MICT-5101: Probability and Stochastic Process[1]

Dr. Md. Rezaul Karim
PhD(KULeuven & UHasselt), MS(Biostatistics), MS(Statistics)
Professor, Department of Statistics and Data Science
Jahangirnagar University (JU), Savar, Dhaka - 1342, Bangladesh

MS-2024

[1] These course slides should not be reproduced nor used by others (without permission).

Lecture Outline I

Lecture Outline II

Introduction

1 Introduction

1.1 Text & Reference Book List

Text Book

1. Ross, S. (2010): *Introduction to Probability Models*, 10th edition, Pearson, Prentice Hall.
2. Anthony J. Hayter (2012): *Probability and Statistics for Engineers and Scientists* 4th Edition, Duxbury Press.

Reference Book List

1. Mehdhi, J. (2009): *Stochastic Processes* , 3rd Revised Edition, New Age Science.
2. Beichelt F. (2016): *Applied probability and stochastic processes*, 7th edition, CRC Press.
3. Ross, S. (2020): *Introduction to Probability and Statistics for Engineers and Scientists*, 6th Edition, Pearson Education Inc.

Fundamentals of Probability Models

1. **Part I:** Probability Theory
   - Basic Concepts of Probability
   - Random Variable
   - Expectation
   - Some Probability Distributions
     - Bernoulli
     - Binomial
     - Poisson
     - Uniform and
     - Normal
     - exponential
     - ...

2. **Part II:** Stochastic Processes
   - Basics of Stochastic Processes
   - Random Point Processes
   - Discrete-Time Markov Chains
   - ...

Part II: Stochastic Processes

Chapter 1: Markov Chains

2 Chapter 1: Markov Chains

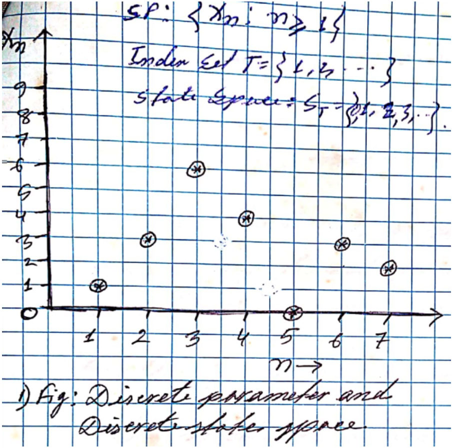
Stochastic Processes

- Stochastic Process: A stochastic process can be defined as a collection (or family) of random variables indexed by time or space.
- Examples:
  - Suppose $X(t)$ be represent the number of customers in the Supermarket at time t. Then $\{X(t); t \in \mathcal{T}\}$ is a family of random variables indexed by the time parameter t, hence the process $X(t)$ is a stochastic process.
  - The daily temperature can be treated as a stochastic process. While there are trends (like seasonal changes), the specific temperature on any given day is influenced by random atmospheric conditions.
  - The daily closing price of a stock can be modeled as a stochastic process. The price changes from day to day are influenced by a variety of random factors, such as market sentiment, news, and economic indicators.
- The indexing parameter can be either continuous or discrete. When the indexing parameter is discrete, generally we denote it by n and represent the stochastic process as $\{X_n; n = 0, 1, 2, \dots\}$.

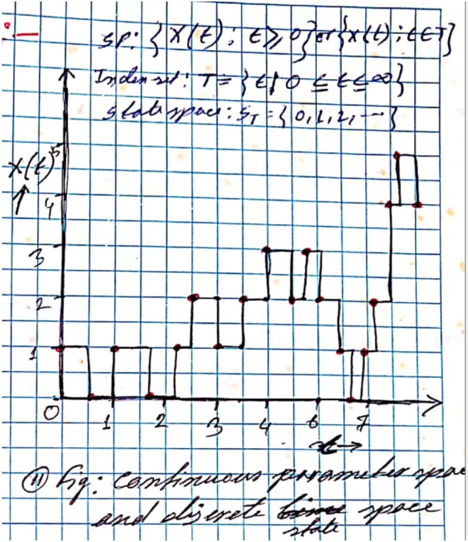
- index set:
  - when $\mathcal{T}$ is a countable set the stochastic process is said to be a discrete-time stochastic process
  - if $\mathcal{T}$ is an interval of the real line, the stochastic process is said to be a continuous-time stochastic process
- State Space: The set of possible values that the random variables $X(t)$ can take, which can be finite, countably infinite, or continuous.
- classification of stochastic process
  - discrete index set and state space
  - continuous index set and discrete state space
  - discrete index set and continuous state space
  - continuous index set and state space

Examples of Stochastic Process

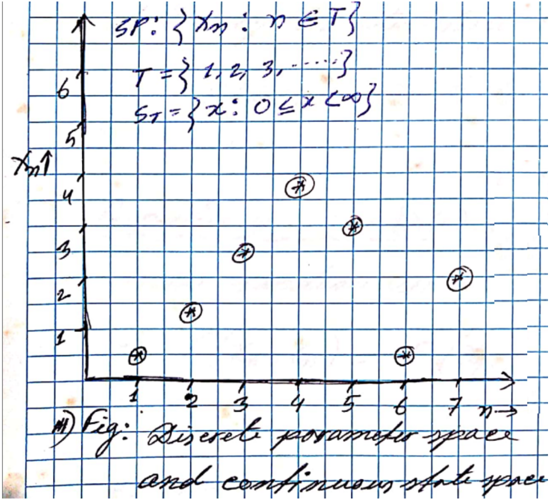
1. **discrete index set and state space**
   - number of the student getting scholarship every year in JU
   - consumer preferences observed on a monthly basis
   - number of the total customer in a supermarket everyday



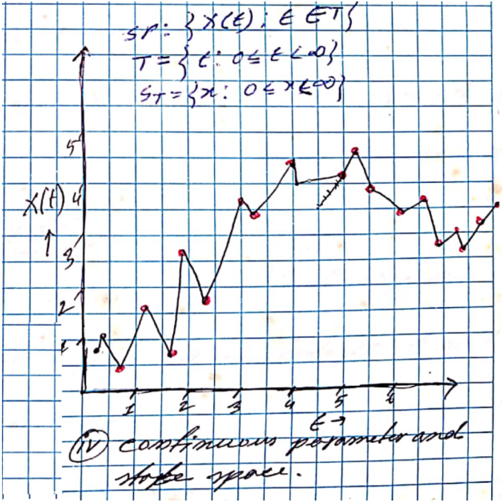
1. **continuous index set and discrete state space**
   - number of incoming calls at a switch board at any time of day
   - number of students waiting for the bus at any time of day



1. **discrete index set and continuous state space**
    - waiting time of the *n*th student arriving at the bus stop
    - maximum temperature at a fixed place every day



1. **continuous index set and state space**
   - maximum temperature in JU at any time of the day
   - content of a dam observed over an interval of time



Stationary Process

- A stationary process is a type of stochastic process whose statistical properties do not change over time. This means that its behavior and characteristics are consistent when observed at different time intervals. Stationarity is a crucial concept in time series analysis and probability theory.
  - Strict Stationarity
  - Weak Stationarity:

Strictly Stationary Process

- A process is strictly stationary if its joint distribution remains the same regardless of shifts in time. This means that all statistical properties, including all moments, are invariant to time shifts. That is, the stochastic process $\{X(t), t \in \mathcal{T}\}$ is said to be stationary in the strict sense if and only if

  $$F(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) = F(x_1, x_2, \ldots, x_n; t_1 + h, t_2 + h, \ldots, t_n + h)$$

  for all integer h and $n \geq 1$
- Example: For example, if you look at the returns from a stock over different time periods (say, from 2000 to 2005 and from 2010 to 2015) and find that they exhibit the same distribution characteristics, this would indicate a strictly stationary process

Weak Stationarity

- A process is weakly stationary (or second-order stationary) if its mean and variance are constant and the autocovariance only depends on the time difference (lag) h. That is, the stochastic process $\{X(t); t \in \mathcal{T}\}$ is said to be weakly stationary (or covariance stationary) if and only if
  - $E[X(t)] = \mu =$ Constant over time
  - $Var[X(t)] = \sigma^2 =$ Constant over time
  - $\gamma(h) = \gamma(t, t + h) =$ Constant over time
- **Monthly Average Temperature:** The monthly average temperature in a city can be modeled as a weakly stationary process.

Example of Weak Stationary: Monthly Average Temperature

- **Mean:** The average temperature over a month is relatively stable from year to year, assuming no significant climate changes.
- **Variance:** The variability of temperatures within each month is also consistent over time. For instance, January may consistently have a certain range of temperatures, and this range doesn't change drastically from year to year.
- **Covariance:** The covariance between the average temperatures of two different months depends only on the time difference between them (e.g., the correlation between January and February would be similar regardless of the specific years being analyzed).

Evolutionary

If the process is not stationary (in any sense), then it is called evolutionary.

Problem 1

Suppose $[X(t), t \in T]$ be a stochastic process where

$$\Pr[X(t) = n] = \frac{e^{-at}(at)^n}{n!} \quad ; \quad n = 0, 1, 2, \ldots, a > 0.$$

Is this process stationary?

Solution

We know that the mean function

$$m(t) = E[X(t)] = \sum_{n=0}^{\infty} n \frac{e^{-at}(at)^n}{n!}$$
$$= at \sum_{n=1}^{\infty} \frac{e^{-at}(at)^{n-1}}{(n-1)!}$$
$$= at\, e^{-at} \sum_{n=1}^{\infty} \frac{(at)^{n-1}}{(n-1)!} \qquad = at\, e^{-at}\, e^{at} \qquad = at$$

and

$$E\left[\{X(t)\}^2\right] = \sum_{n=0}^{\infty} n^2 \frac{e^{-at}(at)^n}{n!} = \sum_{n=0}^{\infty} \{n(n-1)+n\} \frac{e^{-at}(at)^n}{n!}$$
$$= \sum_{n=0}^{\infty} n(n-1) \frac{e^{-at}(at)^n}{n!} + \sum_{n=0}^{\infty} n \frac{e^{-at}(at)^n}{n!}$$
$$= (at)^2\ e^{-at} \sum_{n=2}^{\infty} \frac{(at)^{n-2}}{(n-2)!} + at\, e^{-at} \sum_{n=1}^{\infty} \frac{(at)^{n-1}}{(n-1)!}$$
$$= (at)^2\ e^{-at}\, e^{at} + at\, e^{-at}\, e^{at}$$
$$= (at)^2 + at$$

$$\therefore V[X(t)] = E\left[\{X(t)\}^2\right] - [E\{X(t)\}]^2$$
$$= (at)^2 + at - (at)^2 = at$$

Comment: Since the mean and variance functions of this stochastic process are dependent on time t, so the stochastic process $[X(t), t \in T]$ is evolutionary process.

Problem 2

Consider the process $X(t) = A_1 + A_2 t$ where A_1 *and* A_2 are independent random variables with $E(A_i) = a_i$, $V(A_i) = \sigma_i^2$; $i = 1, 2$. Test the stationarity.

