

# Lecture notes of Stochastic Process

lectured by prof. Hsueh-I Lu

pishen

AlgoLab, CSIE, NTU

March 26, 2012

# Thank list

LeoSW, windker

# Limiting Probabilities

# Outline

- 1 Limiting probabilities
- 2 Stationary distribution
- 3 Long-run proportion
- 4 (Inverse of) Expected return time

# Limiting Probabilities

## Definition

Number  $\pi_j$  is the *limiting probability* of  $j$  if

$$\pi_j = \lim_{n \rightarrow \infty} P_{ij}^n$$

holds for all states  $i \in S$  ( $S \subseteq \mathbb{N}$  is the state space).

- $\pi_j$  is independent of  $i$ .

- $\lim_{n \rightarrow \infty} P^n = \begin{pmatrix} \pi \\ \pi \\ \vdots \end{pmatrix}$ , where  $\pi = (\pi_1, \pi_2, \dots)$

# Stationary Probability Distribution

## Definition

Non-negative row vector  $\pi = (\pi_1, \pi_2, \dots)$  is a *stationary probability distribution* of  $\mathbb{X}$  if  $\pi \times P = \pi$  holds and  $\sum_{i \in S} \pi_i = 1$

- $\pi$  is a normalized left eigenvector with eigenvalue  $= 1$ .
- If  $X(0)$  has distribution  $\pi$ , then  $X(t)$  has the same distribution  $\pi$  for all  $t \geq 1$ .  $\pi$  is also called as *steady-state distribution*.
- It doesn't mean that each  $X(t)$  become independent.  $\pi$  only means the distribution of  $X(t)$  when the previous random variable's value is unknown.

# Theorem 1

## Theorem

*Let  $\mathbb{X}$  be an irreducible, aperiodic, positive recurrent Markov chain, then*

- *The limiting probability  $\pi_j$  of each state  $j$  exists.*
- *$\pi = (\pi_1, \pi_2, \dots)$  is the unique stationary probability distribution.*
- The proof will be stated at page 21.

# Expected return time

## Definition

The *expected return time* of state  $i \in S$  is

$$\mu_i = \sum_{n \geq 1} n \cdot f_i^{(n)}$$

where

$$f_i^{(n)} = P(\min\{t : X(t) = i, t \geq 1\} = n | X(0) = i)$$

$$\blacksquare f_i = \sum_{n \geq 1} f_i^{(n)}$$



# Positive recurrent & null recurrent

## Definition

State  $i$  is *positive recurrent* if  $\mu_i < \infty$

## Definition

State  $i$  is *null recurrent* if  $\mu_i = \infty$

- Both are recurrent states, and are *class properties*, which means that if state  $i$  and  $j$  communicate, they will share this property.
- If  $\mathbb{X}$  is finite, then each recurrent state of  $\mathbb{X}$  is positive recurrent.  
Proof stated at page 46.

# Example of null recurrent

## Example

For a Markov chain with  $n$  states  $(1, \dots, n)$ , if

$$P(X(t+1) = i+1 | X(t) = i) = 1 - 1/n$$

and

$$P(X(t+1) = 1 | X(t) = i) = 1/n$$

According to geometric distribution (taking  $p = 1/n$ ), the expectation value of “steps taken for state 1 to come back” will be  $1/p = n$ , hence  $\lim_{n \rightarrow \infty} n = \infty$ .

# Period of a chain

## Definition

The *period* of state  $i$  is  $d$  if  $d$  is the largest integer such that

$$P_{ii}^n = 0$$

holds for all  $n$  which is not divisible by  $d$ .

## Definition

If each state of  $\mathbb{X}$  has period 1, then  $\mathbb{X}$  is called *aperiodic*.

- If  $P_{ii} > 0$  for all  $i \in S$ , then  $\mathbb{X}$  is aperiodic.
- Period can be seen as the gcd of all  $n$  that have  $P_{ii}^n > 0$ , note that  $P_{ii}^{\text{gcd}} > 0$  is not necessary.
- The period of drunken man problem is 2.

# Lemma 1

## Lemma

*If state  $j$  is aperiodic and positive recurrent, then*

$$\pi_j \equiv \lim_{n \rightarrow \infty} P_{ij}^n$$

*exists and is positive for all states  $i \in S$ .*

- This can be proved by the Blackwell theorem in Renewal theory.
- It doesn't promise that each  $\pi_j$  for different  $i$  will be the same. But they will be the same if we add the irreducible property.

# Property of lim

- The position of lim cannot be switched arbitrarily in an equation.

