

# Chapter 4

Pranav Tikkawar

September 20, 2024

9/3

## Chapter 4

### Markov Property

If the probability of the next state only depends on the current state, it satisfies the "Markov Property".

**Drunkards walk example**

$$\mathbb{P}(x_{i+1} = x_i \pm 1) = \frac{1}{2} \mathbb{P}(x_{i+1} \neq x_i \pm 1) = 0$$

$$\mathbb{P}(x_{i+1} = x + 1 | x_i = x) = 1/2$$

$$\mathbb{P}(x_{i+1} = x - 1 | x_i = x) = 1/2$$

### Formal Definition

Let  $\{X_n, n \in \mathbb{N}\}$  be a stochastic process that takes discrete time values. Suppose  $\mathbb{P}(X_{n+1} = j | X_n = i_n \dots X_0 = i_0) = P_{i,j}$ . Such a stochastic process is called a Markov Chain.  $P_{ij}$  is the transition probability from state  $i$  to state  $j$ .

### Transition Probability Matrix

Let  $i, j \in \mathbb{N}$  be possible states of the Markov Chain. The matrix  $P = [P_{ij}]$  is called the transition probability matrix of the Markov Chain. Where  $P_{ij} = \mathbb{P}(x_{n+1} = j | x_n = i)$ . **Ex 4.1**

$$\mathbb{P}(\text{rain tomorrow} | \text{rain today}) = \alpha$$

$$\mathbb{P}(\text{rain tomorrow} | \text{no rain today}) = \beta$$

$$\text{Let } \begin{cases} 0 = \text{rain} \\ 1 = \text{no rain} \end{cases}$$

$$P = \begin{bmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{bmatrix}$$

**Ex 4.4** Suppose whether it rains tomorrow or not depends on both today's and yesterday's weather.

Today's Weather	Yesterdays's Weather	Value
Rain	Rain	0
Rain	No Rain	1
No Rain	Rain	2
No Rain	No Rain	3

Suppose:

$$\mathbb{P}(\text{rain tomorrow} | \text{rain today, rain yesterday}) = .7$$

$$\mathbb{P}(\text{rain tomorrow} | \text{rain today, no rain yesterday}) = .5$$

$$\mathbb{P}(\text{rain tomorrow} | \text{no rain today, rain yesterday}) = .4$$

$$\mathbb{P}(\text{rain tomorrow} | \text{no rain today, no rain yesterday}) = .2$$

$$P = \begin{bmatrix} .7 & 0 & .3 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & .4 & 0 & .6 \\ 0 & .2 & 0 & .8 \end{bmatrix}$$

## 4.2 Chapman-Kolmogorov Theorem

$P_{ij}$  = probability of going from state i to state j

$P_{ij}^{(n)}$  = probability of going from state i to state j in n steps.

$$P_{ij}^{(n+m)} = \sum_k P_{ik}^{(n)} P_{kj}^{(m)} \quad (\text{pg.197})$$

**Look at example 4.10 for next class**

9/8

### Proof of Chapman-Kolmogorov Theorem

$$\text{Equation: } P_{ij}^{(n+m)} = \sum_k P_{ik}^{(n)} P_{kj}^{(m)}$$

We can visualize this as a graph with n+m steps and we consider all the paths  $i \rightarrow j$  and sum them with the law of total probability.

**Proof:**

$$\begin{aligned} P_{ij}^{(n+m)} &= \mathbb{P}(X_{n+m} = j | X_0 = i) \\ &= \sum_k \mathbb{P}(X_{n+m} = j, X_n = k | X_0 = i) \\ &= \sum_k \mathbb{P}(X_{n+m} = j | X_n = k, X_0 = i) \mathbb{P}(X_n = k | X_0 = i) \end{aligned}$$

Note that this is the probability of going from k to j in m steps (which doesn't depend on  $x_0 = i$  due to the Markov Property) and from i to k in n steps.

Homogeneity of a Markov Chain.

**Example 4.10** An urn always contains 2 balls. Possible ball colors are red and

blue. Each stage of the process we pick a ball and randomly replace it with another ball. Replacement of the same color is .8 and replacement of a different color is .2.