Solution:

We know that the mean function is

$$m(t) = E[X(t)] = E[A_1 + A_2 t] = a_1 + a_2 t$$

$$\begin{aligned} E[X(t)X(s)] &= E[(A_1 + A_2 t)(A_1 + A_2 s)] \\ &= E\left[A_1^2 + A_2^2 ts + (t+s) A_1 A_2\right] \\ &= \sigma_1^2 + a_1^2 + ts\left(\sigma_2^2 + a_2^2\right) + (t+s) a_1 a_2 \end{aligned}$$

$$\text{Sicne } E\left(A_1^2\right) = [E(A_1)]^2 + V(A_1) = a_1^2 + \sigma_1^2$$

$$\begin{aligned} E[X(t)]^2 &= E\left[(A_1 + A_2 t)^2\right] \\ &= E\left[A_1^2 + A_2^2 t^2 + 2A_1 A_2 t\right] \\ &= E\left[A_1^2\right] + t^2 E\left[A_2^2\right] + 2t\, E[A_1 A_2] \\ &= \sigma_1^2 + a_1^2 + t^2\left(\sigma_2^2 + a_2^2\right) + 2t a_1 a_2 \end{aligned}$$

$$
\begin{aligned}
V\left[X\left(t\right)\right] &= E\left[X\left(t\right)\right]^2 - \left[E\left[X\left(t\right)\right]\right]^2 \\
&= \sigma_1^2 + a_1^2 + t^2\left(\sigma_2^2 + a_2^2\right) + 2ta_1a_2 - \left(a_1 + a_2t\right)^2 \\
&= \sigma_1^2 + a_1^2 + t^2\sigma_2^2 + t^2a_2^2 + 2ta_1a_2 - a_1^2 - 2ta_1a_2 - t^2a_2^2 \\
&= \sigma_1^2 + t^2a_2^2
\end{aligned}
$$

$$
\begin{aligned}
C\left(s,t\right) &= cov\left[X\left(s\right), X\left(t\right)\right] \\
&= E\left[X\left(s\right), X\left(t\right)\right] - E\left[X\left(s\right)\right]E\left[X\left(t\right)\right] \\
&= \sigma_1^2 + a_1^2 + \left(s+t\right)a_1a_2 + ts\left(\sigma_2^2 + a_2^2\right) - \left(a_1 + a_2s\right)\left(a_1 + a_2t\right) \\
&= \sigma_1^2 + a_1^2 + sa_1a_2 + ta_1a_2 + ts\sigma_2^2 + tsa_2^2 - a_1^2 - sa_1a_2 - ta_1a_2 - tsa_2^2 \\
&= \sigma_1^2 + ts\sigma_2^2
\end{aligned}
$$

Comment: Since $m\left(t\right) = a_1 + a_2t$ and $V\left[X\left(t\right)\right] = \sigma_1^2 + t^2\sigma_2^2$ are functions of t, so that the process is evolutionary.

Problem 3

Consider the process $X(t) = A\cos\omega t + B\sin\omega t$, where A, B uncorrelated random variables each with are mean 0 and variance 1 and ω is a positive constant. Show that the process is covariance stationary.

Solution:

We have the mean function

$$\begin{aligned} M(t) = E[X(t)] &= \cos\omega t E(A) + \sin\omega t E(B) \\ &= 0 \end{aligned}$$

$$\begin{aligned} &E[X(t)X(s)] = E[(A\cos\omega t + B\sin\omega t)(A\cos\omega s + B\sin\omega s)] \\ &= E\left[A^2\cos\omega t\cos\omega s + AB\cos\omega t\sin\omega s + AB\sin\omega t\cos\omega s + B^2\sin\omega t\sin\omega s\right] \\ &= \cos\omega t\cos\omega s + \sin\omega t\sin\omega s \\ &= \cos(s-t)\omega \end{aligned}$$

$$\begin{aligned} \text{since, } E\left[\{X(t)\}^2\right] &= Var[X(t)] + [E\{X(t)\}]^2 \\ &= 1 \end{aligned}$$

and,

$$\begin{aligned} C(s,t) &= Cov\left\{X(t), X(s)\right\} \\ &= \cos(s-t)\,\omega \end{aligned}$$

Comment: Here the first two moments are finite and the covariance function is a function of $(s-t)$. Thus the process is covariance stationary.

Problem 4

Consider the process $[X(t), t \in T]$ whose probability distribution, under a certain condition, is given by

$$\Pr[X(t) = n] = \frac{(at)^{n-1}}{(1+at)^{n+1}}, \qquad n = 1,\ 2,\ \ldots\ \ldots$$
$$= \frac{at}{1+at}, \qquad n = 0$$

Test the stationarity.

Solution:

We have the mean function

$$\begin{aligned} m(t) = E[X(t)] = \textstyle\sum_{n=0}^{\infty} n \Pr[X(t) = n] = \sum_{n=1}^{\infty} n \frac{(at)^{n-1}}{(1+at)^{n+1}} \\ = \frac{1}{(1+at)^2} \textstyle\sum_{n=1}^{\infty} n \left(\frac{at}{1+at}\right)^{n-1} \\ = \frac{1}{(1+at)^2}\left[1 + 2\frac{at}{1+at} + 3\left(\frac{at}{1+at}\right)^2 + \ \ldots\ \ldots\ \ldots\right] \\ = \frac{1}{(1+at)^2}\left[1 - \frac{at}{1+at}\right]^{-2} = \frac{1}{(1+at)^2}\left[\frac{1}{1+at}\right]^{-2} = 1 \end{aligned}$$

and,

$$E\left\{[X(t)]^2\right\} = \sum_{n=0}^{\infty} n^2 \ \Pr[X(t) = n] = \sum_{n=1}^{\infty} n^2 \ \frac{(at)^{n-1}}{(1+at)^{n+1}}$$
$$= \sum_{n=1}^{\infty} n(n-1) \ \frac{(at)^{n-1}}{(1+at)^{n+1}} + \sum_{n=1}^{\infty} n \frac{(at)^{n-1}}{(1+at)^{n+1}}$$
$$= \frac{at}{(1+at)^3} \sum_{n=2}^{\infty} n(n-1) \left(\frac{at}{1+at}\right)^{n-2} + 1$$
$$= \frac{at}{(1+at)^3} \left[2 + 3 \cdot 2 \frac{at}{1+at} + 4 \cdot 3 \left(\frac{at}{1+at}\right)^2 + 5 \cdot 4 \left(\frac{at}{1+at}\right)^3 + \ \ldots\right.$$
$$= \frac{2at}{(1+at)^3} \left[1 + 3 \frac{at}{1+at} + 6 \left(\frac{at}{1+at}\right)^2 + 10 \left(\frac{at}{1+at}\right)^3 + \ \ldots \ \ldots \ \ldots\right.$$
$$= \frac{2at}{(1+at)^3} \left[1 - \frac{at}{1+at}\right]^{-3} + 1$$
$$= \frac{2at}{(1+at)^3} \left[\frac{1}{1+at}\right]^{-3} + 1 = 2at + 1$$

$\therefore \ Var[X(t)] = 2at$

Comment: Here the first moment is constant but the second moment (and the variance) increases with t. Thus the process $[X(t), t \in T]$ is evolutionary.

What is a Gaussian Process?

- A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian (normal) distribution. That is, if the distribution of $[X(t_1), X(t_2), \ldots, X(t_n)]$ for all $t_1, t_2, \ldots, t_n$ is multivariate normal then $\{X(t); t \in T\}$ is said to be a Gaussian Process.
- Gaussian processes are often used in machine learning and statistics for regression, classification, and spatial modeling.

Martingales Process:

A martingale is a specific type of stochastic process that satisfies certain properties related to conditional expectations. That is, a stochastic process $\{X_n\}_{n=1}^{\infty}$ is called a martingale (or a Martingale process) if,

(i) the expected value of the absolute value of X_n is finite. That is

$$E\left[|X_n|\right] < \infty$$

(ii) the expected value of the next observation, given all previous observations, is equal to the current observation. That is,

$$E\left[X_{n+1}|X_n,\ X_{n-1},\ \cdots,\ X_0\right] = X_n$$

Example:

Let $\{Z_i\}$; $i = 1,\ 2,\ \cdots$ be a sequence of $i.i.d.$ random variables with mean 0 and let $X_n = \sum_{i=1}^{n} Z_i$, then show that $\{X_n\}_{n=1}^{\infty}$ is a martingale.

Solution: We have,

$$E(X_n) = E\left\{\sum_{i=1}^{n} Z_i\right\} = \sum_{i=1}^{n} E\{Z_i\} = 0$$

Since, Z_i's are i.i.d. and with mean 0. We have,

$$\begin{aligned} & X_n = Z_1 + Z_2 + \cdots + Z_n \\ \Rightarrow \quad & X_{n+1} = Z_1 + Z_2 + \cdots + Z_n + Z_{n+1} = X_n + Z_{n+1} \end{aligned}$$

So that,

$$\begin{aligned} & E\{X_{n+1}|X_n,\ X_{n-1},\ \cdots,\ X_1\} = E\{(X_n + Z_{n+1})|X_n,\ X_{n-1},\ \cdots,\ X_1\} \\ & = E\{X_n|X_n,\ X_{n-1},\ \cdots,\ X_1\} + E\{Z_{n+1}|X_n,\ X_{n-1},\ \cdots,\ X_1\} \\ & = X_n + E\{Z_{n+1}\}\ ;\ \text{Since } Z_i \text{ are independent and } E\{Z_{n+1}\} = 0 \end{aligned}$$

So, the process is martingale process because it satisfy the two condition of martingale process. (*Shown*)

Example:

Let $\{Z_i\}$; $i = 1, 2, \cdots$ be a sequence of $i.i.d.$ random variables with $E\{Z_i\} = 1$ and let $X_n = \prod_{i=1}^{n} Z_i$, then show that $\{X_n\}_{n=1}^{\infty}$ is a martingale.

Solution:

We have,

$$\begin{aligned} E(X_n) &= E\{Z_1 \cdot Z_2 \cdot \cdots \cdot Z_n\} \\ &= E\{Z_1\} \cdot E\{Z_2\} \cdot \ \cdots \ \cdot E\{Z_n\} \quad ; \quad \textit{Since they are independent} \\ &= 1 \cdot 1 \cdot \ \cdots \ \cdot 1 = 1 < \infty \end{aligned}$$

Again We have,

$$\begin{aligned} & X_n = Z_1 \cdot Z_2 \cdot \cdots \cdot Z_n \\ \Rightarrow \quad & X_{n+1} = Z_1 \cdot Z_2 \cdot \cdots \cdot Z_n \cdot Z_{n+1} \\ & \quad = X_n \cdot Z_{n+1} \end{aligned}$$

So that,

$$\begin{aligned} E\{X_{n+1}|X_n, X_{n-1}, \cdots, X_1\} &= E\{(X_n \cdot Z_{n+1})|X_n, X_{n-1}, \cdots, X_1\} \\ &= E\{X_n|X_n, X_{n-1}, \cdots, X_1\} \cdot E\{Z_{n+1}|X_n, X_{n-1}, \cdots, X_1\} \\ &= X_n \cdot E\{Z_{n+1}\} \quad ; \quad \textit{Since } Z_{i+1} \textit{ and } X_n, X_{n-1}, \cdots, X_1 \\ &= X_n \quad ; \quad \textit{Since } E\{Z_{n+1}\} = 1 \end{aligned}$$

Markov Chain

A stochastic process is said to be **Markov chain** if the the conditional probability distribution of future states depends only on the present state, not on past states. A discrete-time stochastic process $\{X_n\}_{n\geq 0}$ is said to be discrete Markov chain if the conditional distribution of any future state X_{n+1}, give the past states $X_0, X_1, \ldots, X_{n-1}$ and the present state X_n, is independent of the past states and depends on only on the current state. That is

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i)$$

for all $n \geq 0$ and all $i_0, i_1, \ldots, i_{n-1}, i, j \in S$,

Time-Homogeneous Markov Chain

A time-homogeneous Markov chain is a type of Markov chain in which the transition probabilities are constant over time. This means that the probability of transitioning from one state to another depends only on the current state and the next state, not on the specific time at which the transition occurs. For a time-homogeneous Markov chain, the transition probabilities can be expressed as:

$$P(X_{n+1} = j \mid X_n = i) = p_{ij} \tag{1}$$

where p_{ij} is the probability of transitioning from state i to state j. This probability is the same for all n.

