tau_n > t\}.$$

Hence,

$$P(X(t) = n) = P(\tau_{n+1} > t) - P(\tau_n > t). \qquad (*)$$

- $n = 0$:

$$P(X(t) = 0) = P(\tau_1 > t) - 0 = P(\xi_1 > t) = e^{-\lambda t}.$$

- **To show:**

$$P(\tau_n > t) = e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}, \qquad (**)$$

$$n = 1, 2, \cdots .$$

If so, substituting $(**)$ into $(*)$ gives the theorem.

Proof of $(**)$ by induction:

$n = 1$: $P(\tau_1 > t) = P(\xi_1 > t) = e^{-\lambda t}$. $(**)$ holds.

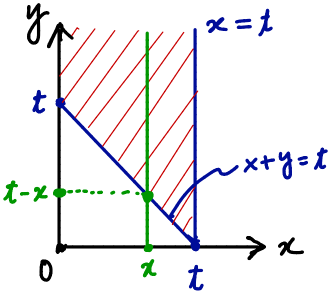
Letting $(**)$ hold for $n \geq 1$, we need to show that $(**)$ is true for $n + 1$. Indeed,

$$
\begin{aligned}
&P(\tau_{n+1} > t) \\
&= P(\tau_n + \xi_{n+1} > t) \\
&= P(\xi_{n+1} > t) + P(\xi_{n+1} \le t, \tau_n + \xi_{n+1} > t) \\
&= e^{-\lambda t} + \int_0^t \lambda e^{-\lambda s} \cdot P(\tau_n > t - s)\, ds \text{ (explain later)} \\
&= e^{-\lambda t} + \int_0^t \lambda e^{-\lambda s} \cdot \sum_{k=0}^{n-1} e^{-\lambda(t-s)} \frac{(\lambda(t-s)^k)}{k!}\, ds \\
&\qquad\qquad\qquad \text{(Use induction assumption!)} \\
&= e^{-\lambda t} + e^{-\lambda t} \sum_{k=0}^{n-1} \frac{\lambda^{k+1}}{k!} \cdot \int_0^t (t-s)^k ds \\
&= e^{-\lambda t} + e^{-\lambda t} \sum_{k=0}^{n-1} \frac{\lambda^{k+1}}{k!} \cdot \frac{t^{k+1}}{(k+1)} \\
&= e^{-\lambda t} + e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^{k+1}}{(k+1)!} = e^{-\lambda t} \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!}. \quad \square
\end{aligned}
$$

Note (See Durrett P93-103):

Let X, Y be independent with densities $f(\cdot)$, $g(\cdot)$ over $[0, \infty)$, resp. Then,

$$P(X < t, X + Y > t) = \int_0^t \int_{t-x}^{\infty} f(x)g(y)\, dydx$$
$$= \int_0^t f(x) \int_{t-x}^{\infty} g(y)\, dydx = \int_0^t f(x)P(Y > t - x)dx.$$



Remarks:

- Note:

$$P(X(0)=k)=\delta_{0k}=\begin{cases}1 & k=0,\\ 0 & \text{otherwise.}\end{cases}$$

$$\begin{aligned}\therefore\ P(X(t)=n) &= \sum_{k=0}^{\infty} P(X(0)=k)P(X(t)=n|X(0)=k)\\ &= \sum_{k=0}^{\infty}\delta_{0k}P_{kn}(t)\\ &= P_{0n}(t).\end{aligned}$$

$$\therefore\ P_{0n}(t)=e^{-\lambda t}\frac{(\lambda t)^n}{n!},\quad n=0,1,2\cdots.$$

- $E(X(t))=\lambda t$ is the **expected no of arrivals in $[0,t]$**. λ is the **arrival rate**.

Corollary. The **Poisson process** $\{X(t)\}_{t\geqslant 0}$ with rate λt satisfies:

(i) $X(0) = 0$.

(ii) For $0 < s < t$, $X(t) - X(s)$ has Poisson distribution with mean $\lambda(t-s)$, and is independent of $X(s)$.

(iii) For $0 \leqslant t_1 \leqslant \cdots \leqslant t_n$,

$$X(t_2) - X(t_1), \cdots, X(t_n) - X(t_{n-1})$$

are independent.

Also, $\{X(t)\}_{t\geqslant 0}$ satisfies the Markov property with

$$E(X(t)) = \lambda t, \quad Var(X(t)) = \lambda t.$$

Remark: Very often, (i)(ii)(iii) are also used as the definition of Poisson process!

IDEA of Proof:

- (i): Obvious.
- (ii): For $0 < s < t$,

$$
\begin{aligned}
&P(X(t) - X(s) = n) \\
&= \sum_{m=0}^{\infty} P(X(s) = m, X(t) = n + m) \\
&= \sum_{m=0}^{\infty} P(X(s) = m) P(X(t) = n + m | X(s) = m) \\
&= \sum_{m=0}^{\infty} P(X(s) = m) P_{m,n+m}(t - s) \\
&= \sum_{m=0}^{\infty} P(X(s) = m) P_{0,n}(t - s) \\
&= P_{0,n}(t - s) \\
&= e^{-\lambda(t-s)} \frac{[\lambda(t - s)]^n}{n!}.
\end{aligned}
$$

$X(t) - X(s)$ is independent of $X(s)$ means

$$P(X(t)-X(s)=n, X(s)=m) = P(X(t)-X(s)=n)P(X(s)=m),$$

equivalently

$$P(X(t)-X(s)=n|X(s)=m) = P(X(t)-X(s)=n).$$

Indeed, note

$$\mathrm{LHS} = P(X(t)=m+n|X(s)=m) = P_{m,m+n}(t-s) = P_{0,n}(t-s).$$

- (iii): Omit the proof. Intuitively clear (See P94-95 in Durrent Chapter 3)