If initially both the first balls are red, what is the probability that the 5th ball is red?

$$P = \begin{bmatrix} .8 & .2 & 0 \\ .1 & .8 & .1 \\ 0 & .2 & .8 \end{bmatrix}$$

Note: for a set up where the probability of changing colors is invariant of the color of the ball, the transition matrix will be visually "radially" symmetric\*\*\*.

$$\begin{aligned} \mathbb{P}(X_5 = \text{red}) &= P_{22}^{(4)} + \frac{1}{2}P_{21}^{(4)} + 0P_{12}^{(4)} \\ &= 0.7048 \end{aligned}$$

Ask what are other Properties of stochastic matrix

$$a_{i,j} = a_{n-i, n-j}$$

#### Example 4.11

In a sequence of independent flips of a fair coin, let  $N$  denote the number of flips until there is a run of 3 heads.

Find (a)  $P(N \leq 8)$  (b)  $P(N = 8)$

Consider 4 states: 0,1,2,3. given by  $n$  = the number of consecutive heads

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

$$(a) = P_{03}^{(8)}$$

$$(b) = \frac{1}{2}P_{02}^{(7)}$$

### 4.3 Classification of States

**Definition:** State  $j$  of is accessible from state  $i$  if  $P_{ij}^{(n)} > 0$  for some  $n \geq 0$ . If the states are accessible from each other, they are said to communicate.

Communication is an equivalence relation.

Reflexive and symmetric are obvious.

Transitive is proven by the Chapman-Kolmogorov Theorem.

This relation divides the states into classes.

### Reccurent and Transient States

**Definition:** A given state  $i$  of a Markov Chain let  $f_i$  denote the probability that the chain will eventually return to state  $i$ .

A state is called **Recurrent** if  $f_i = 1$  and **Transient** if  $f_i < 1$ .

The expected number of revisits to a recurrent state is infinite.

For a transient state the probability of being in state  $i$  for exactly  $n$  times period is  $f_i^n(1 - f_i)$ : Note that this is Geometric distribution

**Lets notice state properties:**

$$\begin{aligned} f_i &= \mathbb{P}(x_{n+N} = i | X_n = i) \\ &= \mathbb{P}(x_N = i | X_0 = i) \end{aligned}$$

A Recurrent state is revisited infinitely often after it is visited once it will be revisited by the markov properties, and it repeats.

A Transient state is revisited only a finite number of times.

**Proof of Transitive state finite recurrence:**

The probability a transient state is revisited exactly  $n$  times is  $f_i^{n-1}(1 - f_i)$

$$\begin{aligned} E(n) &= \sum_{n=1}^{\infty} n f_i^{n-1} (1 - f_i) \\ &= (1 - f_i) \sum_{n=1}^{\infty} \frac{d}{df_i} f_i^n \\ &= (1 - f_i) \frac{d}{df_i} \sum_{n=1}^{\infty} f_i^n \\ &= (1 - f_i) \frac{d}{df_i} \frac{f_i}{1 - f_i} \\ &= \frac{1}{(1 - f_i)} \end{aligned}$$

**Proposition 4.1**

A state is is

1. Recurrent if  $\sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty$
2. Transient if  $\sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty$

**Proof:**

Define  $I_n = \begin{cases} 1 & \text{if } X_n = i \\ 0 & \text{otherwise} \end{cases}$

The number of times period the process is in state  $i$  is  $\sum_{n=0}^{\infty} I_n$

The expected value of the number of times the process is in state  $i$  is

$$\begin{aligned} \mathbb{E}\left(\sum_{n=0}^{\infty} I_n\right) &= \sum_{n=0}^{\infty} \mathbb{E}(I_n) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(x_n = i | x_0 = i) \\ &= \sum_{n=0}^{\infty} P_{ii}^{(n)} \end{aligned}$$

**Corollary 4.2 (pg 207)**

If  $i$  is recurrent and  $i$  communicates with  $j$ , then  $j$  is recurrent.

**Proof:**

$$i \leftrightarrow j \rightarrow \exists k \text{ s.t. } P_{ij}^k > 0 \text{ and } P_{ji}^k$$

For any  $n$ ,

$$\begin{aligned} P_{ij}^{(m+n+k)} &\geq P_{ji}^m P_{ii}^n P_{ij}^k \\ \sum_{n=1}^{\infty} P_{ij}^{(m+n+k)} &\geq \sum_{n=1}^{\infty} P_{ji}^m P_{ii}^n P_{ij}^k \\ \sum_{t=0}^{\infty} P_{jj}^t &\geq \sum_{n=1}^{\infty} P_{ij}^{(m+n+k)} \geq P_{ji}^m P_{ij}^k \sum_{n=1}^{\infty} P_{ii}^n \geq \infty \end{aligned}$$

Thus if  $i$  is recurrent and  $i$  communicates with  $j$ , then  $j$  is recurrent.

