40.520 Stochastic Models

Discrete Time Markov Chains

Xiaotang Yang

Engineering Systems and Design (ESD)
Singapore University of Technology and Design (SUTD)

Spring 2026

Outline

Introduction

Stochastic Process

A stochastic process is a collection of random variables that are indexed by time. Usually, we denote it by

- $(X_1, X_2, \dots, X_n, \dots)$ for a discrete-time random process
- $(X_t, t \geq 0)$ for a continuous-time random process

Why Study Stochastic Processes?

- Characterize temporal relationships between random variables
- Simplest model: X_n are independent
- But independence may not capture real-world complexity
- **Markov chains** — processes where the future depends **only** on the present.

What is a Markov Chain?

Let $(X_n, n = 0, 1, 2, \ldots)$ be a (discrete-time) stochastic process with finite/countable states.

It is a **Markov chain** if:

$$P(X_{n+1} = j \mid X_1 = i_1, \ldots, X_n = i_n) = P(X_{n+1} = j \mid X_n = i_n)$$

for all n.

Markov Property

Future depends only on the present, not the past.

The Markovian Property (Formally)

Equivalent definition:

$$
\begin{aligned}
&P(X_{n+1} = j, X_1 = i_1, \ldots, X_{n-1} = i_{n-1} \mid X_n = i_n) \\
&= P(X_{n+1} = j \mid X_n = i_n) \cdot P(X_1 = i_1, \ldots, X_{n-1} = i_{n-1} \mid X_n = i_n)
\end{aligned}
$$

Interpretation: Given the current state, future and past are independent.

Transition Matrix

Consider **time-homogeneous** Markov chains:

$$P(X_{n+1} = j \mid X_n = i) \text{ is independent of } n.$$

Define:

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

Transition Matrix

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

Properties of Transition Matrix

Stochastic Matrix

Each row sums to 1:

$$\sum_{j=1}^{m} p_{ij} = 1 \quad \text{for all } i.$$

- $p_{ij} \geq 0$ for all i, j
- Represents one-step transition probabilities

Example: Weather Forecasting

- Chance of rain tomorrow depends **only** on today's weather
- Assumptions:
  - If rain today $\Rightarrow$ rain tomorrow with probability 70%
  - If no rain today $\Rightarrow$ rain tomorrow with probability 50%

Characterize the Transation matrix for whether.

Example: 1-D Random Walk

- State space: integers $0, \pm 1, \pm 2, \ldots$
- For some $0 < p < 1$:

  $$P_{i,i+1} = p, \quad P_{i,i-1} = 1 - p$$

  for all i

Symmetric vs. Asymmetric

- **Symmetric**: $p = \frac{1}{2}$
- **Asymmetric**: $p \neq \frac{1}{2}$

Classic model for particle movement, gambling, stock prices, etc.

Chapman-Kolmogorov Equations

The Chapman-Kolmogorov Equations

CK Equations

For $n, m \geq 0$,

$$P_{ij}^{n+m} = P(X_{n+m} = j \mid X_0 = i) = \sum_k P_{ik}^n P_{kj}^m$$

Matrix Form

$$P^{n+m} = P^n \cdot P^m$$

where P^n is the n-step transition matrix.

Meaning: To go from i to j in $n + m$ steps, pass through some intermediate state k at step n.

Example: Rain Probability (Revisited)

- States: $S_1 = \text{Rain}, \quad S_2 = \text{No Rain}$
- Transition matrix:
$$P = \begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix}$$
- Given: Raining today (state 1)
- Find: Probability of rain 4 days from today.

Example: Urn and Balls

- Urn always contains 2 balls (red/blue)
- Each step: randomly choose a ball, replace with:
  - Same color: probability 80%
  - Opposite color: probability 20%
- Initially: both balls red
- Find: Probability 5th ball selected is red

Example: Fair Coin

- Independent flips of fair coin
- Let N = number of flips until 3 consecutive heads
- Find:

$$P(N \leq 8) \quad \text{and} \quad P(N = 8)$$

The Gambler's Ruin Problem

Problem Statement

- Gambler starts with i dollars
- Each play:
  - Win \$1 with probability p
  - Lose \$1 with probability $q = 1 - p$
- Plays are independent
- Goal: Reach N dollars before going broke (0 dollars)

What is the probability P_i of reaching N before 0?