- For Markov property: Check

$$P(X(t)=y|X(t_1)=x_1,\cdots,X(t_n)=x_n,X(s)=x) = P(X(t)=y|X(s)=x)$$

for any $0 \leqslant t_1 < t_2 < \cdots < t_n < s \leqslant t$. □

Sum: We see that the **Poisson process**

$$X(t), \quad t \geq 0,$$

turns out to be a **MJP** (continuous-time JP with the Markov property) with $X(0) = 0$ and the **transition function**:

For any $t \geq 0$ and any $x, y \in S = \{0, 1, 2 \cdots\}$,

$$\boxed{P_{xy}(t) = \begin{cases} 0 & \text{if } x > y, \\ = P_{0,y-x}(t) = e^{-\lambda t} \dfrac{(\lambda t)^{y-x}}{(y-x)!} & \text{if } x \leq y. \end{cases}}$$

Here, $\lambda > 0$ is the arrival rate. □

§3.3 Basic properties of MJP

Let $\{X(t)\}_{t\geqslant 0}$ be a MJP with

$$P_{xy}(t) = P(X(t) = y | X(0) = x).$$

Proposition. (Chapman-Kolmogorov equation)

$$P_{xy}(t+s) = \sum_z P_{xz}(t) P_{zy}(s).$$

In matrix form, letting $P(t) = [P_{xy}(t)]$, the above is

$$\boxed{P(t+s) = P(t)P(s).}$$

Remark: It is similar to the discrete case

$$P^{m+n}(x,y) = \sum_{z\in s} P^m(x,z)P^n(z,y)$$

Pf.: Note:

$$P_{xy}(t+s) = \sum_z P_x(X(t) = z, X(t+s) = y)$$

and

$$P_x(X(t) = z, X(t+s) = y)$$

$$= P_x(X(t) = z) P_x(X(t+s) = y | X(t) = z)$$

$$= P_x(X(t) = z) P(X(t+s) = y | X(0) = x, X(t) = z)$$

$$= P_x(X(t) = z) P(X(s) = y | X(0) = z) \text{ (Markov+Time-Homg)}$$

$$= P_{xz}(t) P_{zy}(s).$$

It follows that

$$P_{xy}(t+s) = \sum_z P_{xz}(t) P_{zy}(s). \quad \square$$

Note: Assume $P(t)$ is differentiable in $[0, \infty)$, and

$$\boxed{D \stackrel{\text{def}}{=} P'(0).}$$

Then, from the C.-K. equation

$$P(t+s) = P(t)P(s),$$

one has

$$\left.\frac{d}{ds}\right|_{s=0} (\cdot) \Rightarrow P'(t) = P(t)D,$$

$$\left.\frac{d}{dt}\right|_{t=0} (\cdot) \Rightarrow P'(s) = DP(s).$$

$$\therefore \boxed{P'(t) = P(t)D = DP(t), \quad t \geq 0.}$$

Fact I.

$$D = P'(0) \stackrel{\text{def}}{=} [q_{xy}]_{x,y\in S} = \begin{bmatrix} - & + & + & + & \cdots \\ + & - & + & + & \cdots \\ + & + & - & + & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix},$$

called the **rate matrix**.

$$+ : \text{entry} \geq 0; \quad - : \text{entry} \leq 0.$$

Indeed, note:

$$
\begin{aligned}
q_{xy} &= P'_{xy}(0) \\
&= \lim_{h\to 0+} \frac{P_{xy}(h) - P_{xy}(0)}{h} \\
&= \lim_{h\to 0+} \frac{P(X(h) = y | X(0) = x) - P(X(0) = y | X(0) = x)}{h} \\
&= \begin{cases} \lim_{h\to 0+} \dfrac{P(X(h) = y | X(0) = x) - 1}{h} (\leqslant 0) & \text{if } x = y, \\ \lim_{h\to 0+} \dfrac{P(X(h) = y | X(0) = x) - 0}{h} (\geqslant 0) & \text{if } x \neq y. \end{cases}
\end{aligned}
$$

Fact II. Each row sum of D is zero:

$$\sum_{y \in S} q_{xy} = 0, \quad \forall\, x \in S. \qquad (*)$$

Indeed, note:

$$\sum_{y \in S} P_{xy}(t) = 1, \forall\, t \geq 0. \;\therefore\; \frac{d}{dt}\Big|_{t=0} \Rightarrow \sum_{y \in S} P'_{xy}(0) = 0. \qquad \square$$

Observe: $(*)$ means $q_{xx} + \sum_{y \neq x} q_{xy} = 0$, that is,

$$\boxed{\underbrace{-q_{xx}}_{\text{the rate to jump away from } x} = \sum_{y \neq x} \underbrace{q_{xy}}_{\text{the rate to jump to } y \text{ from } x}.}$$

Recall:

- $E(\tau_x)$ is the mean waiting time to jump away from x, so $q_x = \frac{1}{E(\tau_x)}$ is **the rate of change. Note:**

  $$q_x = 0 \textbf{ iff } E(\tau_x) = \infty, \textbf{ iff } x \text{ is absorbing.}$$

- $Q = [Q_{xy}]$ is the Markov matrix introduced before. $Q_{xx} = 1$ **iff** x is absorbing. For non-absorbing x,

  $$Q_{xx} = 0, \quad \sum_{y \neq x} Q_{xy} = 1,$$

  and in such case, Q_{xy} is understood to be **the proportion that the chain will jump to y from x**.

Main Theorem:

$$-q_{xx} = q_x; \quad q_{xy} = q_x Q_{xy} \text{ for } y \neq x.$$

Pf.: Case x is absorbing ($q_x = 0$, $Q_{xy} = \delta_{xy}$):

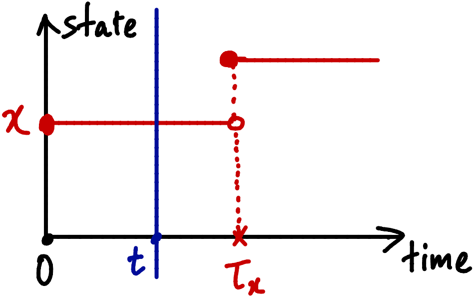
$$P_{xy}(t) = \delta_{xy}. \quad \therefore\ q_{xy} = P'_{xy}(0) = 0.$$

Conclusion is then TRUE.