Transition Probability

A transition probability refers to the probability of moving from one state to another in a Markov chain. It quantifies how likely it is for a system to transition from a current state i to a next state j in a given time step. The transition probability from state i to state j is denoted as:

$$P_{ij} = P(X_{n+1} = j \mid X_n = i)$$

Properties:

- Non-negativity: $P_{ij} \geq 0$ for all states i and j.
- Normalization:

  $$\sum_{j \in S} P_{ij} = 1$$

  This ensures that from any state i, the system will transition to some state j with certainty.

Example: Weather Model

Consider a simple weather model with three states:

- Sunny
- Cloudy
- Rainy

The transition probabilities can be represented in the following matrix P:

$$P = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.4 & 0.2 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$$

where:

- Row 1 represents transitions from Sunny
- Row 2 represents transitions from Cloudy
- Row 3 represents transitions from Rainy

This system is a time-homogeneous Markov chain because these transition probabilities do not change over time.

Understanding the Transition Matrix

Each entry P_{ij} in the matrix represents the probability of transitioning from state i to state j:

- $P_{11} = 0.6$: Probability of staying Sunny.
- $P_{12} = 0.3$: Probability of transitioning from Sunny to Cloudy.
- $P_{13} = 0.1$: Probability of transitioning from Sunny to Rainy.

(Similar interpretations apply for the other rows.)

Forecasting Example

Suppose today is Sunny. We can use the transition matrix to forecast the weather for the next day:

- Probability of Sunny tomorrow:

$$P_{11} = 0.6$$

- Probability of Cloudy tomorrow:

$$P_{12} = 0.3$$

- Probability of Rainy tomorrow:

$$P_{13} = 0.1$$

Thus, the forecast for tomorrow would be:

- Sunny: 60%
- Cloudy: 30%
- Rainy: 10%

Multiple Days Forecast

To forecast multiple days ahead, you can raise the transition matrix to a power n corresponding to the number of days:

$$P^n$$

For example, to forecast the weather for two days ahead, calculate P^2 to see the probabilities for the states after two transitions.

Time-Homogeneous vs. Time-Inhomogeneous

- Time-Homogeneous: Transition probabilities do not change over time (constant P_{ij}).
- Time-Inhomogeneous: Transition probabilities can vary with time and may depend on the time step n.

Time-Homogeneous Markov Chain

Example: Weather Model

Consider a weather model with three states: Sunny (S), Cloudy (C), and Rainy (R).

- Transition Probability Matrix:

$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

- Interpretation:
  - If today is Sunny, there is a 60% chance it will be Sunny tomorrow.
  - If today is Cloudy, there is a 40% chance it will be Sunny tomorrow.
  - These probabilities remain constant over time.

Time-Nonhomogeneous Markov Chain

Example: Weather Model Consider the same weather model but with changing probabilities over time.

- **Day 1 Transition Matrix**:

$$P^{(1)} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

- **Day 2 Transition Matrix**:

$$P^{(2)} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$

- Interpretation:
  - On Day 1, if it is Sunny, there is a 60% chance it will be Sunny again.
  - On Day 2, that chance increases to 70%.
  - Transition probabilities change based on the day.

Example 4.1: Forecasting the Weather (See textbook)

Suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today, and not on past weather conditions.

- If it rains today, then it will rain tomorrow with probability α; and if it does not rain today, then it will rain tomorrow with probability β.
- If we say that the process is in state 0 when it rains and state 1 when it does not rain, then the preceding is a two-state Markov chain whose transition probabilities are given by:

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & \alpha & 1-\alpha \\ \hline 1 & \beta & 1-\beta \end{array}$$

Example 4.3 On any given day Gary is either cheerful (C), so-so (S), or glum (G). If he is cheerful today, then he will be C, S, or G tomorrow with respective probabilities 0.5, 0.4, 0.1. If he is feeling so-so today, then he will be C, S, or G tomorrow with probabilities 0.3, 0.4, 0.3. If he is glum today, then he will be C, S, or G tomorrow with probabilities 0.2, 0.3, 0.5.

Letting X_n denote Gary's mood on the nth day, then $\{X_n, n \geqslant 0\}$ is a three-state Markov chain (state $0 = C$, state $1 = S$, state $2 = G$) with transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix} \qquad \blacksquare$$

Example 4.4 (Transforming a Process into a Markov Chain) Suppose that whether or not it rains today depends on previous weather conditions through the last two days. Specifically, suppose that if it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

If we let the state at time n depend only on whether or not it is raining at time n, then the preceding model is not a Markov chain (why not?). However, we can transform this model into a Markov chain by saying that the state at any time is determined by the weather conditions during both that day and the previous day. In other words, we can say that the process is in

state 0	if it rained both today and yesterday,
state 1	if it rained today but not yesterday,
state 2	if it rained yesterday but not today,
state 3	if it did not rain either yesterday or today.

The preceding would then represent a four-state Markov chain having a transition probability matrix

The Weather Prediction Problem

How to model this problem as a Markov Process ?

The state space: 0 = (RR) 1 = (NR) 2 = (RN) 3 = (NN)

The transition matrix:

$$\mathbf{P} = \begin{array}{c|cccc|} & 0(RR) & 1(NR) & 2(RN) & 3(NN) \\ \hline 0\ (RR) & 0.7 & 0 & 0.3 & 0 \\ 1\ (NR) & 0.5 & 0 & 0.5 & 0 \\ 2\ (RN) & 0 & 0.4 & 0 & 0.6 \\ 3\ (NN) & 0 & 0.2 & 0 & 0.8 \\ \hline \end{array}$$

This is a discrete-time Markov process.

Random Walk

- A **random walk** is a stochastic process involving a sequence of steps in a discrete state space, where each step is determined by random choices. It can be defined as a Markov chain where the next state depends only on the current state. For a state i, the probabilities are given by:

$$P(X_{n+1} = i + 1 \mid X_n = i) = p \quad \text{(move right)}$$

$$P(X_{n+1} = i - 1 \mid X_n = i) = 1 - p \quad \text{(move left)}.$$

  Each step is independent, and the process usually starts from an initial state, such as $X_0 = 0$. Random walks are fundamental in various fields, modeling phenomena like diffusion and market fluctuations.
- It is called a symmetric random walk if $p = 0.5$.
- Imagine a person taking steps either forward or backward with equal probability. The position of the person after a number of steps is a classic example of a random walk.

Gambler's Ruin Problem

In the **Gambler's Ruin** problem, a gambler starts with A dollars and plays a series of bets. The outcomes of each bet are determined by random choices. The gambler can either win or lose money in each round.

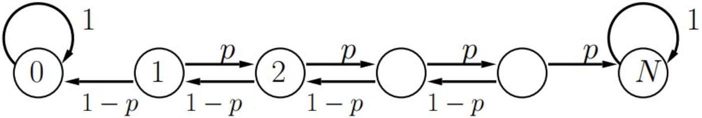
- **Initial State**: The gambler starts with A dollars.
- **Target State**: The gambler aims to reach $N = 5$ dollars.
- **States**: The states of the process are $0, 1, 2, 3, 4, 5$:
  - 0: Absorbing state representing ruin.
  - 5: Absorbing state representing winning.

Transition Probabilities: At each step, the transition probabilities are:

$$P(X_{n+1} = i + 1 \mid X_n = i) = p \quad \text{(win)}$$
$$P(X_{n+1} = i - 1 \mid X_n = i) = 1 - p \quad \text{(lose)}$$

Example 0.1 (Gambler's Ruin). This can be modeled as a Markov chain with state space $S = \{0, 1, 2, \ldots, N\}$ and transition probabilities

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \quad 1 \le i \le N-1$$
$$P_{00} = 1 = P_{NN} \quad \text{(absorbing states)}$$



It is a random walk on a finite state space and with two absorbing barriers.

Transition Matrix

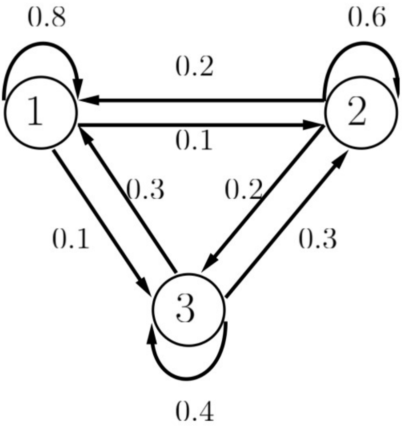
For $N = 5$, the transition matrix P can be represented as follows:

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

Where the rows and columns represent the states $0, 1, 2, 3, 4, 5$ respectively.

Example 0.2 (Social mobility). Let X_n be a family's social class: 1 (lower), 2 (middle), 3 (upper) in the nth generation. We can model this process as a Markov chain with certain kind of transition probabilities such as

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$



Ehrenfest Diffusion Model

- Imagine two urns (or boxes), A and B, that can hold a certain number of balls (particles).
- Initially, some balls are placed in one of the urns.
- Let K be the total number of balls, and let X_n be the number of balls in urn A at step n.

Transition:

- At each step, one ball is chosen uniformly at random and moved to the other urn.
- This process continues indefinitely.

States:

- The state space consists of the possible configurations of the balls in the urns.
- For example, if there are 4 balls, the possible states for X_n (number of balls in urn A) range from 0 to 4.

Transition Probabilities

If there are i balls in urn A (i.e., $X_n = i$), the transition probabilities are:

- $P_{i,i+1} = P(X_{n+1} = i + 1 | X_n = i) = \frac{K-i}{K}$
  (probability of moving a ball from urn B to urn A)
- $P_{i,i-1} = P(X_{n+1} = i - 1 | X_n = i) = \frac{i}{K}$
  (probability of moving a ball from urn A to urn B)
- $P_{ii} = P(X_{n+1} = i | X_n = i) = 0$
  (no self-transition)

Transition Probability Matrix

For example, if there are 4 balls, the possible states for X_n (number of balls in urn A) range from 0 to 4. For $K = 4$ balls, the transition probability matrix P is given by:

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

Joint Probability in Terms of Conditional Probability

The joint probability $P(A, B, C)$ can be expressed as:

$$P(A, B, C) = P(A) \cdot P(B|A) \cdot P(C|A, B)$$

Where:

- $P(A)$ is the probability of event A.
- $P(B|A)$ is the conditional probability of event B given that event A has occurred.
- $P(C|A, B)$ is the conditional probability of event C given that both events A and B have occurred.

Alternative Expressions:

$$P(A, B, C) = P(B) \cdot P(A|B) \cdot P(C|A, B)$$
$$P(A, B, C) = P(C) \cdot P(A|C) \cdot P(B|A, C)$$

Joint Distribution of Random Variables in a Markov Chain

Suppose $\{X_n : \; n = 0, 1, 2, \ldots\}$ is a stationary Markov chain with

- state space $\mathfrak{X}$ and
- transition probabilities $\{P_{ij} : i, j \in \mathfrak{X}\}$.

Define $\pi_0(i) = \mathrm{P}(X_0 = i)$, $i \in \mathfrak{X}$ to be the distribution of X_0.