## Example

$$1 = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{m}{m+n} \neq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{m}{m+n} = 0$$

- lim would not influence the inequality.

## Example

$$\text{If } f(n) \geq g(n), \text{ then } \lim_{n \rightarrow \infty} f(n) \geq \lim_{n \rightarrow \infty} g(n)$$

# Property of $\lim$ (cont.)

- $\lim$  is linear operator under finite number of functions.

## Example

For  $m < \infty$ ,

$$\sum_{i=1}^m \lim_{n \rightarrow \infty} f_i(n) = \lim_{n \rightarrow \infty} \sum_{i=1}^m f_i(n)$$

need an example of  $m = \infty$

# Inequality 1

## Inequality

$$\sum_{j \in S} \pi_j \leq 1$$

Proof.

$$\begin{aligned}\lim_{m \rightarrow \infty} \sum_{j=1}^m \pi_j &= \lim_{m \rightarrow \infty} \sum_{j=1}^m \lim_{n \rightarrow \infty} P_{ij}^n \\ &= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \sum_{j=1}^m P_{ij}^n \\ &\leq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \sum_{j \in S} P_{ij}^n = 1\end{aligned}$$



- The last equation works since  $\sum_{j \in S} P_{ij}^n = 1$ .



# Inequality 2

## Inequality

*For state  $j \in S$ , we have*

$$\pi_j \geq \sum_{i \in S} \pi_i P_{ij}$$

Proof.

For  $m \geq 1$  and  $n \geq 1$ ,

$$P_{ij}^{n+1} = \sum_{k \in S} P_{ik}^n P_{kj} \geq \sum_{k=1}^m P_{ik}^n P_{kj}$$

then

$$\begin{aligned} \pi_j &= \lim_{n \rightarrow \infty} P_{ij}^{n+1} \\ &\geq \lim_{n \rightarrow \infty} \sum_{k=1}^m P_{ik}^n P_{kj} = \sum_{k=1}^m \lim_{n \rightarrow \infty} P_{ik}^n P_{kj} = \sum_{k=1}^m \pi_k P_{kj} \end{aligned}$$

hence, we know

$$\lim_{m \rightarrow \infty} \sum_{k=1}^m \pi_k P_{kj} = \sum_{k \in S} \pi_k P_{kj} \leq \pi_j$$



# Equality 1

## Equality

$$\pi_j = \sum_{i \in S} \pi_i P_{ij}$$

Proof.

Assume that for some  $j \in S$ ,  $\pi_j > \sum_{i \in S} \pi_i P_{ij}$ , then

$$\begin{aligned} \sum_{j \in S} \pi_j &> \sum_{j \in S} \sum_{i \in S} \pi_i P_{ij} \\ &= \sum_{i \in S} \pi_i \sum_{j \in S} P_{ij} = \sum_{i \in S} \pi_i \end{aligned}$$

Since a value cannot be greater than itself, we got contradiction. □

- $\sum$  should be represented by  $\lim \sum$ , the equation will still work.

# Proof of theorem 1

- **Step 0:** existence of limiting probability.
- **Step 1:** existence of stationary probability distribution.
- **Step 2:** uniqueness.

## 0. Existence of limiting probability

### Proof.

By lemma 1, we know that there exists a  $\pi_j$  for row  $i$ . Since the Markov chain is irreducible and all the states are positive recurrent, for any state  $i'$  other than  $i$ , we know that  $i'$  surely will visit  $i$  in finite steps. Therefore, the  $\pi_j$  value at row  $i'$  will equal to the  $\pi_j$  value at row  $i$ , which means that all the  $\pi_j$  for column  $j$  are the same, and is the limiting probability.  $\square$

still not clear enough

# 1. Existence of stationary probability distribution

We want to prove that

Target

*There's a vector  $s = (s_1, s_2, \dots)$  such that*

1  $\sum_{i \in S} s_i = 1$

2  $s \times P = s$

## Proof.

By lemma 1, we know that there exists a  $\pi = (\pi_1, \pi_2, \dots)$ .

And by equality 1, we know that

$$(\pi_1, \pi_2, \dots) \times P = (\pi_1, \pi_2, \dots)$$

Hence  $\pi$  can satisfy the 2nd part of our target.

Then, we take  $k = \sum_{i \in S} \pi_i$ . By inequality 1, we know that  $k < \infty$ , and can get

$$\left(\frac{\pi_1}{k}, \frac{\pi_2}{k}, \dots\right) \times P = \left(\frac{\pi_1}{k}, \frac{\pi_2}{k}, \dots\right)$$

where  $\sum_{i \in S} \frac{\pi_i}{k} = 1$  also satisfy the 1st part of our target.