Case x is non-absorbing:

$$\begin{aligned} P_{xy}(t) &= P_x(X(t) = y) \\ &= \underbrace{P_x(\tau_x > t, X(t) = y)}_{I:\ \text{no jump yet}} \\ &\quad + \underbrace{P_x(\tau_x \leqslant t, X(t) = y)}_{II:\ \text{it has jumped}}. \end{aligned}$$

For I (no jump yet):

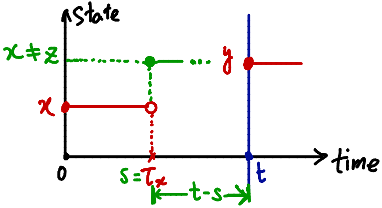


$$I = P_x(\tau_x > t, X(t) = y)$$

$$= \begin{cases} 0 & \text{for } y \neq x, \\ P_x(\tau_x > t) = e^{-q_x t} & \text{for } y = x \end{cases}$$

$$= \delta_{xy} e^{-q_x t}.$$

For II (it has jumped):



$$II = P_x(\tau_x \leqslant t, X(t) = y)$$
$$= \sum_{z \neq x} P_x(\tau_x \leqslant t, X(\tau_x) = z, X(t) = y)$$
$$= \sum_{z \neq x} \int_0^t P_x(\tau_x = s) Q_{xz} P_{zy}(t-s) ds$$
$$= \sum_{z \neq x} \int_0^t q_x e^{-q_x s} Q_{xz} P_{zy}(t-s) ds.$$

$$\therefore P_{xy}(t) = I + II$$

$$= \delta_{xy} e^{-q_x t} + \sum_{z \neq x} \int_0^t q_x e^{-q_x s} Q_{xz} P_{zy}(t-s) ds$$

$$= \delta_{xy} e^{-q_x t} + q_x e^{-q_x t} \sum_{z \neq x} \int_0^t Q_{xz} P_{zy}(u) e^{q_x u} du$$

(Change of variable: $t - s = u$)

$$\therefore P'_{xy}(t) = -q_x P_{xy}(t) + q_x \sum_{z \neq x} Q_{xz} P_{zy}(t)$$

$$\therefore P'_{xy}(0) = -q_x P_{xy}(0) + q_x \sum_{z \neq x} Q_{xz} P_{zy}(0)$$

$$= -q_x \delta_{xy} + q_z \sum_{z \neq x} Q_{xz} \delta_{zy}$$

$$= -q_x \delta_{xy} + q_x Q_{xy}$$

$$= \begin{cases} -q_x + 0 = -q_x & \text{for } y = x, \\ q_x Q_{xy} & \text{for } y \neq x. \end{cases} \quad \square$$

Example 1. Poisson process with rate λt:

$$P_{0n}(t) = P(X(t) = n | X(0) = 0) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \cdots$$

$$P(t) = \begin{bmatrix} e^{-\lambda t} & e^{-\lambda t}\frac{\lambda t}{1!} & e^{-\lambda t}\frac{(\lambda t)^2}{2!} & \cdots \\ 0 & e^{-\lambda t} & e^{-\lambda t}\frac{\lambda t}{1!} & \cdots \\ 0 & 0 & \ddots & \ddots \\ 0 & 0 & 0 & \ddots \end{bmatrix}$$ **(transition function)**

Then

$$D = P'(0) = \begin{bmatrix} -\lambda & \lambda & 0 & \cdots & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots \\ 0 & 0 & 1 & \cdots & \cdots \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix}.$$

Example 2. Car check up with 3 operations in sequence:
(1) Engine time up $\to$ (2) air condition repair $\to$ (3) break system replacement $\to$ (4) leave.

Assume that this is a MJP with the mean time in each operation 1.2, 1.5, 2.5 hours.

$S = \{1, 2, 3, 4\}$. The rate of moving up to the next stage is $\frac{1}{1.2}$, $\frac{1}{1.5}$, $\frac{1}{2.5}$. Thus,

$$D = \begin{bmatrix} -\frac{1}{1.2} & \frac{1}{1.2} & 0 & 0 \\ 0 & -\frac{1}{1.5} & \frac{1}{1.5} & 0 \\ 0 & 0 & -\frac{1}{2.5} & \frac{1}{2.5} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \\ 0\,0\,0\,1 \end{bmatrix}.$$

Further questions:

(a) What is the prob that after 4hour the car is in step (3)? That is to find $P(X(4) = 3|X(0) = 1)$.

(b) What is the prob that after 4hour the car is still in the shop? That is to find $P(X(4) = 4|X(0) = 1)$.

Generally, need to find

$$P(t) = \begin{bmatrix} P_{11}(t) & P_{12}(t) & P_{13}(t) & P_{14}(t) \\ P_{21}(t) & P_{22}(t) & P_{23}(t) & P_{24}(t) \\ P_{31}(t) & P_{32}(t) & P_{33}(t) & P_{34}(t) \\ P_{41}(t) & P_{42}(t) & P_{43}(t) & P_{44}(t) \end{bmatrix}.$$

Method: Solve the linear ODE system:

$$P'(t) = DP(t), \quad P(0) = I.$$

Example 3. A barbar shop with two barbars and two waiting chains. Customers arrives at a rate 5 per hr. Each barbar serves at a rate 2 per hr. If the waiting chains are full the customer will leave.

$$X(t) \stackrel{\text{def}}{=} \text{the no of customers in the shop.}$$
$$S = \{0, 1, 2, 3, 4\}.$$

$$D = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} -5 & 5 & 0 & 0 & 0 \\ 2 & -7 & 5 & 0 & 0 \\ 0 & 4 & -9 & 5 & 0 \\ 0 & 0 & 4 & -9 & 5 \\ 0 & 0 & 0 & 4 & -4 \end{bmatrix} \end{array}, \quad Q = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{2}{7} & 0 & \frac{5}{7} & 0 & 0 \\ 0 & \frac{4}{9} & 0 & \frac{5}{9} & 0 \\ 0 & 0 & \frac{4}{9} & 0 & \frac{5}{9} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Further questions:

(a) In the long run, what is the prob to have one customer, two customers, etc.? That is to find

$$\lim_{t\to\infty} P(X(t) = k), \quad k \in S.$$

(b) Find the expected time for it to be full, counting from the opening time. That is to find

$$E(T_y),$$

where $T_y = \inf\{t : X(t) = y, X(0) = 0\}$.