What is the joint distribution of X_0, X_1, X_2?

$$
\begin{aligned}
&\mathrm{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2)\\
&= \mathrm{P}(X_0 = i_0)\mathrm{P}(X_1 = i_1 | X_0 = i_0)\mathrm{P}(X_2 = i_2 | X_1 = i_1, X_0 = i_0)\\
&= \mathrm{P}(X_0 = i_0)\mathrm{P}(X_1 = i_1 | X_0 = i_0)\mathrm{P}(X_2 = i_2 | X_1 = i_1) \quad (\because \text{ Markov})\\
&= \pi_0(i_0) P_{i_0 i_1} P_{i_1 i_2}
\end{aligned}
$$

In general,

$$
\begin{aligned}
&\mathrm{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, \ldots, X_{n-1} = i_{n-1}, X_n = i_n)\\
&= \pi_0(i_0) P_{i_0 i_1} P_{i_1 i_2} \ldots P_{i_{n-1} i_n}
\end{aligned}
$$

n-Step Transition Probabilities

Suppose $\{X_n\}$ is a stationary Markov chain with state space $\mathfrak{X}$.
Define the n-step transition probabilities

$$P_{ij}^{(n)} = \mathrm{P}(X_{n+k} = j | X_k = i) \quad \text{for } i, j \in \mathfrak{X} \text{ and } n, k = 0, 1, 2, \ldots$$

How to calculate $P_{ij}^{(n)}$?

n-step Transition Probability

In a Markov chain, the n-step transition probability is defined as:

$$P_{ij}^{(n)} = P(X_n = j | X_0 = i) \text{ or, } P_{ij}^{(n)} = P(X_{n+k} = j | X_k = i)$$

This represents the probability of transitioning from state i to state j in n steps.

How to calculate $P_{ij}^{(n)}$?

Example: Ehrenfest Model, 4 Balls

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 2/4 & 0 & 2/4 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Q1 Find $P_{4,2}^{(2)} = \mathrm{P}(X_2 = 2|X_0 = 4)$.
Only one possible path: $4 \to 3 \to 2$,
so $P_{4,2}^{(2)} = P_{4,3}P_{3,2} = 1 \cdot (3/4) = 3/4$.

Q2 Find $P_{4,2}^{(3)} = \mathrm{P}(X_3 = 2|X_0 = 4)$.
Impossible to go from 4 to 2 in odd number of steps,
so $P_{4,2}^{(3)} = 0$.

$$\mathbb{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 2/4 & 0 & 2/4 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Q3 Find $P_{4,2}^{(4)} = \mathrm{P}(X_4 = 2|X_0 = 4)$.

Possible paths:

$$\begin{array}{ccccccccc} 4 & \rightarrow & 3 & \rightarrow & 4 & \rightarrow & 3 & \rightarrow & 2 \\ & & & \searrow & & \nearrow & & \nearrow & \\ & & & & 2 & \rightarrow & 1 & & \end{array}$$

$$P_{4,2}^{(4)} = P_{4,3}P_{3,4}P_{4,3}P_{3,2} + P_{4,3}P_{3,2}P_{2,3}P_{3,2} + P_{4,3}P_{3,2}P_{2,1}P_{1,2}$$
$$= 1 \cdot \frac{1}{4} \cdot 1 \cdot \frac{3}{4} + 1 \cdot \frac{3}{4} \cdot \frac{2}{4} \cdot \frac{3}{4} + 1 \cdot \frac{3}{4} \cdot \frac{2}{4} \cdot \frac{3}{4} = \frac{3}{4}$$

Q4 Find $P_{4,2}^{(10)} = \mathrm{P}(X_{10} = 2|X_0 = 4)$.

Too many paths to list, likely to miss a few.

Chapman-Kolmogorov Equation for Higher Order Transition Probabilities

The Chapman-Kolmogorov equation relates n-step and m-step transition probabilities as follows:

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

Where:

- $P_{ij}^{(n+m)}$: Probability of transitioning from state i to state j in $n+m$ steps.
- $P_{ik}^{(n)}$: Probability of transitioning from state i to intermediate state k in n steps.
- $P_{kj}^{(m)}$: Probability of transitioning from intermediate state k to state j in m steps.
- The sum is taken over all possible intermediate states k in the state space S.

Proof of the Chapman-Kolmogorov Equation

We know,

$$P_{ik}^{(n)} = P(X_n = k|X_0 = i), \quad P_{kj}^{(m)} = P(X_m = j|X_0 = k)$$

Using the Law of Total Probability:

$$P_{ij}^{(n+m)} = P(X_{n+m} = j|X_0 = i) = \sum_{k\in S} P(X_{n+m} = j|X_n = k)P(X_n = k|X_0 = i)$$

Applying the Markov Property:

$$P(X_{n+m} = j|X_n = k) = P_{kj}^{(m)}$$

Thus:

$$P_{ij}^{(n+m)} = \sum_{k\in S} P_{kj}^{(m)} P(X_n = k|X_0 = i) = \sum_{k\in S} P_{ik}^{(n)} P_{kj}^{(m)}$$

Therefore:

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

Chapman-Kolmogorov Equation in Matrix Notation

For $n = 1, 2, 3, \ldots$, let

$$\mathbb{P}^{(n)} = \begin{pmatrix} P_{00}^{(n)} & P_{01}^{(n)} & P_{02}^{(n)} & \cdots & P_{0j}^{(n)} & \cdots \\ P_{10}^{(n)} & P_{11}^{(n)} & P_{12}^{(n)} & \cdots & P_{1j}^{(n)} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ P_{i0}^{(n)} & P_{i1}^{(n)} & P_{i2}^{(n)} & \cdots & P_{ij}^{(n)} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}$$

be the n-step transition probability matrix.

The Chapman-Kolmogorov equation just asserts that

$$\mathbb{P}^{(m+n)} = \mathbb{P}^{(m)} \times \mathbb{P}^{(n)}$$

Note $\mathbb{P}^{(1)} = \mathbb{P}$, $\Rightarrow \mathbb{P}^{(2)} = \mathbb{P}^{(1)} \times \mathbb{P}^{(1)} = \mathbb{P} \times \mathbb{P} = \mathbb{P}^2$.

By induction,

$$\mathbb{P}^{(n)} = \mathbb{P}^{(n-1)} \times \mathbb{P}^{(1)} = \mathbb{P}^{n-1} \times \mathbb{P} = \mathbb{P}^n$$

Example:

Given the transition matrix:

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{bmatrix}$$

To find the 2-step transition probabilities:

$$P^{(2)} = P \cdot P = \begin{bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{bmatrix} \cdot \begin{bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.35 & 0.65 \\ 0.34 & 0.66 \end{bmatrix}$$

Example 4.8 Consider Example 4.1 in which the weather is considered as a two-state Markov chain. If $\alpha = 0.7$ and $\beta = 0.4$, then calculate the probability that it will rain four days from today given that it is raining today.

Solution: The one-step transition probability matrix is given by

$$\mathbf{P} = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix}$$

Hence,

$$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix} \cdot \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix}$$

$$= \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix},$$

$$\mathbf{P}^{(4)} = (\mathbf{P}^2)^2 = \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix} \cdot \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix}$$

$$= \begin{Vmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{Vmatrix}$$

and the desired probability P_{00}^4 equals 0.5749. ■

Example 4.9 Consider Example 4.4. Given that it rained on Monday and Tuesday, what is the probability that it will rain on Thursday?

Solution: The two-step transition matrix is given by

$$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix} \cdot \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix}$$

$$= \begin{Vmatrix} 0.49 & 0.12 & 0.21 & 0.18 \\ 0.35 & 0.20 & 0.15 & 0.30 \\ 0.20 & 0.12 & 0.20 & 0.48 \\ 0.10 & 0.16 & 0.10 & 0.64 \end{Vmatrix}$$

Since rain on Thursday is equivalent to the process being in either state 0 or state 1 on Thursday, the desired probability is given by $P_{00}^2 + P_{01}^2 = 0.49 + 0.12 = 0.61$. ■

Law of Total Probability i n-th Step

Define $\pi_n(i) = \mathrm{P}(X_n = i)$, $i \in \mathfrak{X}$ to be the marginal distribution of X_n, $n = 1, 2, \ldots$. Then again by the law of total probabilities,

$$\begin{aligned}
\pi_n(j) &= \mathrm{P}(X_n = j) \\
&= \sum\nolimits_{k\in\mathfrak{X}} \mathrm{P}(X_0 = k)\mathrm{P}(X_n = j | X_0 = k) \qquad (1) \\
&= \sum\nolimits_{k\in\mathfrak{X}} \pi_0(k) P_{kj}^{(n)}
\end{aligned}$$

Suppose the state space $\mathfrak{X}$ is $\{0, 1, 2, \ldots\}$.
If we write the marginal distribution of X_n as a row vector

$$\pi_n = (\pi_n(0), \pi_n(1), \pi_n(2), \ldots),$$

then the equation (1) is

$$\pi_n = \pi_0 \mathbb{P}^{(n)} = \pi_0 \mathbb{P}^n$$

Example: Ehrenfest Model, 4 Balls

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 4/4 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 2/4 & 0 & 2/4 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 4/4 & 0 \end{pmatrix} \end{array}$$

Q3 Find $P_{4,2}^{(4)} = \mathrm{P}(X_4 = 2 | X_0 = 4)$.

Q4 Find $P_{4,2}^{(10)} = \mathrm{P}(X_{10} = 2 | X_0 = 4)$.

Q5 Given $\mathrm{P}(X_0 = i) = 1/5$ for $i = 0, 1, 2, 3, 4$, find $\mathrm{P}(X_4 = 2)$

$$\mathbb{P}^2 = \mathbb{P} \times \mathbb{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 5/8 & 0 & 3/8 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 3/8 & 0 & 5/8 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \end{pmatrix} \end{array}$$

$$\mathbb{P}^3 = \mathbb{P} \times \mathbb{P}^2 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 5/8 & 0 & 3/8 & 0 \\ 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \\ 0 & 3/8 & 0 & 5/8 & 0 \end{pmatrix} \end{array}$$

$$\mathbb{P}^4 = \mathbb{P}^2 \times \mathbb{P}^2 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 17/32 & 0 & 15/32 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 15/32 & 0 & 5/32 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \end{pmatrix} \end{array}$$

Example: Ehrenfest Model, 4 Balls (Cont'd)

$$\mathbb{P}^4 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 17/32 & 0 & 15/32 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 15/32 & 0 & 5/32 & 0 \\ 3/32 & 0 & \boxed{3/4} & 0 & 5/32 \end{pmatrix} \end{array}$$

For Q3, $\mathrm{P}(X_4 = 2|X_0 = 4) = P_{42}^{(4)} = 3/4.$
which agrees with our previous calculation.

Example: Ehrenfest Model, 4 Balls (Cont'd)

To find $P_{4,2}^{(10)}$ for Q4, it's awful lots of work to compute $\mathbb{P}^{10}\ldots$

There are ways to save some work. By the C-K equation,

$$\mathbb{P}_{4,2}^{(10)} = \underbrace{\mathbb{P}_{4,0}^{(5)}\mathbb{P}_{0,2}^{(5)}}_{=0} + \mathbb{P}_{4,1}^{(5)}\mathbb{P}_{1,2}^{(5)} + \underbrace{\mathbb{P}_{4,2}^{(5)}\mathbb{P}_{2,2}^{(5)}}_{=0} + \mathbb{P}_{4,3}^{(5)}\mathbb{P}_{3,2}^{(5)} + \underbrace{\mathbb{P}_{4,4}^{(5)}\mathbb{P}_{4,2}^{(5)}}_{=0}$$

because it's impossible to move between even states in odd number of moves.

We just need to find $\mathbb{P}_{4,1}^{(5)}$, $\mathbb{P}_{4,3}^{(5)}$, $\mathbb{P}_{1,2}^{(5)}$, and $\mathbb{P}_{3,2}^{(5)}$.