Markov Description of the Model

- If we let X_n denote the player's fortune at time n, then the process $\{X_n, n = 1, 2, \ldots\}$ is a Markov chain with transition probabilities

$$P_{00} = P_{NN} = 1$$

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \quad i = 1, 2, \ldots, N-1$$

Solution

- Let P_i be the probability that, starting with i dollars, the gambler fortune will eventually reach N.
- By conditioning on the outcome of the initial play

$$P_i = pP_{i+1} + qP_{i-1}, \quad i = 1, 2, \ldots, N-1.$$

and $P_0 = 0, P_N = 1$.

Solution (Continued)

- Hence, we obtain from the preceding slide that

$$
\begin{aligned}
P_2 - P_1 &= \frac{q}{p}(P_1 - P_0) = \frac{q}{p} P_1 \\
P_3 - P_2 &= \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1 \\
&\cdots \\
P_N - P_{N-1} &= \frac{q}{p}(P_{N-1} - P_{N-2}) = \left(\frac{q}{p}\right)^{N-1} P_1
\end{aligned}
$$

Solution (Continued)

- Adding all the equalities up, we obtain

$$P_1 = \begin{cases} 1/N & p = 1/2 \\ \frac{1-(q/p)}{1-(q/p)^N} & p \neq 1/2 \end{cases}$$

$$P_i = \begin{cases} i/N & p = 1/2 \\ \frac{1-(q/p)^i}{1-(q/p)^N} & p \neq 1/2 \end{cases}$$

Solution (Continued)

Note that, as $N \to \infty$

$$P_i \to \begin{cases} 0, & p \leq 1/2; \\ 1 - \left(\frac{q}{p}\right)^i & p > 1/2. \end{cases}$$

- If $p > 1/2$, there is a positive probability that the gambler's fortune will increase indefinitely
- If $p \leq 1/2$, the gambler will, with probability 1, go ruin against an infinitely rich adversary (say, a casino).

Hitting Times

Probability of Reaching (Hitting) A Set?

Let $(X_n)_{n\geq 0}$ be a Markov chain with state space S. The hitting time of a subset A of S is the random variable $H^A : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ given by

$$H^A(\omega) = \inf\{n \geq 0 : X_n(\omega) \in A\}$$

The probability starting from i that $(X_n)_{n>0}$ ever hits A is then

$$h_i^A = P_i(H^A < \infty).$$

When A is a closed class, h_i^A is called the **absorption probability**.

A General Result for Probabilities of Hitting A Set

Theorem
The vector of hitting probabilities $h^A = (h_i^A : i \in S)$ is the **minimal non-negative solution** to the system of linear equations

$$\begin{cases} h_i^A = 1 & \text{for } i \in A \\ h_i^A = \sum_{j \in S} p_{ij} h_j^A & \text{for } i \notin A. \end{cases}$$

How long does it take to hit a set?

The mean time taken for $(X_n)_{n\geq 0}$ to reach A is given by

$$k_i^A = E_i(H^A) = \sum_{n<\infty} n\, P_i(H^A = n) + \infty \cdot P_i(H^A = \infty).$$

Remark: The concept of expectation breaks down at infinity.

A General Result for Expected Time to Hit A Set

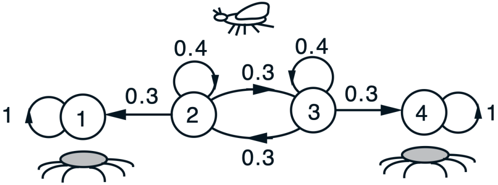
Theorem

The vector of mean hitting times $k^A = (k_i^A : i \in S)$ is the minimal non-negative solution to the system of linear equations

$$\begin{cases} k_i^A = 0 & \text{for } i \in A \\ k_i^A = 1 + \sum_{j \notin A} p_{ij} k_j^A & \text{for } i \notin A. \end{cases}$$

Example

Consider the chain with the following diagram:



	1	2	3	4
1	1.0	0	0	0
2	0.3	0.4	0.3	0
3	0	0.3	0.4	0.3
4	0	0	0	1.0

p_{ij}

Suppose that at time 0, the fly is at location i.

- What is the probability that the fly gets caught?
- How long does the fly live on an average before getting caught?

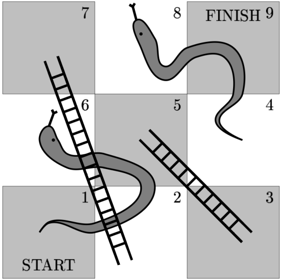
Example: Snakes and Ladders

A simple game of 'snakes and ladders' is played on a board of nine squares. At each turn a player tosses a fair coin and advances one or two places according to whether the coin lands heads or tails.