How to solve:

$$P'(t) = DP(t), \quad P(0) = I.$$

Case when S **is finite**:

$$P(t) = e^{tD} \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} \frac{(tD)^n}{n!} \quad \text{(convergent!)}.$$

Informal Proof: At $t = 0$, $e^{tD} = e^{0D} = I$ (Convention: $0^0 = 1$, $D^0 = I$), and for $t > 0$,

$$\begin{aligned}(e^{tD})' &= \sum_{n=1}^{\infty} \frac{t^{n-1} D^n}{(n-1)!} \\ &= D \left[\sum_{n=1}^{\infty} \frac{(tD)^{n-1}}{(n-1)!} \right] \\ &= De^{tD}. \quad \square\end{aligned}$$

Example. Let $D = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$. **Q.:** Find $P(t)$.

Sol.: Look for $D = Q\,\mathrm{diag}\{\lambda_1, \lambda_2\}\,Q^{-1}$.

(i) **Eigenvalues:** $\det(D - \lambda I) = 0$,
i.e., $0 = \det \begin{bmatrix} -1-\lambda & 1 \\ 2 & -2-\lambda \end{bmatrix} = (-1-\lambda)(-2-\lambda) - 2$,
i.e., $\lambda^2 + 3\lambda = 0$. $\therefore \lambda = 0, -3$.

(ii) **Eigenvectors:** $\lambda = 0 : D - \lambda I = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}, e_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$\lambda = -3 : D - \lambda I = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}, e_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

Let $Q \overset{\text{def}}{=} [e_1, e_2] = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$, $Q^{-1} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$. Then,

$$\begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix} = Q^{-1} D Q, \quad i.e., \boxed{D = Q \begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix} Q^{-1}.}$$

Hence

$$P(t) = e^{tD} = \sum_{n=0}^{\infty} \frac{(tD)^n}{n!} = Q \left(\sum_{n=0}^{\infty} \frac{\left(t \begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix} \right)^n}{n!} \right) Q^{-1}$$

$$= Q \begin{bmatrix} \sum_{n=0}^{\infty} \frac{0^n}{n!} & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{(-3t)^n}{n!} \end{bmatrix} Q^{-1}$$

$$= \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} + e^{-3t} \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{bmatrix},$$

$$\therefore \quad \lim_{t\to\infty} P(t) = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}, \quad \text{namely,}$$

$$\lim_{t\to\infty} P(X(t)=0) = 2/3, \ \lim_{t\to\infty} P(X(t)=1) = 1/3. \quad \square$$

Remark: Set $\pi = [2/3, 1/3]$. Then, $\pi P(t) = \pi$, $\forall\, t \geq 0$, so π is a **SD** for $P(t)$.

§3.4 The birth and death process

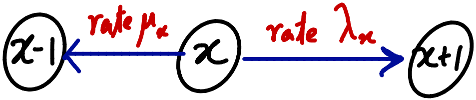
Setup:

Let $S = \{0, 1, \cdots\}$,

$$D = [q_{xy}] = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix}.$$

Assume that all $\lambda_x, \mu_x \neq 0 \, (> 0)$.

λ_x: birth rate, μ_x: death rate



Example 1. Revisit the Poisson process.

We already derived earlier $P(t) = [P_{xy}(t)]$ for a Poisson process $X(t)$, $t \geqslant 0$, using

$$X(t) = \max\{n : \tau_n \leqslant t\}.$$

We further have derived:

$$P'(t) = P(t)D, \quad D = \begin{bmatrix} -\lambda & \lambda & & \\ & -\lambda & \lambda & \\ & & \ddots & \ddots \end{bmatrix},$$

$\lambda > 0$: arrival rate.

Here we want to derive the **inverse**:

Proposition. If $X(t)$ is a MJP with rate matrix

$$D = \begin{bmatrix} -\lambda & \lambda & & \\ & -\lambda & \lambda & \\ & & \ddots & \ddots \end{bmatrix},$$

then $X(t)$ has the Poisson distribution, i.e.

$$P_{xy}(t) = \begin{cases} e^{-\lambda t} \frac{(\lambda t)^{(y-x)}}{(y-x)!} & \text{if } y \geqslant x \geqslant 0, \\ 0 & \text{otherwise.} \end{cases} \qquad \square$$

It is **another way** of obtaining the Poisson process.