Example: Ehrenfest Model, 4 Balls (Cont'd)

$$\mathbb{P}^5 = \mathbb{P}^2 \times \mathbb{P}^3$$

$$= \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 5/8 & 0 & 3/8 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 3/8 & 0 & 5/8 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \end{pmatrix} \end{array} \times \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 5/8 & 0 & 3/8 & 0 \\ 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \\ 0 & 3/8 & 0 & 5/8 & 0 \end{pmatrix} \end{array}$$

$$= \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} & & 0 & & \\ & & 3/4 & & \\ & & 0 & & \\ & & 3/4 & & \\ 0 & 15/32 & 0 & 17/32 & 0 \end{pmatrix} \end{array}$$

So

$$\mathbb{P}_{4,2}^{(10)} = \mathbb{P}_{4,1}^{(5)}\mathbb{P}_{1,2}^{(5)} + \mathbb{P}_{4,3}^{(5)}\mathbb{P}_{3,2}^{(5)} = \frac{15}{32} \times \frac{3}{4} + \frac{17}{32} \times \frac{3}{4} = \frac{3}{4}.$$

Example: Ehrenfest Model, 4 Balls (Cont'd)

Q5: Given $\mathrm{P}(X_0 = i) = 1/5$ for $i = 0, 1, 2, 3, 4$, find $\mathrm{P}(X_4 = 2)$.

$$\pi_0 = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}).$$

$$\pi_4 = \pi_0 \mathbb{P}^4 = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}) \begin{pmatrix} 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 17/32 & 0 & 15/32 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 15/32 & 0 & 17/32 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \end{pmatrix}$$

$$\pi_4(2) = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}) \begin{pmatrix} 3/4 \\ 0 \\ 3/4 \\ 0 \\ 3/4 \end{pmatrix}$$

$$= \frac{1}{5} \cdot \frac{3}{4} + \frac{1}{5} \cdot 0 + \frac{1}{5} \cdot \frac{3}{4} + \frac{1}{5} \cdot 0 + \frac{1}{5} \cdot \frac{3}{4} = \frac{9}{20}$$

Classification of States

Definition. Consider a Markov chain $\{X_n, n \geq 0\}$ with state space $\mathfrak{X}$. For two states $i, j \in \mathfrak{X}$, we say state j is **accessible** from state i if $P_{ij}^{(n)} > 0$ for some n, and we denote it as

$$i \rightarrow j.$$

Note that **accessibility is transitive**: for $i, j, k \in \mathfrak{X}$,

if $i \rightarrow j$ and $j \rightarrow k$, then $i \rightarrow k$.

Proof.

$$\begin{aligned} i \rightarrow j \quad &\Rightarrow \quad P_{ij}^{(m)} > 0 \text{ for some } m \\ j \rightarrow k \quad &\Rightarrow \quad P_{jk}^{(n)} > 0 \text{ for some } n \end{aligned}$$

By Chapman-Kolmogorov Equation:

$$P_{ik}^{(m+n)} = \sum_{l \in \mathfrak{X}} P_{il}^{(m)} P_{lk}^{(n)} \geq P_{ij}^{(m)} P_{jk}^{(n)} > 0,$$

which shows $i \rightarrow k$.

Definition. Consider a Markov chain $\{X_n, n \geq 0\}$ chain with state space $\mathfrak{X}$. Two states $i, j \in \mathfrak{X}$ are said to **communicate** if $i \rightarrow j$, and $j \rightarrow i$. We denote it as

$$i \longleftrightarrow j.$$

Fact. Communicability is also transitive, meaning that

$$\text{if } i \longleftrightarrow j \text{ and } j \longleftrightarrow k, \text{ then } i \longleftrightarrow k.$$

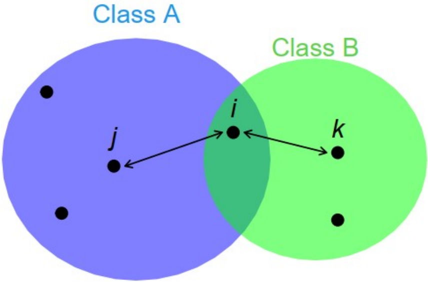
The proof is straight forward from the transitivity of accessibility.

Communicative Class

Definition. Two states that communicate with each other are in the same **class**. A state that communicates with no other states itself is a class.

Fact. Two classes are either identical or disjoint.

Proof. If two classes A and B have one state i in common, then all states in A communicate with i and all states in B do too. Consequently, all states with A can communicate with states in B (through state i). Class A and Class B must be identical.

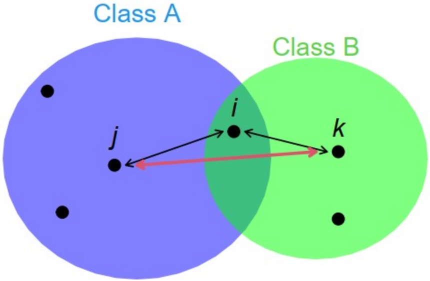


Communicative Class

Definition. Two states that communicate with each other are in the same **class**. A state that communicates with no other states itself is a class.

Fact. Two classes are either identical or disjoint.

Proof. If two classes A and B have one state i in common, then all states in A communicate with i and all states in B do too. Consequently, all states with A can communicate with states in B (through state i). Class A and Class B must be identical.

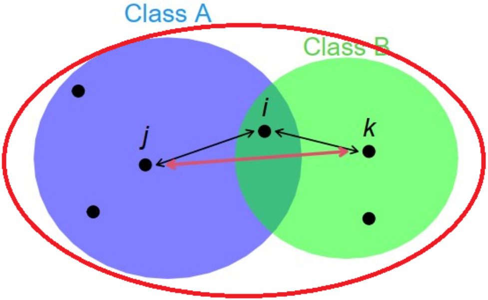


Communicative Class

Definition. Two states that communicate with each other are in the same **class**. A state that communicates with no other states itself is a class.

Fact. Two classes are either identical or disjoint.

Proof. If two classes A and B have one state i in common, then all states in A communicate with i and all states in B do too. Consequently, all states with A can communicate with states in B (through state i). Class A and Class B must be identical.



Example 1. Specify the classes of the following Markov chains.

$$\mathbb{P}_1 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array} \qquad \mathbb{P}_2 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

Example 1. Specify the classes of the following Markov chains.

$$\mathbb{P}_1 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array} \quad \mathbb{P}_2 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

For $\mathbb{P}_1$, $1 \leftrightarrow 2 \rightarrow 3 \leftrightarrow 4$. Classes: $\{1,2\}$, $\{3,4\}$.

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

Example 1. Specify the classes of the following Markov chains.

$$\mathbb{P}_1 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array} \qquad \mathbb{P}_2 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

For $\mathbb{P}_1$, $1 \leftrightarrow 2 \rightarrow 3 \leftrightarrow 4$. Classes: $\{1,2\}$, $\{3,4\}$.

For $\mathbb{P}_2$,
$$\begin{tikzcd} 1 \arrow[r, leftrightarrow] & 2 \\ 4 \arrow[loop below] & 3 \arrow[l] \arrow[u] \arrow[ul] \arrow[loop below] \end{tikzcd}$$
. Classes: $\{1,2\}$, $\{3\}$, $\{4\}$.

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

Example 1. Specify the classes of the following Markov chains.

$$\mathbb{P}_1 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array} \quad \mathbb{P}_2 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

For $\mathbb{P}_1$, $1 \leftrightarrow 2 \rightarrow 3 \leftrightarrow 4$. Classes: $\{1,2\}$, $\{3,4\}$.

For $\mathbb{P}_2$, $\begin{array}{ccc} 1 & \longleftrightarrow & 2 \\ & \nwarrow & \uparrow \\ 4 & \leftarrow & 3 \\ \circlearrowleft & & \circlearrowleft \end{array}$. Classes: $\{1,2\}$, $\{3\}$, $\{4\}$.

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

All states communicate. Only one class.

Closed Classes

Definition. A class C is said to be closed if

$$P_{ij} = 0 \quad \text{for all } i \text{ in } C \text{ and } j \text{ not in } C.$$

Once the process gets into a closed class. It will never leave the class since the outgoing probabilities from the class are all 0.

Examples.

- For $\mathbb{P}_1$ in the previous slide, the class $\{1,2\}$ is not closed because it has a non-zero outgoing probability $P_{23} = 0.1 > 0$. The class $\{3,4\}$ is closed.
- For $\mathbb{P}_2$ in the previous slide, the classes $\{1,2\}$ and $\{4\}$ are closed, and $\{3\}$ is not closed.

A Markov Chain Restricted to a Closed Class is Also a Markov Chain

Example.

$$\mathbb{P}_1 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array}$$

- For $\mathbb{P}_1$ above, the Markov chain restricted to the class $\{3,4\}$ is also a Markov chain, with transition matrix

$$\begin{array}{c c} & \begin{array}{cc} 3 & 4 \end{array} \\ \begin{array}{c} 3 \\ 4 \end{array} & \begin{pmatrix} 0.2 & 0.8 \\ 0.9 & 0.1 \end{pmatrix} \end{array}$$

- The Markov chain for $\mathbb{P}_1$ can not be restricted to $\{1,2\}$ as it may transit out of the state space from 2 to 3.

Irreducibility

A Markov chain is said to be **irreducible** if it has only 1 class.

Recurrence & Transience

Consider a Markov chain $\{X_n, n \geq 0\}$ chain with state space $\mathfrak{X}$.
For $i \in \mathfrak{X}$, define

$$f_i = \mathrm{P}(X_n = i \text{ for some n} > 0 | X_0 = i)$$

If $f_i = 1$, we say state i is **recurrent**
If $f_i < 1$, we say state i is **transient**

- It's generally difficult to compute f_i directly.
  We need other tools to determine whether a state is recurrent or transient.

States in a Non-Closed Class Are Always Transient

For a class A that is NOT closed, there must exists some state k not in A such that

$$P_{i_0,k} > 0, \quad \text{for some state } i_0 \text{ in class } A$$

but

$$P_{ki}^{(n)} = 0 \quad \text{for all state } i \text{ in class } A \text{ and for all } n.$$

Otherwise, state i would be accessible from state k ($k \longrightarrow i$). As i and i_0 are in the same class, we know $i \longleftrightarrow i_0$. Combining all the above, we have

$$k \longrightarrow i \longleftrightarrow i_0 \xrightarrow{P_{i_0,k}>0} k.$$

and hence k would communicate with i_0, contradicting to the assumption that k is not in A.

States in a Non-Closed Class Are Always Transient

For a class A that is NOT closed, there must exists some state k not in A such that

$$P_{i_0,k} > 0, \quad \text{for some state } i_0 \text{ in class } A$$

but

$$P_{ki}^{(n)} = 0 \quad \text{for all state } i \text{ in class } A \text{ and for all } n.$$

Otherwise, state i would be accessible from state k ($k \longrightarrow i$). As i and i_0 are in the same class, we know $i \longleftrightarrow i_0$. Combining all the above, we have

$$k \longrightarrow i \longleftrightarrow i_0 \overset{P_{i_0,k}>0}{\longrightarrow} k.$$

and hence k would communicate with i_0, contradicting to the assumption that k is not in A.

Starting from a state j in a non-closed class A, there is a positive probability that the Markov chain will move to state k and never comes back to the class. Hence state j must be transient.

States in a Non-Closed Class Are Always Transient

For a class A that is NOT closed, there must exists some state k not in A such that

$$P_{i_0,k} > 0, \quad \text{for some state } i_0 \text{ in class } A$$

but

$$P_{ki}^{(n)} = 0 \quad \text{for all state } i \text{ in class } A \text{ and for all } n.$$

Otherwise, state i would be accessible from state k ($k \longrightarrow i$). As i and i_0 are in the same class, we know $i \longleftrightarrow i_0$. Combining all the above, we have

$$k \longrightarrow i \longleftrightarrow i_0 \xrightarrow{P_{i_0,k}>0} k.$$

and hence k would communicate with i_0, contradicting to the assumption that k is not in A.

Starting from a state j in a non-closed class A, there is a positive probability that the Markov chain will move to state k and never comes back to the class. Hence state j must be transient.