Therefore, this vector can be  $s$ , which means that it exists. □



## 2. Uniqueness

### Target

*If  $s = (s_1, s_2, \dots)$  is a stationary distribution of  $\mathbb{X}$ , then  $s = \pi$ .*

- We'll prove this by inequality 3 & 4.

# Inequality 3

Inequality

$$s_j \geq \pi_j, \forall j \in S$$

## Proof.

Let the distribution of  $X(0)$  be  $s$ , by the property of stationary distribution, we have

$$\begin{aligned}s_j &= P(X(n) = j) = \sum_{i \in S} P(X(n) = j | X(0) = i) P(X(0) = i) \\&= \sum_{i \in S} P_{ij}^n \cdot s_i \\&\geq \sum_{i=1}^m P_{ij}^n \cdot s_i \\&\Rightarrow s_j = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} s_j \\&\geq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \sum_{i=1}^m P_{ij}^n \cdot s_i = \lim_{m \rightarrow \infty} \sum_{i=1}^m \pi_j \cdot s_i = \pi_j\end{aligned}$$



# Inequality 4

Inequality

$$s_j \leq \pi_j, \forall j \in S$$

Proof.

Similar in the proof above,  $\forall m, n \geq 1$ , we have

$$\begin{aligned} s_j &= \sum_{i \in S} P_{ij}^n \cdot s_i \\ &\leq \sum_{i=1}^m P_{ij}^n \cdot s_i + \sum_{i=m+1}^{\infty} s_i \\ \Rightarrow s_j &= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} s_j \\ &\leq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \left( \sum_{i=1}^m P_{ij}^n \cdot s_i + \sum_{i=m+1}^{\infty} s_i \right) \\ &= \pi_j \end{aligned}$$



# An example Markov chain

## Example

$$P = \begin{pmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{pmatrix}, 0 < \alpha, \beta < 1$$

$$\pi = \left( \frac{\beta}{1 + \beta - \alpha}, \frac{1 - \alpha}{1 + \beta - \alpha} \right)$$

# Real world example: Hardy-Weinberg Law

## Example

There're two kinds of allele:

- dominant: **A**
- recessive: **a**

And three kinds of senotype with population proportion as follow:

- AA:  $p$
- aa:  $q$
- Aa:  $r = 1 - (p + q)$

## Example (cont.)

$$P = \begin{array}{cc} & \begin{array}{ccc} AA & aa & Aa \end{array} \\ \begin{array}{c} AA \\ aa \\ Aa \end{array} & \left( \begin{array}{ccc} p + \frac{r}{2} & 0 & q + \frac{r}{2} \\ 0 & q + \frac{r}{2} & p + \frac{r}{2} \\ \frac{p}{2} + \frac{r}{4} & \frac{p}{2} + \frac{r}{4} & \frac{p+q+r}{2} \end{array} \right) \end{array}$$

we get  $\pi = (p, q, r)$  when

- $p = (p + \frac{r}{2})^2$
- $q = (q + \frac{r}{2})^2$
- $r = 2(p + \frac{r}{2})(q + \frac{r}{2})$



# Long-run proportion

## Definition

We say that  $r_j$  is the *long-run proportion* of state  $j \in S$  if

$$r_j = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq t \leq n} P_{ij}^t$$

holds for each state  $i \in S$ .

- It represents the average appearance times of state  $j$  in the whole process.
- We will show that (in theorem 3) if  $\mathbb{X}$  is irreducible, then the long-run proportion of all states exist.

# Theorem 2

## Theorem (type 1)

*If  $r_j$  exists for each  $j \in S$  and  $\sum_{j \in S} r_j > 0$ , then  $r = (r_1, r_2, \dots)$  is the unique stationary distribution of  $\mathbb{X}$ .*

or

## Theorem (type 2)

*If  $r_j$  exists for each  $j \in S$  and **a stationary distribution exists**, then  $r = (r_1, r_2, \dots)$  is the unique stationary distribution of  $\mathbb{X}$ .*

# Proof

## Existence of stationary distribution in type 1:

Let

$$R = \begin{pmatrix} r \\ r \\ \vdots \end{pmatrix} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq t \leq n} P^t$$

then

$$\begin{aligned} R \times P &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq t \leq n} P^{t+1} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq t \leq n} P^t + \lim_{n \rightarrow \infty} \frac{1}{n} (P^{n+1} - P) \\ &= R \end{aligned}$$