Pf.: Recall

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

Observe

(i) If $y = 0$, then

$$P'_{x0}(t) = -\lambda P_{x0}(t), \quad P_{x0}(0) = \delta_{x0}.$$

$$\therefore P_{x0}(t) = \delta_{x0} e^{-\lambda t}.$$

(ii) If $y \geq 1$, then

$$P'_{xy}(t) = \lambda P_{x,y-1}(t) - \lambda P_{xy}(t), \quad P_{xy}(0) = \delta_{xy}.$$

$$\therefore P_{xy}(t) = e^{-\lambda t}\delta_{xy} + \int_0^t e^{-\lambda(t-s)} \lambda P_{x,y-1}(s)\, ds.$$

Claim #1. $P_{xy}(t) = 0, \forall\, y < x$. Indeed,

if $y = 0$ $(x \geq 1)$, $P_{x0}(t) = 0$.

if $y = 1$ $(x \geq 2)$,

$$P'_{x,1}(t) = \lambda P_{x,0}(t) - \lambda P_{x,1}(t) = -\lambda P_{x,1}(t),\ P_{x,1}(0) = 0.$$

$$\therefore P_{x,1}(t) = 0.$$

If $y = 2$ $(x \geq 3)$,

$$P'_{x,2}(t) = \lambda P_{x,1}(t) - \lambda P_{x,2}(t) = -\lambda P_{x,2}(t),\ P_{x,2}(0) = 0.$$

$$\therefore P_{x,2}(t) = 0.$$

Inductively,

$$P_{xy}(t) = 0, \quad \forall\, x > y \geq 0.$$

Claim #2. $P_{xy}(t) = e^{-\lambda t}\frac{(\lambda t)^{y-x}}{(y-x)!}$, $\forall\, y \geq x \geq 0$.

Indeed, let $x \geq 0$ be fixed.
For $y = x$,

$$P_{xx}(t) = e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}\lambda \underbrace{P_{x,x-1}(s)}_{=0}\, ds = e^{-\lambda t}.$$

For $y = x + 1$,

$$P_{x,x+1}(t) = e^{-\lambda t} \underbrace{\delta_{x,x+1}}_{=0} + \int_0^t e^{-\lambda(t-s)} \underbrace{\lambda P_{x,x}(s)}_{=e^{-\lambda s}} ds$$
$$= \cdots = e^{-\lambda t}\lambda t.$$

Inductively, we get the desired result. □

Exercise:

(1) Extend the above to

$$D = \begin{bmatrix} -\lambda_0 & \lambda_0 & & & \\ & -\lambda_1 & \lambda_1 & & \\ & & -\lambda_2 & \lambda_2 & \\ & & & \ddots & \ddots \end{bmatrix}, \quad \text{(see. P.98).}$$

It is a general **pure birth** process.

(2) Think about the more general **BD process**:

$$D = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots \end{bmatrix}. \quad \square$$

Example 2. Branching Process:

– A collection of particles

– each waiting to
either split into two particles with prob p
or vanish with prob $(1-p)$

– the waiting time is exp. r.v. with rate λ.

$$X(t) \stackrel{\text{def}}{=} \text{be the no of particles at time } t.$$

Q.: Find the rate matrix D.

Lemma. Let $\xi_1, \cdots, \xi_n$ be independent r.v. having exponential distribution with rate $\alpha_1, \cdots, \alpha_n$, resp. Then,

$$\min\{\xi_1, \cdots, \xi_n\}$$

is an exponential r.v. with rate

$$\alpha_1 + \cdots + \alpha_n,$$

and for each $k = 1, \cdots, n$

$$P(\xi_k = \min\{\xi_1, \cdots, \xi_n\}) = \frac{\alpha_k}{\alpha_1 + \cdots + \alpha_n}. \quad \square$$

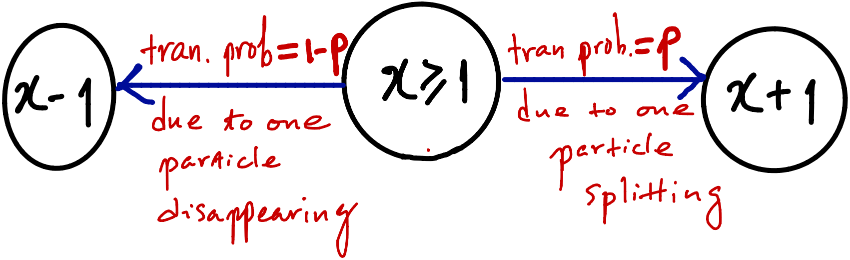
If so, then

$$Q = \begin{bmatrix} 1 & 0 & \cdots & & \\ 1-p & 0 & p & & \\ & 1-p & 0 & p & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \text{ (Markov matrix for state transition)},$$

$$D = \begin{bmatrix} 0 & 0 & 0 & & & \\ (1-p)\lambda & -\lambda & p\lambda & & & \\ & 2\lambda(1-p) & -2\lambda & 2\lambda p & & \\ & & 3\lambda(1-p) & -3\lambda & 3\lambda p & \\ & & & \ddots & \ddots & \ddots \end{bmatrix} \text{ (rate matrix)}.$$

Indeed,

- Let $X(0) = x$, and $\xi_1, \cdots, \xi_x$ be **the time** any one of the particles splits or disappears.
- At time $\tau_1 = \min\{\xi_1, \cdots, \xi_x\}$, the no of particles will be $x+1$ or $x-1$.
- By lemma above, τ_1 is an exp. r.v. with rate λx:
  the portion to $x+1$ $= p \cdot \lambda x$; the portion to $x-1$ $= (1-p) \cdot \lambda x$.



$$\lambda x = \boxed{\text{the rate to jump away from } x}$$

$$p \cdot \lambda x = \boxed{\text{the rate to jump to } x+1}$$

(Birth rate)

$$(1-p) \cdot \lambda x = \boxed{\text{the rate to jump to } x-1}$$

(Death rate)

Proof of Lemma:

$$\begin{aligned}
&P(\min\{\xi_1, \cdots, \xi_n\} > t) \\
&= P(\xi_1 > t, \cdots, \xi_n > t) \\
&= P(\xi_1 > t) \times \cdots \times P(\xi_n > t) \\
&= e^{-\alpha_1 t} \times \cdots \times e^{-\alpha_n t} \\
&= e^{-(\alpha_1 + \cdots + \alpha_n)t}.
\end{aligned}$$