Are states in an closed class always recurrent?

Fact 1 If state i is recurrent, then starting from state i, the process will revisit state i infinitely often.

Fact 2 If state i is transient, then starting from state i, the number of times the process revisits state i is finite, with expected value $1/(1-f_i)$.

Reason: Let N_i be the number of times the process revisits state i after starting from i. Observe that

$$\begin{aligned} \mathrm{P}(N_i = k) = & \ \mathrm{P}(\text{returns to } i \text{ after 1st departure}) \\ & \times \cdots \times \mathrm{P}(\text{returns to } i \text{ after } k\text{th departure}) \\ & \times \mathrm{P}(\text{never returns to } i \text{ after k+1st departure}) \\ = & \ f_i^k(1-f_i), \quad k = 0, 1, 2, \ldots \end{aligned}$$

i.e., N_i has a geometric distribution with mean $1/(1-f_i)$.

Claim:

$$\mathbb{E}(\#\text{ of visit to state } i|X_0 = i) = \sum_{n=1}^{\infty} P_{ii}^{(n)}$$

Proof. Define

$$I_{ni} = \begin{cases} 1 & \text{if } X_n = i \\ 0 & \text{if } X_n \neq i \end{cases}, \quad n \geq 0,\ i \in \mathfrak{X}.$$

Observe that $\sum_{n=1}^{\infty} I_{ni}$ is the number of visits to state i.

$$\begin{aligned} \mathbb{E}\left[\sum_{n=1}^{\infty} I_{ni} \middle| X_0 = i\right] &= \sum_{n=1}^{\infty} \mathbb{E}[I_{ni}|X_0 = i] \\ &= \sum_{n=1}^{\infty} \mathrm{P}(X_n = i|X_0 = i) \\ &= \sum_{n=1}^{\infty} P_{ii}^{(n)} \end{aligned}$$

Conclusion from Fact 1, Fact 2 and the Claim Above:

$$\begin{aligned} \text{State } i \text{ is recurrent} &\iff \mathbb{E}(\#\text{ of visit to state } i|X_0 = i) = \infty. \\ &\iff \sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty \end{aligned}$$

Proposition 4.1

$$\text{State } i \text{ is } \begin{cases} \text{recurrent if} & \sum_{n=1}^{\infty} P_{ii}^{(n)} = \infty \\ \text{transient if} & \sum_{n=1}^{\infty} P_{ii}^{(n)} < \infty \end{cases}$$

Implication of Proposition 4.1:

States in a <u>finite-state</u> Markov chain CANNOT be all transient.

Reason. Observe that $\sum_{i \in X} I_{ni} = 1$ for all n since X_n must be in one of the states. Thus

$$\sum_{n=1}^{\infty} \sum_{i \in X} I_{ni} = \sum_{n=1}^{\infty} 1 = \infty.$$

Since $\mathfrak{X}$ is finite, there exists at least one state i such that

$$\sum_{n=1}^{\infty} I_{ni} = \infty.$$

Such states are recurrent. Otherwise $\sum_{n=1}^{\infty} \sum_{i \in X} I_{ni}$ will be $< \infty$.

Corollary 4.2

If $i \longleftrightarrow j$, and i is recurrent, then j is also recurrent.

Proof.

$$\begin{aligned} i \to j \quad &\Rightarrow \quad P_{ij}^{(k)} > 0 \text{ for some } k \\ j \to i \quad &\Rightarrow \quad P_{ji}^{(l)} > 0 \text{ for some } l \end{aligned}$$

By Chapman-Kolmogorov Equation:

$$P_{jj}^{(l+n+k)} \geq P_{ji}^{(l)} P_{ii}^{(n)} P_{ij}^{(k)}, \quad \text{for all } k = 0, 1, 2, \ldots$$

Thus

$$\sum\nolimits_{n=1}^{\infty} P_{jj}^{(n)} \geq \sum\nolimits_{n=1}^{\infty} P_{jj}^{(l+n+k)} \geq \underbrace{P_{ji}^{(l)}}_{>0} \underbrace{\sum\nolimits_{n=1}^{\infty} P_{ii}^{(n)}}_{=\infty} \underbrace{P_{ij}^{(k)}}_{>0} = \infty$$

Corollary 4.2 implies that all states of a finite irreducible Markov chain are recurrent.

Example: One-Dimensional Random Walk

$$X_{n+1} = \begin{cases} X_n + 1 & \text{with prob. } p \\ X_n - 1 & \text{with prob. } 1-p \end{cases}$$

- State space $\{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$
- All states communicate

$$\cdots \longleftrightarrow -2 \longleftrightarrow -1 \longleftrightarrow 0 \longleftrightarrow 1 \longleftrightarrow 2 \longleftrightarrow \cdots$$

Only one class $\Rightarrow$ Irreducible

$\Rightarrow$ States are all transient or all recurrent.

It suffices to check whether 0 is recurrent or transient, i.e., whether

$$\sum_{n=1}^{\infty} P_{00}^{(n)} = \infty \text{ or } < \infty$$

Example: One-Dimensional Random Walk (Cont'd)

$$P_{00}^{(2n+1)} = 0 \quad \text{(Why?)}$$

$$\begin{aligned}
P_{00}^{(2n)} &= \binom{2n}{n} p^n(1-p)^n \\
&= \frac{(2n)!}{n!\,n!} p^n(1-p)^n \qquad \boxed{\text{Stirlin's Formula: } n! \approx n^{n+0.5}e^{-n}\sqrt{2\pi}} \\
&\approx \frac{(2n)^{2n+0.5}e^{-2n}\sqrt{2\pi}}{(n^{n+0.5}e^{-n}\sqrt{2\pi})^2} p^n(1-p)^n \\
&= \frac{1}{\sqrt{\pi n}}[4p(1-p)]^n
\end{aligned}$$

Thus

$$\sum_{n=1}^{\infty} P_{ii}^{2n} \approx \sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}}[4p(1-p)]^n \begin{cases} < \infty & \text{if } p \neq 1/2 \\ = \infty & \text{if } p = 1/2 \end{cases}$$

Conclusion: One-dimensional random walk is recurrent if $p = 1/2$, and transient otherwise.

Stationary Distribution

Define $\pi_i^{(n)} = \mathrm{P}(X_n = i)$, $i \in \mathfrak{X}$ to be the marginal distribution of X_n, $n = 1, 2, \ldots,$ and let $\pi^{(n)}$ be the row vector

$$\pi^{(n)} = (\pi_0^{(n)}, \pi_1^{(n)}, \pi_2^{(n)}, \ldots),$$

From Chapman-Kolmogrov Equation, we know that

$$\pi^{(n)} = \pi^{(n-1)}\mathbb{P} \quad \text{i.e.} \quad \pi_j^{(n)} = \sum_{i \in \mathfrak{X}} \pi_i^{(n-1)} P_{ij} \text{ for all } j \in \mathfrak{X},$$

If π is a distribution on $\mathfrak{X}$ satisfying

$$\pi\mathbb{P} = \pi \quad \text{i.e.} \quad \pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij} \text{ for all } j \in \mathfrak{X},$$

then $\pi^{(0)} = \pi$ implies $\pi^{(n)} = \pi$ for all n.

We say π is a **stationary distribution** of the Markov chain.

Example 1: 2-state Markov Chain

$$\mathfrak{X} = \{0, 1\}, \qquad \mathbb{P} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

$$\pi\mathbb{P} = \pi \Rightarrow \begin{cases} \pi_0 & = (1-\alpha)\pi_0 + \beta\pi_1 \\ \pi_1 & = \alpha\pi_0 + (1-\beta)\pi_1 \end{cases}$$

$$\Rightarrow \begin{cases} \alpha\pi_0 & = \beta\pi_1 \\ \beta\pi_1 & = \alpha\pi_0 \end{cases}$$

Need one more constraint: $\pi_0 + \pi_1 = 1$

$$\Rightarrow \pi = (\pi_0, \pi_1) = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta} \right)$$

Example 2: Ehrenfest Diffusion Model with N Balls

$$P_{ij} = \begin{cases} \dfrac{i}{N} & \text{if } j = i-1 \\ \dfrac{N-i}{N} & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned}
\pi_0 &= \pi_1 P_{10} = \tfrac{1}{N}\pi_1 \Rightarrow \pi_1 = N\pi_0 = \tbinom{N}{1}\pi_0 \\
\pi_1 &= \pi_0 P_{01} + \pi_2 P_{21} = \pi_0 + \tfrac{2}{N}\pi_2 \Rightarrow \pi_2 = \tfrac{N(N-1)}{2}\pi_0 = \tbinom{N}{2}\pi_0 \\
\pi_2 &= \pi_1 P_{12} + \pi_3 P_{32} = \tfrac{N-1}{N}\pi_1 + \tfrac{3}{N}\pi_3 \Rightarrow \pi_3 = \tfrac{N(N-1)(N-2)}{6}\pi_0 = \tbinom{N}{3}\pi_0 \\
&\ \vdots \qquad\qquad \vdots
\end{aligned}$$

In general, you'll get $\pi_i = \binom{N}{i}\pi_0$.
As $1 = \sum_{i=0}^{N}\pi_i = \pi_0 \sum_{i=0}^{N}\binom{N}{i}$ and $\sum_{i=0}^{N}\binom{N}{i} = 2^N$, we have

$$\pi_i = \binom{N}{i}\left(\frac{1}{2}\right)^N \quad \text{for } i = 0, 1, 2, \ldots, N.$$

Stationary Distribution May Not Be Unique

Consider a Markov chain with transition matrix $\mathbb{P}$ of the form

$$\mathbb{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} * & * & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix} \end{array} = \begin{pmatrix} \mathbb{P}_x & 0 \\ 0 & \mathbb{P}_y \end{pmatrix}$$

This Markov chain has 2 classes $\{0,1\}$ and $\{2, 3, 4\}$, both closed and recurrent. So this Markov chain can be reduced to two sub-Markov chains, one with state space $\{0,1\}$ and the other $\{2, 3, 4\}$. Their transition matrices are respectively $\mathbb{P}_x$ and $\mathbb{P}_y$.

Say $\pi_x = (\pi_0, \pi_1)$ and $\pi_y = (\pi_2, \pi_3, \pi_4)$ be respectively the stationary distributions of the two sub-Markov chains, i.e.,

$$\pi_x \mathbb{P}_x = \pi_x, \quad \pi_y \mathbb{P}_y = \pi_y$$

Verify that $\pi = (c\pi_0, c\pi_1, (1-c)\pi_2, (1-c)\pi_3, (1-c)\pi_4)$ is a stationary distribution of $\{X_n\}$ for any c between 0 and 1.

Not All Markov Chains Have a Stationary Distribution

For one-dimensional symmetric random walk, the transition probabilities are

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

The stationary distribution $\{\pi_j\}$ would satisfy the equation:

$$\pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij} = \frac{1}{2}\pi_{j-1} + \frac{1}{2}\pi_{j+1}.$$

Once π_0 and π_1 are determined, all π_j's can be determined from the equations as

$$\pi_j = \pi_0 + (\pi_1 - \pi_0)j, \quad \text{for all integer } j.$$

As $\pi_j \geq 0$ for all integer j, $\Rightarrow \pi_1 = \pi_0$. Thus

$$\pi_j = \pi_0 \quad \text{for all integer } j$$

Impossible to make $\sum_{j=-\infty}^{\infty} \pi_j = 1$.

Conclusion: 1-dim symmetric random walk does not have a stationary distribution.