**Remark:** If the state  $i$  is transient and if the state  $j$  communicates with  $i$ , then  $j$  is transient.

**Proof:** Assume the if, Suppose  $j$  is not transient. Then  $j$  is recurrent. Then  $i$  is recurrent. This is a contradiction. Thus  $j$  is transient.

**Remark:** Transience and Recurrence are class properties.

**Remark:** Suppose we have a Markov Chain with a finite number of states. Then at least one state is recurrent.

A Markov Chain with exactly one communication class is called irreducible.

A finite state irreducible Markov Chain must have all states recurrent.

**Example 4.18** Consider a Markov Chain with 5 states.

$$P = \begin{bmatrix} .5 & .5 & 0 & 0 & 0 \\ .5 & .5 & 0 & 0 & 0 \\ 0 & 0 & .5 & .5 & 0 \\ 0 & 0 & .5 & .5 & 0 \\ .25 & .25 & 0 & 0 & .5 \end{bmatrix}$$

Find the equivalence classes, classify them as recurrent or transient.

**Solution:**

The equivalence classes are  $\{0, 1\}$  and  $\{2, 3\}$  and  $\{4\}$

4 is its own class due to the fact the communication between 0 and 1 is not

symmetric.

Does number of nonzero eigenvectors equal the number of equivalence classes?

**Example 4.19**

Markov Chain with states  $(0, \pm 1, \dots)$

$P_{i,i+1} = p$  and  $P_{i,i-1} = 1 - p$

pg 209-211

A random walk is symmetric if  $p = 1/2$

Can prove that the random walk recurrent in that case **Remark:**

Definition of Recurrence:

$$f_i = P(\text{Ever coming back to state } i \text{ — starting at state } i)$$

$$f_i = P\left(\sum_n (x_n = i) \mid x_0 = i\right)$$

**Random Walk**

State space is  $\mathcal{Z}$

Transition probabilities are:  $P_{i,i+1} = p$  and  $P_{i,i-1} = 1 - p$

Note: when there is only one equivalence class, the Markov Chain is irreducible.

Find  $f_0 = \beta = P(\text{ever returning to } 0 \text{ — starting at } 0)$

Condition the probability  $\beta$  on the next transition.

$$\beta = (p)P(\text{ever returns to } 0 \mid x_1 = 1) + (1 - p)P(\text{returns to } 0 \mid x_1 = -1)$$

Let  $\alpha = P(\text{ever returns to } 0 \mid x_1 = 1)$

$$\alpha = (p)P(\text{ever returns to } 0 \mid x_1 = 1, x_2 = 2) + (1 - p)P(\text{ever returns to } 0 \mid x_1 = 1, x_2 = 0)$$

$$\alpha = 1 - p + p\alpha^2$$

Solving gives

$$\alpha = \frac{1 - p}{p}$$

If the random walk is symmetric then  $\alpha = 1$  is the only Solution. Substitution in equation with beta gives  $f_0 = 1$

## 4.4 Long Run Proportions and Limiting Probability

Let  $i \neq j$  be states of a Markov Chain.

Define  $f_{ij}$  as the probability that the Markov chain, starting in state  $i$ , will ever reach state  $j$ .

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

**Proposition 4.3**

If the state  $i$  is recurrent and  $i$  communicates with  $j$ , then probability of eventually reaching  $j$  is 1,  $f_{ij} = 1$ .

**Proof:**

$$i \leftrightarrow j \rightarrow \exists n > 0 \text{ s.t. } P_{ij}^{(n)} > 0$$

Assume  $n$  is the minimum such integer.

Since  $i$  is recurrent, the infinite sequence  $0 = k_0 < k_1 < k_2 < \dots$  exists such that  $X_{k_r} = i$  for  $r = 0, 1, 2, 3, \dots$

Define  $z = \min(r > 0, X_{k_r+n} = j)$

Then  $P(Z = z) = P_{ij}^n (1 - P_{ij}^{(n)})^{z-1}$

Thus  $f_{ij} = 1$  by sandwich Theorem

Assume  $j$  is a recurrent state.

Define  $N_j = \min(n > 0 | X_n = j)$

Let  $m_j = E[N_j | x_0 = j]$

It is the expected number of steps to return to  $j$ .

Since we know that  $P(N_j < \infty | x_0 = j) = 1$

Still it may happen that  $m_j = E[N_j | x_0 = j] = \infty$

Definition: if  $m_j < \infty$  then  $j$  is positive recurrent.