To consider $P(\xi_k = \min\{\xi_1, \cdots, \xi_n\})$, W.L.G. take $k = 1$. Set

$$\eta = \min\{\xi_2, \cdots, \xi_n\}.$$

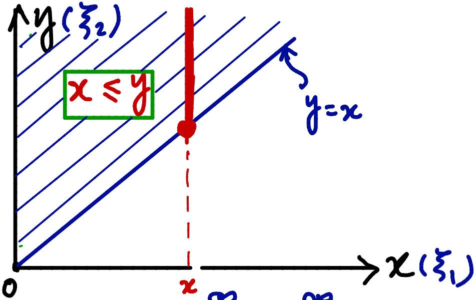
Then by above, η is an exp.r.v. with rate

$$\beta_1 \stackrel{\text{def}}{=} \sum_{y=2}^{n} \alpha_y.$$

$$
\begin{aligned}
& P(\xi_1 = \min\{\xi_1, \cdots, \xi_n\}) \\
&= P(\xi_1 \leqslant \eta) \\
&= \iint\limits_{x \leqslant y} \alpha_1 e^{-\alpha_1 x} \cdot \beta_1 e^{-\beta_1 y} dxdy \\
&= \int_0^\infty \left(\int_x^\infty \cdots dy \right) dx \\
&= \frac{\alpha_1}{\alpha_1 + \beta_1} \\
&= \frac{\alpha_1}{\alpha_1 + \sum_{y=2}^n \alpha_y} \\
&= \frac{\alpha_1}{\alpha_1 + \alpha_2 + \cdots + \alpha_n}. \quad \square
\end{aligned}
$$

For instance, consider ξ_1, ξ_2 only:

$$P(\xi_1 = \min\{\xi_1, \xi_2\}) = P(\xi_1 \leqslant \xi_2) = \iint_{x \leqslant y} \alpha_1 e^{-\alpha_1 x} \cdot \alpha_2 e^{-\alpha_2 y} dxdy = \cdots = \frac{\alpha_1}{\alpha_1 + \alpha_2}.$$



$$\iint_{x \leq y} = \int_0^\infty dx \int_x^\infty dy \cdots$$

Remark: Suppose that we allow **new particles** to immigrate into the system at **rate** α, and then give succeeding generation.

$$\eta \overset{\text{def}}{=} \textbf{the first time} \text{ a new particle arrives.}$$

$$\tau_1 = \min\{\xi_1, \cdots, \xi_x, \eta\}: \text{ the waiting time to change.}$$

the rate of changing **away from** x particles $= x\lambda + \alpha$.v

$$D = \begin{bmatrix} -\alpha & \alpha & & & \\ (1-p)\lambda & -(\lambda+\alpha) & p\lambda+\alpha & & \\ & 2(1-p)\lambda & -(2\lambda+\alpha) & 2p\lambda+\alpha & \\ 0 & & \ddots & \ddots & \ddots \end{bmatrix}$$

See the textbook P92. □

Example 3. Queuing Model.

$X(t) \stackrel{\text{def}}{=}$ the no of persons on the line at time t waiting for service.

$$\begin{cases} \text{arrival rate } \lambda : \text{Poisson} \\ \text{service rate } \mu \text{: exponential distr} \end{cases}$$

There are several models for queueing.

- **M/M/1 queue:**

M stands for memoyless,

1^{st} M stands for waiting time for the arrival,

2^{nd} M stands for waiting time for service,

The last number is for the number of servers.

$$D = \begin{bmatrix} -\lambda & \lambda & 0 & 0 \\ \mu & (-\lambda + \mu) & \lambda & 0 \\ 0 & \ddots & \ddots & \ddots \end{bmatrix}.$$

• **M/M/k queue** (k servers):

Note: $\mu_n = \begin{cases} n\mu & \text{if } n \leqslant k, \\ k\mu & \text{if } n > k. \end{cases}$

$$D = \begin{bmatrix} -\lambda & \lambda & & & & & & \\ \mu & -(\mu+\lambda) & \lambda & & & & \mathbf{0} & \\ & 2\mu & -(2\mu+\lambda) & \lambda & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & k\mu & -(k\mu+\lambda) & \lambda & & \\ & \mathbf{0} & & & k\mu & -(k\mu+\lambda) & \lambda & \\ & & & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

- **M/M/∞ queue** (∞ servers):

$$D = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\mu+\lambda) & \lambda & & \\ & 2\mu & -(2\mu+\lambda) & \lambda & \\ & & \ddots & \ddots & \ddots \end{bmatrix}$$

arrival rate $= \lambda$, service rate $= \mu$,

$X(t) \stackrel{\text{def}}{=}$ the no of customers on the line at time t.

(e.g., in the telephone exchange, this is a continuous-time version of a previous example in the Markov chain).

Q.: Find $P_{xy}(t)$ and $\lim_{t\to\infty} P_{xy}(t)$.

Lemma. Let $Y(t)$ be a Poisson process with rate λ. Then for $0 \leqslant s \leqslant t$ (t fixed),

$$P(\tau_1 \leqslant s | Y(t) = 1) = \frac{s}{t},$$

i.e. the density function is $\frac{1}{t}$ on $[0, t]$, namely, given that the arrival (one) is within $[0, t]$, the arrival time is a uniform distr on $[0, t]$.

Note: This is a special case of Ex 6 with $Y(t) = n$.