Limiting Distribution

A Markov chain is said to have a **limiting distribution** if for all $i, j \in \mathfrak{X}$, $\pi_j = \lim_{n\to\infty} P_{ij}^{(n)}$ exists, independent of the initial state X_0, and π_j's satisfy $\sum_{j\in\mathfrak{X}} \pi_j = 1$.

$$\text{i.e.,} \qquad \lim_{n\to\infty} \mathbb{P}^{(n)} = \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Limiting Distribution is a Stationary Distribution

The limiting distribution of a Markov chain is a stationary distribution of the Markov chain.

Proof *(not rigorous)*. By Chapman Kolmogorov Equation,

$$P_{ij}^{(n+1)} = \sum_{k \in \mathfrak{X}} P_{ik}^{(n)} P_{kj}$$

Letting $n \to \infty$, we get

$$\begin{aligned}
\pi_j = \lim_{n \to \infty} P_{ij}^{(n+1)} &= \lim_{n \to \infty} \sum_{k \in \mathfrak{X}} P_{ik}^{(n)} P_{kj} \\
&=^* \sum_{k \in \mathfrak{X}} \lim_{n \to \infty} P_{ik}^{(n)} P_{kj} \quad \text{(needs justification)} \\
&= \sum_{k \in \mathfrak{X}} \pi_k P_{kj}
\end{aligned}$$

Thus the limiting distribution π_j's satisfies the equations $\pi_j = \sum_{k \in \mathfrak{X}} \pi_k P_{kj}$ for all $j \in \mathfrak{X}$ and is a stationary distribution.

See Karlin & Taylor (1975), Theorem 1.3 on p.85-86 for a rigorous proof.

Limiting Distribution is Unique

If a Markov chain has a limiting distribution π, then

$$\lim_{n\to\infty} \pi_j^{(n)} = \pi_j \text{ for all } j \in \mathfrak{X}, \text{ whatever } \pi^{(0)} \text{ is}$$

Proof *(not rigorous)*. Since

$$\pi_j^{(n)} = \sum_{k\in\mathfrak{X}} \pi_k^{(0)} P_{kj}^{(n)}$$

Letting $n \to \infty$, we get

$$\begin{aligned}
\lim_{n\to\infty} \pi_j^{(n)} &= \lim_{n\to\infty} \sum_{k\in\mathfrak{X}} \pi_k^{(0)} P_{kj}^{(n)} \\
&=^* \sum_{k\in\mathfrak{X}} \pi_k^{(0)} \lim_{n\to\infty} P_{kj}^{(n)} \quad \text{(needs justification)} \\
&= \underbrace{\sum_{k\in\mathfrak{X}} \pi_k^{(0)}}_{=1} \pi_j = \pi_j
\end{aligned}$$

i.e., if a limiting distribution exists, it is the unique stationary distribution.

Example: Two-State Markov Chain

$$\mathfrak{X} = \{0, 1\}, \qquad \mathbb{P} = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

By induction, one can show that

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

$$\to \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix} \quad \text{as } n \to \infty$$

The limiting distribution π is $\left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$.

Not All Markov Chains Have Limiting Distributions

Consider the simple random walk X_n on $\{0, 1, 2, 3, 4\}$ with absorbing boundary at 0 and 4. That is,

$$X_{n+1} = \begin{cases} X_n + 1 & \text{with probability } 0.5 \text{ if } 0 < X_n < 4 \\ X_n - 1 & \text{with probability } 0.5 \text{ if } 0 < X_n < 4 \\ X_n & \text{if } X_n = 0 \text{ or } 4 \end{cases}$$

The transition matrix is hence

$$\mathbb{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Not All Markov Chains Have Limiting Distributions

The n-step transition matrix of the simple random walk X_n on $\{0, 1, 2, 3, 4\}$ with absorbing boundary at 0 and 4 can by shown by induction using the Chapman-Kolmogorov Equation to be

$$\mathbb{P}^{(2n-1)} = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.75 - 0.5^{n+1} & 0 & 0.5^n & 0 & 0.25 - 0.5^{n+1} \\ 0.5 - 0.5^n & 0.5^n & 0 & 0.5^n & 0.5 - 0.5^n \\ 0.25 - 0.5^{n+1} & 0 & 0.5^n & 0 & 0.75 - 0.5^{n+1} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\mathbb{P}^{(2n)} = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.75 - 0.5^{n+1} & 0.5^{n+1} & 0 & 0.5^{n+1} & 0.25 - 0.5^{n+1} \\ 0.5 - 0.5^{n+1} & 0 & 0.5^n & 0 & 0.5 - 0.5^{n+1} \\ 0.25 - 0.5^{n+1} & 0.5^{n+1} & 0 & 0.5^{n+1} & 0.75 - 0.5^{n+1} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Not All Markov Chains Have Limiting Distributions

The limit of the n-step transition matrix as $n \to \infty$ is

$$\mathbb{P}^{(n)} \to \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \!\!\begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.75 & 0 & 0 & 0 & 0.25 \\ 0.5 & 0 & 0 & 0 & 0.5 \\ 0.25 & 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Though $\lim_{n\to\infty} P_{ij}^{(n)}$ exists but the limit depends on the initial state i, this Markov chain has no limiting distribution.

This Markov chain has two distinct absorbing states 0 and 4. Other transient states may be absorbed to either 0 or 4 with different probabilities depending how close those states are to 0 or 4.

Periodicity

A state of a Markov chain is said to have **period** d if

$$P_{ii}^{(n)} = 0, \quad \text{whenever } n \text{ is not a multiple of } d$$

In other words, d is the *greatest common divisor* of all the n's such that

$$P_{ii}^{(n)} > 0$$

We say a state is **aperiodic** if $d = 1$, and **periodic** if $d > 1$.

Fact: Periodicity is a class property.
That is, all states in the same class have the same period.

For a proof, see Problem 2&3 on p.77 of Karlin & Taylor (1975).

Examples (Periodicity)

- All states in the Ehrenfest diffusion model are of period $d = 2$ since it's impossible to move back to the initial state in odd number of steps.
- 1-D (2-D) Simple random walk on all integers (grids on a 2-d plane) are of period $d = 2$
- Suppose a 2-D random walk can move to the nearest grid point in any direction, horizontally, vertically or diagonally, each with probability $1/8$.

$$\begin{array}{ccccc} \bullet & & \bullet & & \bullet \\ & \nwarrow & \uparrow & \nearrow & \\ \bullet & \leftarrow & \bullet & \rightarrow & \bullet \\ & \swarrow & \downarrow & \searrow & \\ \bullet & & \bullet & & \bullet \end{array}$$

What is the period of this Markov chain?

Example (Periodicity)

Specify the classes of a Markov chain with the following transition matrix, and find the periodicity for each state.

$$
\begin{array}{c}
\\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array}
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{pmatrix}
0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 0 & 0 & 0.5 & 0 \\
0 & 0 & 0.5 & 0 & 0.5 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \\
0 & 0 & 0 & 0 & 0 & 0.7 & 0.3
\end{pmatrix}
\end{array}
$$

$$
\begin{array}{ccccc}
5 & \rightarrow & 1 & \rightarrow & 2 \\
\uparrow & \swarrow & \uparrow & \swarrow & \\
4 & \rightarrow & 3 & & \\
 & & \downarrow & & \\
7 & \leftrightarrow & 6 & &
\end{array}
$$

Classes: $\{1,2,3,4,5\}$, $\{6,7\}$.
Period is $d = 1$ for state 6 and 7.
Period is $d = 3$ for state 1,2,3,4,5 since
$\{1\} \rightarrow \{2,4\} \rightarrow \{3,5\} \rightarrow \{1\}$.

Periodic Markov Chains Have No Limiting Distributions

For example, in the Ehrenfest diffusion model with 4 balls, it can be shown by induction that the $(2n-1)$-step transition matrix is

$$
\mathbb{P}^{(2n-1)} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}
\begin{array}{c}
\begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\
\begin{pmatrix}
0 & 1/2+1/2^{2n-1} & 0 & 1/2-1/2^{2n-1} & 0 \\
1/8+1/2^{2n+1} & 0 & 3/4 & 0 & 1/8-1/2^{2n+1} \\
0 & 1/2 & 0 & 1/2 & 0 \\
1/8-1/2^{2n+1} & 0 & 3/4 & 0 & 1/8+1/2^{2n+1} \\
0 & 1/2-1/2^{2n-1} & 0 & 1/2+1/2^{2n-1} & 0
\end{pmatrix}
\end{array}
$$

$$
\to \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}
\begin{array}{c}
\begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\
\begin{pmatrix}
0 & 1/2 & 0 & 1/2 & 0 \\
1/8 & 0 & 3/4 & 0 & 1/8 \\
0 & 1/2 & 0 & 1/2 & 0 \\
1/8 & 0 & 3/4 & 0 & 1/8 \\
0 & 1/2 & 0 & 1/2 & 0
\end{pmatrix}
\end{array}
\quad \text{as } n \to \infty.
$$

Periodic Markov Chains Have No Limiting Distributions

and the $2n$-step transition matrix is

$$\mathbb{P}^{(2n)} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/8+1/2^{2n+1} & 0 & 3/4 & 0 & 1/8-1/2^{2n+1} \\ 0 & 1/2+1/2^{2n+1} & 0 & 1/2-1/2^{2n+1} & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 1/2-1/2^{2n+1} & 0 & 1/2+1/2^{2n+1} & 0 \\ 1/8-1/2^{2n+1} & 0 & 3/4 & 0 & 1/8+1/2^{2n+1} \end{pmatrix} \end{array}$$

$$\to \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \end{pmatrix} \end{array} \quad \text{as } n \to \infty.$$

Periodic Markov Chains Have No Limiting Distributions

In general for Ehrenfest diffusion model with N balls, as $n \to \infty$,

$$P_{ij}^{(2n)} \to \begin{cases} 2\binom{N}{j}\left(\frac{1}{2}\right)^N & \text{if } i+j \text{ is even} \\ 0 & \text{if } i+j \text{ is odd} \end{cases}$$

$$P_{ij}^{(2n+1)} \to \begin{cases} 0 & \text{if } i+j \text{ is even} \\ 2\binom{N}{j}\left(\frac{1}{2}\right)^N & \text{if } i+j \text{ is odd} \end{cases}$$

$\lim_{n\to\infty} P_{ij}^{(n)}$ doesn't exist for all $i, j \in \mathfrak{X}$

Summary

- Stationary distribution may not be unique if the Markov chain is not irreducible
- Stationary distribution may not exist
- A limiting distribution is always a stationary distribution
- If it exists, limiting distribution is unique
- Limiting distribution do not exist if the Markov chain is periodic

Positive Recurrence and Null Recurrence

For a Markov chain, consider the return time to a recurrent state i

$$T_i = \min\{n > 0 : X_n = i | X_0 = i\}$$

We say a state i is

- **positive recurrent** if $\mathbb{E}[T_i] < \infty$.
- **null recurrent** if $\mathrm{P}(T_i < \infty) = 1$ but $\mathbb{E}[T_i] = \infty$.
- **transient** if $\mathrm{P}(T_i < \infty) < 1$

We say a state is **ergodic** if it is aperiodic and positive recurrent.

The Fundamental Limit Theorem of Markov Chains I

Consider a recurrent irreducible aperiodic Markov chain. Then

$$\lim_{n\to\infty} P_{ij}^{(n)} = \frac{1}{\mathbb{E}[T_j]}$$

Moreover, if a Markov chain is irreducible and ergodic,

$$\pi_j = \lim_{n\to\infty} P_{ij}^{(n)} = \frac{1}{\mathbb{E}[T_j]}$$

is uniquely determined by the set of equations

$$\pi_j \geq 0, \quad \sum\nolimits_{j\in\mathfrak{X}} \pi_j = 1, \quad \pi_j = \sum\nolimits_{i\in\mathfrak{X}} \pi_i P_{ij}$$

Proof. See Theorem 1.1, 1.2, 1.3 on p.81-86 in Karlin & Taylor (1975).