As stated later,  $\sum_{j \in S} r_j \leq 1$ , hence by normalizing  $r$ , we prove that stationary distribution exist.

$$\blacksquare (\lim f(n)) \cdot g(n) = \lim f(n) \cdot g(n)?$$

## Proof (cont.)

### Uniqueness:

Let  $\pi$  be an arbitrary stationary distribution, then

$$\begin{aligned}r &= \pi \times R \\&= \pi \times \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq t \leq n} P^t \\&= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq t \leq n} \pi \times P^t \\&= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq t \leq n} \pi \\&= \pi\end{aligned}$$

## Proof (cont.)

**Prove that**  $\sum_{j \in S} r_j \leq 1$ :

$$\begin{aligned}\sum_{j \in S} r_j &= \lim_{m \rightarrow \infty} \sum_{j=1}^m \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n P_{ij}^t \\ &= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^m P_{ij}^t \\ &\leq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sum_{j \in S} P_{ij}^t \\ &= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n 1 = 1\end{aligned}$$

# Example 1

On a highway, if we know the probability that

- A truck is followed by a truck:  $1/4$
- A truck is followed by a car:  $3/4$
- A car is followed by a truck:  $1/5$
- A car is followed by a car:  $4/5$

We can construct a matrix

$$\begin{array}{cc} & \begin{array}{cc} T & C \end{array} \\ \begin{array}{c} T \\ C \end{array} & \begin{pmatrix} 1/4 & 3/4 \\ 1/5 & 4/5 \end{pmatrix} \end{array}$$

and get the portion of trucks and cars on the whole highway as the eigenvector  $(4/19, 15/19)$  (we will know that long-run proportion exists by Theorem 3).

## Example 2

For a system which has several good and bad states, we have a matrix  $P$ :

$$\begin{matrix} & g_1 & g_2 & \cdots & b_1 & b_2 & \cdots \\ \begin{matrix} g_1 \\ g_2 \\ \vdots \\ b_1 \\ b_2 \\ \vdots \end{matrix} & \left( \begin{array}{cccccc} & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \end{array} \right) \end{matrix}$$

## Example 2 (cont.)

**Q1:** Breakdown rate (breakdown times / total time)

The long-run frequency of going to a bad state from a good state is

$$\sum_{i \in g} \sum_{j \in b} r_i P_{ij}$$



## Example 2 (cont.)

**Q2:** The expected time  $\mu_G$  (resp.  $\mu_B$ ) of staying in good (resp. bad) states once we reach a good (resp. bad) state?

**Ans:**

For each  $t = 1, 2, \dots$ , let  $G_t$  (resp.  $B_t$ ) be the length of the  $t$ -th good (resp. bad) phase of consecutive good (resp. bad) states. By the strong law of large numbers,

$$P\left(\lim_{t \rightarrow \infty} \frac{G_1 + B_1 + G_2 + B_2 + \dots + G_t + B_t}{t} = \mu_G + \mu_B\right) = 1$$

Since the reciprocal of above is the breakdown rate, we get equation (1):

$$P\left(\sum_{i \in G} \sum_{j \in B} \pi_i P_{ij} = \frac{1}{\mu_G + \mu_B}\right) = 1$$

## Example 2 (cont.)

Also, with probability 1, we get equation (2):

$$P \left( \sum_{i \in G} r_i = \lim_{t \rightarrow \infty} \frac{G_1 + G_2 + \cdots + G_t}{G_1 + B_1 + \cdots + G_t + B_t} = \frac{\mu_G}{\mu_G + \mu_B} \right) = 1$$

Then, by (2)/(1), we get that

$$P \left( \mu_G = \frac{\sum_{i \in G} r_i}{\sum_{i \in G} \sum_{j \in B} r_i P_{ij}} \right) = 1$$

■  $\lim \frac{f(n)}{g(n)} = \frac{\lim f(n)}{\lim g(n)} ?$

# Theorem 3

## Theorem

*If  $\mathbb{X}$  is irreducible, then the long-run proportion  $r_i$  exists with probability 1, moreover,*

- 1** *If state  $i$  is positive recurrent (i.e.  $0 < \mu_i < \infty$ ), then  $P(r_i = \frac{1}{\mu_i}) = 1$ .*
- 2** *If state  $i$  is null recurrent (i.e.  $\mu_i = \infty$ ) or transient, then  $P(r_i = 0) = 1$ .*

## Part 1:

Suppose  $X(0) = i$ ,  $T_k$  is the number of steps required for the  $k$ -th  $i$  goes to  $(k+1)$ -st  $i$ , then by the strong law of large number,

$$P\left(\lim_{k \rightarrow \infty} \frac{T_1 + T_2 + \cdots + T_k}{k} = \mu_i\right) = 1$$
$$\Rightarrow P\left(r_i = \lim_{k \rightarrow \infty} \frac{k}{T_1 + T_2 + \cdots + T_k} = \frac{1}{\mu_i}\right) = 1$$

■  $\lim(A/B) = \frac{1}{\lim(B/A)}$ ?