If  $m_j = \infty$  then  $j$  is null recurrent.

We define  $\pi_j$  to be the long run proportion of time the Markov chain is in state  $j$ .

$$\pi_j = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} I_k$$

Where  $I_k = 1$  if  $X_k = j$  and 0 otherwise.

Proposition 4.4 :

If the Markov chain is irreducible and recurrent then any initial state  $x_0$  will have  $\pi_j = \frac{1}{m_j}$

At time  $T_0 + \sum T_k$  the chain enters state  $j$  for the  $(n+1)$ th time, the proportion of the time the chain is in state  $j$  during this is  $\frac{n+1}{T_0 + \sum T_k}$

$$\pi_j = \lim_{n \rightarrow \infty} \frac{n+1}{T_0 + \sum T_k}$$

Prop 4.4 if a MC is irreducible and recurrent, then  $\pi_j = \frac{1}{m_j}$

Prop 4.5 if the state  $i$  is positive recurrent and if the state  $j$  communicates with  $i$  then the state  $j$  is also positive recurrent.

Proof:  $i \leftrightarrow j \rightarrow \exists n > 0$  such that  $P_{ij}^n > 0, \pi_i P_{ij}^n =$  Proportion of times that the process will be in state  $j$ ,  $n$  steps after it was in state  $i$ .  $< \pi_j$

$$\pi_i P_{ij}^n \leq \pi_j$$

$\pi_i > 0$  since it is positive recurrent. since  $P$  is also finite thus  $\pi_j$  is positive recurrent.

Remark: null recurrence is also a class property.

**Claim:** an irreducible finite state markov chain is positive recurrent.

**Proof:** Let  $m_j$  be the expected return time to state j.

If you have a finite MC then there is one EC, and if one is null recurrent then all are null recurrent.

Suppose that state i in such MC is null recurrent.

Then  $\pi_j = 0$  Since null recurrence is a class property and there is only one class, thus all states are null recurrent.

$\pi_i = 0$  for all states.

$\sum p_i = 0$  with probability one. This is a contradiction.

**Theorem 4.1** Consider a irreducible Markov Chain. If the chain is positive recurrent then the long run proportions are unique solutions of the system of equations

$$\sum_i \pi_i P_{ij} = \pi_j$$

and  $\sum_j \pi_j = 1$

Think of it like all ways the  $\pi$  go to state j.

Similar to conservation of flow.

**Matrix Intuition:**

Write  $\vec{\pi} = [\pi_0 \ \pi_1 \ \pi_2 \ \dots]$

be the row vector with entries  $\pi_j$

Then  $\vec{\pi}P = \vec{\pi}$

And  $\sum_j \pi_j = 1$

**Example 1**

Consider a two state Markov Chain with transition matrix

$$P = \begin{bmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{bmatrix}$$

Compute the long run proportions  $\pi_0$  and  $\pi_1$

Assume  $\alpha, \beta \neq 0, 1$

$$\pi_0 P_{01} + \pi_1 P_{11} = \pi_1$$

$$\pi_0 P_{00} + \pi_1 P_{10} = \pi_0$$

$$\pi_0 + \pi_1 = 1$$

In matrix formulation we have  $[\pi_0 \ \pi_1] \begin{bmatrix} \alpha & 1 - \alpha \\ \beta & 1 - \beta \end{bmatrix} = [\pi_0 \ \pi_1]$

Short cut to remember

$$\begin{bmatrix} 0 & 1 \\ 0 & P_{00} & P_{01} \\ 1 & P_{10} & P_{11} \\ & \pi_0 & \pi_1 \end{bmatrix}$$

Note that  $\pi_0 = \frac{\beta}{1 - \alpha + \beta}$  and  $\pi_1 = \frac{1 - \alpha}{1 - \alpha + \beta}$

**Example 2**

Doubly stochastic matrix. If the sum of the columns and rows are equal to 1.



If the transition matrix of a Markov Chain with  $n$ -states is a doubly stochastic matrix, then the long run proportions are  $\pi_j = 1/n$  for all  $j$ .

**Proof:**

$$[1/n, 1/n, 1/n, \dots]P = [1/n, 1/n, 1/n, \dots]$$

$$1/n \sum_j P_{ij} = 1/n$$

$$\sum_j P_{ij} = 1$$

Since solutions are unique as in Theorem 4.1, thus  $\pi_j = 1/n$  for all  $j$ .

**Example 3**

Simple random symmetric walk. This is also reflecting at the edges.