Why $\pi_i = 1/\mathbb{E}(T_i)$?

Consider a Markov chain started from state j. Let S_k be the time till the k-th visit to state i. Then

$$S_k = T_{ji} + T_{ii}(1) + \ldots + T_{ii}(k-1)$$

Here

- T_{ji} = the first time the process visits state i from state j, and
- $T_{ii}(m)$ = the time between the mth and $(m+1)$st visit to state i.

Observe that $T_{ii}(1)$, $T_{ii}(2), \ldots T_{ii}(k-1)$ are i.i.d. and have the same distribution as T_i.

For k large, the Law of Large Numbers tells us

$$\frac{1}{k}[T_{ji} + T_{ii}(1) + T_{ii}(2) + \cdots + T_{ii}(k-1)] \approx \mathbb{E}(T_i)$$

i.e., the chain visits state i about k times in $k\mathbb{E}(T_i)$ steps.

We have just seen that in n steps, we expect about $n\pi(i)$ visits to the state i. Hence setting $n = k\mathbb{E}(T_i)$, we get the relation

$$\pi_i = 1/\mathbb{E}(T_i).$$

Remark

From the result in the previous page, we can see that a state i is **null recurrent**, i.e., $\mathbb{E}(T_i) = \infty$, if and only if

$$\lim_{n\to\infty} P_{ji}^{(n)} = 0, \quad \text{for all } j \in \mathfrak{X}.$$

Proposition 4.5 Positive Recurrence is a Class Property

- From the Fundamental Limit Theorem of Markov Chains I

$$\pi_i = 1/\mathbb{E}[T_i]$$

  and that a state i is positive recurrent if and only if $\mathbb{E}[T_i] < \infty$ it follows that a state i is positive current if and only if $\pi_i > 0$
- If a state j communicate with a positive recurrent state i, then state j is also positive recurrent.

  Proof. Since $i \leftrightarrow j$, there exists n such that $P_{ij}^{(n)} > 0$. Along with the fact that i is positive recurrent, $\pi_i > 0$, we know $\pi_j = \sum_k \pi_k P_{kj}^{(n)} \geq \pi_i P_{ij}^{(n)} > 0$. So j is also positive recurrent.

Corollary: Null Recurrence is a Class Property

If state i is null recurrent and $i \leftrightarrow j$, then state j is also null recurrent.

Proof. Since recurrence is a class property, state j can only be positive or null recurrent as it communicates with a null recurrent state i. Suppose state j is positive recurrent. As positive recurrence is a class property, state i must also be positive recurrent not null recurrent if it communicates with state j. So state j can only be null recurrent.

Finite-State Markov Chains Have No Null Recurrent States

In a finite-state Markov chain all recurrent states are positive recurrent.

Proof.

It suffices to consider irreducible Markov chains only since a Markov chain restricted to one of its recurrent class is also a Markov chain.
Recall an irreducible Markov chain must be recurrent. Also recall that positive/null recurrence is a class property. Thus if one state is null recurrent, then all states are null recurrent. However, since $\sum_{j\in\mathfrak{X}} P_{ij}^{(n)} = 1$. As there are only finite number of states, it is impossible that $\lim_{n\to\infty} P_{ij}^{(n)} = 0$ for all $j \in \mathfrak{X}$. Thus no state can be null recurrent.

Remark. For a finite state Markov chain, a limiting distribution exists if it is irreducible and aperiodic

The Fundamental Limit Theorem of Markov Chain II ($\star\star\star\star\star$)

If a Markov chain is **irreducible**, then the Markov chain is **positive recurrent** if and only if there exists a solution to the set of equations:

$$\pi_i \geq 0, \quad \sum_{i\in\mathfrak{X}} \pi_i = 1, \quad \pi_j = \sum_{i\in\mathfrak{X}} \pi_i P_{ij}$$

If a solution exists then

- it will be unique, and

$$\pi_j = \begin{cases} \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P_{ij}^{(k)} & \text{if the chain is periodic} \\ \lim_{n\to\infty} P_{ij}^{(n)} & \text{if the chain is aperiodic} \end{cases}$$

Remark. When a Markov chain is periodic, though its limiting distribution $\lim_{n\to\infty} P_{ij}^{(n)}$ doesn't exist, another limit $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} P_{ij}^{(k)}$ exists and is equal to the stationary distribution. The later limit can be interpret as the **long run proportion of time that the Markov chain is in state j.**

Example 1: One-Dimensional Random Walk

In Lecture 4, we have shown that 1-dim symmetric random walk has no stationary distribution.

- Conclusion: 1-dim symmetric random walk is null recurrent, i.e.

$$\mathbb{E}[T_i] = \infty \quad \text{for all state } i$$

In fact, in Lecture 3 we have shown that

$$P_{ii}^{(n)} = \begin{cases} 0 & \text{if } n \text{ is odd} \\ \binom{n}{n/2}\left(\frac{1}{2}\right)^n \approx \sqrt{\frac{2}{\pi n}} & \text{if } n \text{ is even} \end{cases}$$

Thus $\pi_i = \lim_{n\to\infty} P_{ii}^{(n)} = 0$, and hence $\mathbb{E}[T_i] = 1/\pi_i = \infty$.

Ex 2: 1-D Random Walk w/ Partially Reflective Boundary

$$
\begin{aligned}
P_{i,i+1} &= p \quad \text{for all } i = 0, 1, 2, \ldots \\
P_{i,i-1} &= 1 - p \quad \text{for all } i = 1, 2, \ldots \\
p_{00} &= 1 - p
\end{aligned}
$$

Try to solve $\pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij}$

$$
\begin{aligned}
\pi_0 &= \pi_0 P_{00} + \pi_1 P_{10} = (1-p)(\pi_0 + \pi_1) \Rightarrow \pi_1 = \frac{p}{1-p}\pi_0 \\
\pi_1 &= \pi_0 P_{01} + \pi_2 P_{21} = p\pi_0 + (1-p)\pi_2 \Rightarrow \pi_2 = \left(\frac{p}{1-p}\right)^2 \pi_0 \\
\pi_2 &= \pi_0 P_{12} + \pi_3 P_{32} = p\pi_1 + (1-p)\pi_3 \Rightarrow \pi_3 = \left(\frac{p}{1-p}\right)^3 \pi_0 \\
&\vdots \\
\pi_j &= p\pi_{j-1} + (1-p)\pi_{j+1} \qquad \Rightarrow \pi_{j+1} = \left(\frac{p}{1-p}\right)^{j+1} \pi_0
\end{aligned}
$$

Ex 2: 1-D Random Walk w/ Partially Reflective Boundary

$$\sum_{i=0}^{\infty} \pi_i = \pi_0 \sum_{i=0}^{\infty} \left(\frac{p}{1-p}\right)^i = \begin{cases} \pi_0 \left(\frac{1-p}{1-2p}\right) & \text{if } p < 1/2 \\ \infty & \text{if } p \geq 1/2 \end{cases}$$

Conclusion: The process is positive recurrent iff $p < 1/2$, in which case

$$\pi_i = \frac{1-2p}{1-p}\left(\frac{p}{1-p}\right)^i, \quad i = 0, 1, 2, \ldots$$

Ex 3: Ehrenfest Diffusion Model with N Balls

Recall that in Lecture 4, we show that Ehrenfest Diffusion Model is irreducible, has period = 2, and there exists a solution to the set of equations

$$\pi_i \geq 0, \quad \sum_{i \in \mathfrak{X}} \pi_i = 1, \quad \pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij}$$

which is

$$\pi_i = \binom{N}{i} \left(\frac{1}{2}\right)^N \quad \text{for } i = 0, 1, 2, \ldots, N$$

Though the limiting distribution $\lim_{n\to\infty} P_{ij}^{(n)}$ does not exist, we can show that

$$\lim_{n\to\infty} P_{ij}^{(2n)} = 2\binom{N}{j}\left(\frac{1}{2}\right)^N, \quad \lim_{n\to\infty} P_{ij}^{(2n+1)} = 0 \quad \text{if } i+j \text{ is even}$$

$$\lim_{n\to\infty} P_{ij}^{(2n)} = 0, \quad \lim_{n\to\infty} P_{ij}^{(2n+1)} = 2\binom{N}{j}\left(\frac{1}{2}\right)^N \quad \text{if } i+j \text{ is odd}$$

From the above, one can verify that
$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} P_{ij}^{(k)} = \binom{N}{j}\left(\frac{1}{2}\right)^N = \pi_j$.

Exercise 4.50 on p.284

A Markov chain has transition probability matrix

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 0.2 & 0.4 & 0 & 0.3 & 0 & 0.1 \\ 0.1 & 0.3 & 0 & 0.4 & 0 & 0.2 \\ 0 & 0 & 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.2 & 0.8 \end{pmatrix} \end{array}$$

Communicating classes:

$$\underset{\text{transient}}{\underset{\uparrow}{\{1,2\}}} \qquad \underset{\text{recurrent}}{\underset{\uparrow}{\{3,4\}}} \qquad \underset{\text{recurrent}}{\underset{\uparrow}{\{5,6\}}}$$

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

Exercise 4.50 on p.284 (Cont'd)

Observe that $\lim_{n\to\infty} P_{ij}^{(n)} = 0$ if j is transient, hence,

$$\lim_{n\to\infty} P^{(n)} = \begin{array}{c c} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{pmatrix} 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \end{pmatrix} \end{array}$$

Exercise 4.50 on p.284 (Cont'd)

Observe that $\lim_{n\to\infty} P_{ij}^{(n)} = 0$ if j is NOT accessible from i

$$\lim_{n\to\infty} P^{(n)} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & 0 & 0 \\ 0 & 0 & ? & ? & 0 & 0 \\ 0 & 0 & 0 & 0 & ? & ? \\ 0 & 0 & 0 & 0 & ? & ? \end{pmatrix} \end{array}$$

The two classes $\{3,4\}$ and $\{5,6\}$ do not communicate and hence the transition probabilities in between are all 0.

Exercise 4.50 on p.284 (Cont'd)

Recall we have shown that the limiting distribution of a two-state Markov chain with the transition matrix $\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$ is $\left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$. As the Markov chain restricted to the closed class $\{3,4\}$ is also a Markov chain with the transition matrix

$$\begin{array}{cc} & \begin{array}{cc} 3 & 4 \end{array} \\ \begin{array}{c} 3 \\ 4 \end{array} & \begin{pmatrix} 0.3 & 0.7 \\ 0.6 & 0.4 \end{pmatrix} \end{array}.$$

Hence,

$$\lim_{n\to\infty} P^{(n)} = \begin{array}{cc} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{pmatrix} 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & 6/13 & 7/13 & 0 & 0 \\ 0 & 0 & 6/13 & 7/13 & 0 & 0 \\ 0 & 0 & 0 & 0 & ? & ? \\ 0 & 0 & 0 & 0 & ? & ? \end{pmatrix} \end{array}$$

Exercise 4.50 on p.284 (Cont'd)

$$P = \begin{array}{c c} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{pmatrix} 0.2 & 0.4 & 0 & 0.3 & 0 & 0.1 \\ 0.1 & 0.3 & 0 & 0.4 & 0 & 0.2 \\ 0 & 0 & 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.2 & 0.8 \end{pmatrix} \end{array}$$

For the same reason,

$$\lim_{n\to\infty} P^{(n)} = \begin{array}{c c} & \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{pmatrix} 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & 6/13 & 7/13 & 0 & 0 \\ 0 & 0 & 6/13 & 7/13 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2/7 & 5/7 \\ 0 & 0 & 0 & 0 & 2/7 & 5/7 \end{pmatrix} \end{array}$$