# Proof (cont.)

## Part 2:

- 1 If  $i$  is transient,  $i$  will only appear finite times in the long-run, hence

$$r_i = \frac{\text{finite}}{\infty} = 0$$

- 2 If  $i$  is null recurrent,  $\mu_i$  is  $\infty$ , then

$$P\left(\lim_{k \rightarrow \infty} \frac{T_1 + T_2 + \cdots + T_k}{k} = \infty\right) = 1$$

$$P\left(r_i = \lim_{k \rightarrow \infty} \frac{k}{T_1 + T_2 + \cdots + T_k} = 0\right) = 1$$

(The first equation is not promised by the strong law of large number. But if it's not  $\infty$ , we can say that  $\mu_i$  is not  $\infty$ , which is a contradiction.)

# Example 1

## Example (type 1)

If  $\mathbb{X}$  is **irreducible** and finite, then  $\mathbb{X}$  has no null recurrent states.

## Example (type 2)

If  $\mathbb{X}$  is finite, then  $\mathbb{X}$  has no null recurrent states.

- Finite irreducible imply positive recurrent.

## ■ Type 1:

If there's a state which is null recurrent, by irreducible, all the states will be null recurrent. Then, all states have  $P(r_i = 0) = 1$ . By changing the proof in page 37 into finite states version, we know that  $\sum r_i = 1$ . So it's impossible for finite  $r_i$ , which are all close to 0, to sum up to 1.

## ■ Type 2:

If it's not irreducible, the finite set of communicated null recurrent states still form an irreducible and finite Markov chain, which can fit the requirement of type 1.

## Example 2

### Example

In the drunken man problem with infinite states, no state will be positive recurrent.

- Infinite drunken man imply no positive recurrent. Note that it doesn't mean all infinite irreducible Markov chain has no positive recurrent state.



If all the states are positive recurrent, then by theorem 3, we know that all the  $r_i > 0$  and is a finite value. Since each state of drunken man problem has the same structure, all the  $r_i$  has same value. We then set  $r = \epsilon \cdot \min(r_1, r_2, \dots)$  ( $0 < \epsilon < 1$ ) such that  $r_i > r > 0, \forall i$ . And get

$$\sum_{i \in S} r_i > \sum_{i \in S} r = \infty > 1$$

which is contradiction to page 37.

## Example 3: Grand Hotel

### Example

There's a hotel, with  $N$  representing the number of newly occupied rooms each day ( $N$  is a poisson distribution with parameter  $\lambda$ ). And the number of consecutive check-in days of each room is a geometric distribution with probability  $p$  ( $p$  is the probability of check-out).  $X(t)$  is the number of occupied rooms in day  $t$ .

## Q1: $P_{ij} = ?$

We set  $R_i$  as a binomial distribution with parameter  $(i, 1 - p)$ , which represents the number of rooms which will remain occupied in the next day, then

$$\begin{aligned} P_{ij} &= P(R_i + N = j) \\ &= \sum_{k \geq 0} P(R_i + N = j | R_i = k) P(R_i = k) \\ &= \sum_{k \geq 0} P(N = j - k) P(R_i = k) \\ &= \sum_{0 \leq k \leq \min(i, j)} \frac{e^{-\lambda} \cdot \lambda^{j-k}}{(j-k)!} \binom{i}{k} (1-p)^k p^{1-k} \end{aligned}$$

## Q2: $r_i = ?$

We guess (by a dream?) there's a stationary distribution which is a poisson distribution with parameter  $\lambda_0$ . Setting  $X(0)$  with this distribution. And let  $R$  as the number of rooms in  $X(0)$  which remain check-in in the next day ( $R$  is a poisson distribution with parameter  $\lambda_0(1 - p)$ ).  $X(1)$  will have distribution  $R + N$ , which is a poisson distribution with parameter  $\lambda_0(1 - p) + \lambda$ . Then since  $X(0)$  is a stationary distribution, it will have the same distribution with  $X(1)$ , which means that  $\lambda_0 = \lambda_0(1 - p) + \lambda$ , and we get  $\lambda_0 = \lambda/p$ .