States are  $(0, 1, 2, \dots, L)$

$$P_{01} = 1, P_{L,L-1} = 1$$

$$P_{i,i-1} = 1/2 \text{ and } P_{i,i+1} = 1/2$$

Simple case:  $L = 2$

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

$$\pi_0 = 0, \pi_1 = 1/2, \pi_2 = 1/2$$

Case  $L = 3$

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

$$\pi_0 = 1/6, \pi_1 = 1/3, \pi_2 = 1/3, \pi_3 = 1/6$$

Prove that for  $L = n$ ,  $\pi_i = \frac{1}{n}$  for all  $i$  other than  $\pi_0, \pi_L = 1/2L$ .

**Example**

For  $L = 1000$ , what is the probability of revisiting state 0?

it is 2000 as it is  $\pi = 1/m$

If the system is inconsistent the MC is transient or null recurrent.

$\pi_j = 0$  for all  $j$ .

**Example 4.26:**

MC with acceptable status in the set  $A$  and unacceptable status  $A^C$

If  $x_n \in A$  process is "up", if  $x_n \in A^C$  process is "down".

Find:

i: Rate at which the process goes from up to down.

ii: Average length of time process remains down when it goes down.

iii: Average length of time process remains up when it goes up.

**Solution:**

Let  $i \in A$  and  $j \in A^C$

The rate at which the process enters  $j$  from  $i$  is

$$= \pi_i P_{ij}$$

Rare ar which process enters  $j$  from any acceptable state is

$$\sum_{i \in A} \pi_i P_{ij}$$

Rate at which the process goes from  $A \rightarrow A^C$  is

$$\sum_{j \in A^C} \sum_{i \in A} \pi_i P_{ij}$$

Let  $u$  be the average time process stays up.

Let  $d$  be the average time process stays down.

rate at which a breakdown occurs is  $\frac{1}{u+d}$

Therefore

$$\frac{1}{u+d} = \sum_{j \in A^C} \sum_{i \in A} \pi_i P_{ij}$$

Proportion the process is up is

$$\frac{u}{u+d} = \sum_{i \in A} \pi_i$$

Get  $u$  and  $d$  in terms of  $\pi$

$$u = \frac{\sum_{i \in A} \pi_i}{\sum_{j \in A^C} \sum_{i \in A} \pi_i P_{ij}}$$

$$d = \frac{\sum_{j \in A^C} \sum_{i \in A} \pi_i P_{ij}}{\sum_{i \in A^C} \pi_i}$$

### Stationary Probability

If the initial distribution of states is chosen according to long run proportions  $\pi_j$ , then the future distribution of the state of the system will be the same if

$P(x_0 = j) = \pi_j$  then  $P(x_n = j) = \pi_j$

Using induction we can see this true for all  $n$ . (4.4)

**4.4.1: Limiting probabilities: Example:** Consider a two state MC with

$$P = \begin{bmatrix} .7 & .3 \\ .4 & .6 \end{bmatrix}$$

Numerical calculation show that  $P^{(n)}$  converges to a limiting distribution

$$P^n = \begin{bmatrix} 4/7 & 3/7 \\ 4/7 & 3/7 \end{bmatrix}$$

**Claim:** The limiting probabilities  $\lim_{n \rightarrow \infty} P(x_n = j)$  if they exist are equal to the long run proportions  $\pi_j$

**Proof:** Assume  $\alpha_j = \lim_{n \rightarrow \infty} P(x_n = j)$  exists.

Then  $P(x_{n+1} = j) = \sum_i P(x_{n+1} = j | x_n = i) P(x_n = i)$

Gives  $\lim_{n \rightarrow \infty} P(x_{n+1} = j) = \lim_{n \rightarrow \infty} \sum_i P_{ij} P(x_n = i)$

Thus  $\alpha_j = \sum_i P_{ij} \alpha_i$  for all  $j$

Also  $\sum_j \alpha_j = 1$

Recall that  $p_{ij}$  are the unique solutions of the system of equations

$$\sum_i \pi_i P_{ij}, \sum_i \pi_j = 1$$

Therefore  $\lim_{n \rightarrow \infty} \alpha_j = \pi_j$  if  $\alpha_j$  exists

When do limits not exist? When  $n \rightarrow \infty$  diverges or collates

A chain that can only return to a state a multiple of  $d > 1$  times is called a periodic chain. And does not have limiting probabilities.

**Definition:**

An irreducible, positive recurrent, aperiodic Markov Chain is said to be ergodic.