Pf.: For $0 \leqslant s \leqslant t$,

$$\begin{aligned}
&P(\tau_1 \leqslant s | Y(t) = 1) \\
&= P(Y(s) = 1 | Y(t) = 1) \\
&= \frac{P(Y(s) = 1, Y(t) = 1)}{P(Y(t) = 1)} \\
&= \frac{P(Y(s) = 1, Y(t) - Y(s) = 0)}{P(Y(t) = 1)} \\
&= \frac{e^{-\lambda s}\frac{(\lambda s)}{1!} \cdot e^{-\lambda(t-s)}\frac{(\lambda(t-s))^0}{0!}}{e^{-\lambda t}\frac{(\lambda t)}{1!}} \\
&= \frac{s}{t}. \quad \square
\end{aligned}$$

Assume $X(0) = x$.

$$Y(t) \stackrel{\text{def}}{=} \text{the total no that arrived in time } (0, t].$$

Let

$$X(t) = R(t) + N(t),$$

$$R(t) \stackrel{\text{def}}{=} \text{the no of the original } x \text{ (at } t = 0\text{) that are still being served,}$$

$$N(t) \stackrel{\text{def}}{=} \text{the no of those from } Y(t) \text{ that are still being served.}$$

Fact 1. $R(t)$, i.e., the no of the original x (at $t = 0$) that are still being served,

is a binomial r.v.:

$$P(R(t) = k) = \binom{x}{k} (e^{-\mu t})^k (1 - e^{-\mu t})^{x-k},$$

$$0 \le k \le x,$$

$$x = \text{the total no at } t = 0,$$

$e^{-\mu t}$ = **the success prob** of still being served.

Fact 2. Recall: $Y(t)$ is the total no that arrived in time $(0, t]$. We want to consider

$$P(N(t) = n | Y(t) = k).$$

Note: Fix t.

- Given $Y(t) = k$, $N(t)$ should be a binomial r.v., but we have to find "**the success prob**":

$$p_t = P(N(t) = 1 | Y(t) = 1).$$

- For one that arrived at time $s \in (0, t]$, the prob of still being served at time t is $e^{-\mu(t-s)}$.
- By lemma, the arrival time s subject to one arrival in $(0, t]$ is uniform dist $1/t$.
- Then the prob that he is still being served at time t is

$$p_t = \int_0^t \frac{1}{t} \cdot e^{-\mu(t-s)} ds = \frac{1 - e^{-\mu t}}{\mu t}.$$

Hence,

$$P(N(t) = n | Y(t) = k) = \binom{k}{n} p_t^n (1 - p_t)^{k-n},$$

$$0 \leq n \leq k.$$

$$\begin{aligned} \therefore\ P(N(t) = n) &= \sum_{k=n}^{\infty} P(Y(t) = k, N(t) = n) \\ &= \sum_{k=n}^{\infty} P(Y(t) = k) P(N(t) = n | Y(t) = k) \\ &= \cdots \\ &= \frac{(\lambda t p_t)^n}{n!} e^{-\lambda t p_t}. \quad \text{(see P101)} \end{aligned}$$

The same as in last Chap (P55). □

We conclude that (Recall $X(t) = R(t) + N(t)$)

$$\begin{aligned} P_{xy}(t) &= P_x(X(t) = y) \\ &= \sum_{k=0}^{\min\{x,y\}} P_x(R(t) = k) P(N(t) = y - k) \\ &= \sum_{k=0}^{\min\{x,y\}} \binom{x}{k} e^{-k\mu t} (1 - e^{-\mu t})^{x-k} \frac{(\lambda t P_t)^{y-k} e^{-\lambda t p_t}}{(y-k)!}. \end{aligned}$$

For $t \to \infty$, all the terms vanish except $k = 0$:

$$\lim_{t \to \infty} P_{xy}(t) = e^{-\lambda/\mu} \frac{(\lambda/\mu)^y}{y!} \quad (tp_t \to 1/\mu \quad \text{as } t \to \infty).$$

Note: Compare it with the "telephone exchange" example last chapter.

§3.5 Limiting properties of MJP

The definitions of

– stationary distribution (SD)

– recurrence or transience

– etc

are the same as Markov chain.

Let us only sketch some of them.

- **SD:**

Let

$$X(t), \quad t \geqslant 0,$$

be a MJP.

Def.: π is called a **SD if**

(i) (distribution)

$$\pi(y) \geqslant 0, \forall\, y \in S; \sum_{y} \pi(y) = 1.$$

(ii) (stationary)

$$\sum_{x \in S} \pi(x) P_{xy}(t) = \pi(y), \forall\, y \in S, \forall\, t \geqslant 0.$$

How to find the SD π?

In fact,

$$0 = \left(\sum_x \pi(x) P_{xy}(t) \right)' = \sum_x \pi(x) P'_{xy}(t).$$

(**Note:** there is a technical point to interchange $\sum_x$ and $(\cdot)'$ for the **infinite** sum)

Let $t \to 0+$, then $\sum_x \pi(x) q_{xy} = 0$, i.e. in matrix form

$$\pi D = 0,$$

where $D = [q_{xy}]$ is the rate matrix. The converse is also true.

Example. Find the SD of the birth and death process with rate

$$D = \begin{bmatrix} -\lambda_0 & \lambda_0 & & & \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & & \\ & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Sol.: Let $\pi = (x_0, x_1, \cdots)$. $\pi D = 0$ is

$$[x_0, x_1, \cdots] \begin{bmatrix} -\lambda_0 & \lambda_0 & & & \\ \mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & & \\ & & \ddots & \ddots & \ddots \end{bmatrix} = [0, 0, \cdots].$$

Hence

$$\begin{cases} -\lambda_0 x_0 + \mu_1 x_1 = 0, \\ \lambda_{k-1}x_{k-1} - (\lambda_k + \mu_k)x_k + \mu_{k+1}x_{k+1} = 0, \ \ k \geq 1. \end{cases}$$

Note: For $k \geq 1$,

$$\begin{aligned} \lambda_k x_k - \mu_{k+1}x_{k+1} &= \lambda_{k-1}x_{k-1} - \mu_k x_k \\ &= \cdots = \lambda_0 x_0 - \mu_1 x_1 = 0. \end{aligned}$$

$$\therefore x_k = \frac{\lambda_{k-1}}{\mu_k} x_{k-1} = \cdots = \frac{\lambda_{k-1}}{\mu_k} \cdot \frac{\lambda_{k-2}}{\mu_{k-1}} \cdots \frac{\lambda_0}{\mu_1} x_0,$$

$$\therefore x_k = \beta_k x_0, \quad \beta_k \stackrel{\text{def}}{=} \frac{\lambda_0 \cdots \lambda_{k-1}}{\mu_1 \cdots \mu_k} \, (k \geq 1).$$

Formally, $\sum_{k=0}^{\infty} x_k = (\sum_{k=0}^{\infty} \beta_k) x_0$ (Convention: $\beta_0 = 1$).
Then,

- if $\beta \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \beta_k < \infty$, then choosing $x_0 = \frac{1}{\beta}$,

$$\pi = \left(\frac{1}{\beta}, \frac{\beta_1}{\beta}, \frac{\beta_2}{\beta}, \cdots \right) \text{ is a SD.}$$

- if $\sum_{k=0}^{\infty} \beta_k = \infty$, then, no SD! $\square$

Exercise: Use this to check the queue models:

$$M/M/1,\ M/M/2,\ M/M/\infty.$$

For instance, $M/M/\infty$ case:

$$\begin{cases} \lambda_k = \lambda \ (k \geqslant 0) \\ \mu_k = k\mu \ (k \geqslant 1) \end{cases} \quad \therefore \beta_k = \left(\frac{\lambda_0}{\mu_1}\right) \cdots \left(\frac{\lambda_{k-1}}{\mu_k}\right) = \frac{\lambda^k}{k!\mu^k}.$$

$$\sum_{k \geq 0} \beta_k = e^{\lambda/\mu}.$$

$$\therefore \pi = \left(e^{-\frac{\lambda}{\mu}}, e^{-\frac{\lambda}{\mu}} \frac{\lambda}{\mu}, \frac{e^{-\frac{\lambda}{\mu}} (\frac{\lambda}{\mu})^2}{2!}, \cdots, \frac{e^{-\frac{\lambda}{\mu}} (\frac{\lambda}{\mu})^k}{k!}, \cdots \right).$$

"the same as the one by looking for the limit distribution $\lim_{t \to \infty} P_{xy}(t)$"

- **Recurrence and transience.**

$\tau_1 \overset{\text{def}}{=}$ the first time to jump

$T_y \overset{\text{def}}{=} \min\{t \geqslant \tau_1 : X(t) = y\}$ (hitting time)

$$(= \infty \text{ if } X(t) \neq y, \forall\, t \geqslant \tau_1)$$

$\rho_{xy} \overset{\text{def}}{=} P_x(T_y < \infty)$

(the prob that the process starting from x eventually hits y)

Recurrent: $\rho_{yy} = 1$.
Transient: $\rho_{yy} < 1$.
Process is irreducible: $\rho_{xy} > 0$, $\forall\, x, y \in S$.

Let Q be the matrix in the MJP, i.e.

$$P_x(\tau_1 \leqslant t, X(\tau_1) = y) = F_x(t) Q_{xy},\ \ y \neq x,$$

$$F_x(t) = 1 - e^{-q_x t}.$$

Assume irreducible, i.e. $q_x > 0$, $\forall\, x$. Then

$$P(X(\tau_1) = y | X(0) = x) = Q_{xy} (= \frac{q_{xy}}{q_x}),\ \forall\, y \neq x.$$

Let $\tau_0 = 1$, and

$$Z_n = X(\tau_n),\ \ n = 0, 1, 2, \cdots$$

(Only count the jump each time, but ignore the length of waiting time).

Then,

$\{Z_n\}_{n=0}^{\infty}$ is a Markov chain with Q as transition matrix.

Note:

$$T_y \stackrel{\text{def}}{=} \inf\{t \geqslant \tau_1 : X(t) = y\} < \infty$$

iff

$T'_y \stackrel{\text{def}}{=} \inf\{n \geqslant 1 : Z_n = y\} < \infty$ (as Markov chain).

$\therefore$ ρ_{xy} for $\{Z_n\}_{n=0}^{\infty}$ is the same as ρ_{xy} for $\{X(t)\}_{t\geqslant 0}$.

$\therefore$To check recurrent/transience,

we need only consider Q!

Example: In the birth & death process

$$Q = \begin{bmatrix} 0 & 1 & & & \\ \frac{\mu_1}{\lambda_1+\mu_1} & 0 & \frac{\lambda_1}{\lambda_1+\mu_1} & & \\ & \frac{\mu_2}{\lambda_2+\mu_2} & 0 & \frac{\mu_2}{\lambda_2+\mu_2} & \\ & \ddots & \ddots & \ddots \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} 0 & 1 & & & & \\ q_1 & 0 & p_1 & & & \\ & q_2 & 0 & p_2 & & \\ & & \ddots & \ddots & \ddots \end{bmatrix}.$$

It follows from Chapter 1 (P33) that the chain is recurrent **iff**

$$\sum_{n=1}^{\infty} \frac{\mu_1 \cdots \mu_n}{\lambda_1 \cdots \lambda_n} = \sum_{n=1}^{\infty} \frac{q_1 \cdots q_n}{p_1 \cdots p_n} = \infty.$$

- **Long-run behavior.**

Let $m_x \stackrel{\text{def}}{=} E_x(T_x)$ (the mean return time).

– Null recurrent: $m_x = \infty$

– Positive recurrent: $m_x < \infty$. In this case

$$\boxed{\pi(x) = \frac{1}{q_x m_x}.} \qquad (*)$$

Intuitive Proof of $(*)$:

– In $[0, t]$ for large t, the process will visit x for $\frac{t}{m_x}$ times and the average time staying at x (waiting time to jump way) per visit is $1/q_x$.

– The total time spent in x during $[0, t]$ is $\frac{t}{m_x} \cdot \frac{1}{q_x}$.

– The proportion of time spent in x is $\frac{1}{q_x m_x}$.

Note: Any MJP is **aperiodic**.

For an irreducible, positive recurrent MJP,

$$\lim_{t \to \infty} P_{xy}(t) = \pi(y) = \frac{1}{q_y m_y}, \quad x, y \in S.$$

The end of lectures