- If you land at the foot of a ladder you climb to the top.
- If you land at the head of a snake you slide down to the tail.

How many turns on average does it take to complete the game?

What is the probability that a player who has reached the middle square will complete the game without slipping back to square 1?



Stopping Times and Strong Markov Property

A Question

Let $(X_n)_{n\geq 0}$ be a Markov chain with transition matrix P. Instead of starting at time $n = 0$, imagine you start observing at a **random time** T that depends on the chain's behavior:

- Start when the chain first enters a certain state.
- Start after it visits a state for the second time.

Define:

$$Y_n = X_{T+n}, \quad n \geq 0$$

(The process you see *after* time T.)

Is $(Y_n)_{n\geq 0}$ still a Markov chain with the same transition probabilities P?

Stopping Times

A random variable $T : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$ is called a **stopping time** if the event $\{T = n\}$ depends only on $X_0, X_1, \dots, X_n$ for $n = 0, 1, 2, \dots$.

Example

Are the following stopping times?

First passage time to state j:

$$T_j = \inf\{n \geq 1 : X_n = j\},$$

$$\{T_j = n\} = \{X_1 \neq j, \ldots, X_{n-1} \neq j, X_n = j\}$$

Hitting time to set A:

$$H_A = \inf\{n \geq 0 : X_n \in A\},$$

$$\{H_A = n\} = \{X_0 \notin A, \ldots, X_{n-1} \notin A, X_n \in A\}$$

Last exit time from set A:

$$L^A = \sup\{n \geq 0 : X_n \in A\}$$

Strong Markov Property

Let $(X_n)_{n \geq 0}$ be a Markov chain with transition matrix P. Then, conditional on $T < \infty$ and $X_T = i$, $(X_{T+n})_{n \geq 0}$ is a Markov chain with the same transition matrix P and independent of $X_0, X_1, \ldots, X_{T-1}$.

Example

Consider a simple symmetric random walk (S_n) starting at 0, and let T be the first hitting time of level $a > 0$ (i.e., $T = \min\{n \geq 0 : S_n = a\}$).

The strong Markov property implies that the process $\{S_{T+n} - a\}_{n \geq 0}$, after time T, is **again a simple symmetric random walk starting at 0**, and is **independent** of the path up to time T.

This allows us to treat the walk **after T as a fresh restart**, a key step in problems like computing ruin probabilities or using the reflection principle.

State Classification

Asymptotic Behavior of Markov Chains

- The asymptotic behavior of p_{ij}^n as $n \to \infty$.
- If the influence of the initial state recedes in time, then as $n \to \infty$, p_{ij}^n approaches a limit which is independent of i?
- Principles of classifying states of a Markov chain.

Example

Consider a Markov chain consisting of the 4 states 0, 1, 2, 3 and having the following transition probability matrix

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

What is the most improbable state after 1,000 steps by your intuition?

Accessibility and Communication

- State j is said to be **accessible** from state i if for some n, $P_{ij}^n > 0$.
- In the previous slide, state 3 is accessible from state 2.
- But state 2 is *not* accessible from state 3.
- Two states i and j are said to **communicate** if they are accessible to each other. We write $i \leftrightarrow j$.
- Which pairs of states communicate in the previous example?

Simple Properties of Communication

- The relation of communication satisfies the following three properties:
    1. State i communicates with itself;
    2. If state i communicates with state j, then state j communicates with state i;
    3. If state i communicates with state j, and state j communicates with state k, then state i communicates with state k.

State Classes

- Two states that communicate are said to be in the same **class**.
- Any two classes are either identical or disjoint.
- The concept of communication divides the state space into a number of separate classes.
- In the previous example, how many classes do we have?

Example: Irreducible Markov Chain

Consider the Markov chain consisting of the three states 0, 1, 2, and having transition probability matrix

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

How many classes does it contain?

Definition

The Markov chain is said to be **irreducible** if there is only one class.

Transience and Recurrence

Recurrence and Transience

- Let f_i represent the probability that, starting from state i, the process will ever return to state i.
- We say a state i is **recurrent** if $f_i = 1$.
- If a state is recurrent, starting from this state, the Markov chain will return to it again, and again, and again — in fact, infinitely often.

Recurrence and Transience (Continued)

A non-recurrent state is said to be **transient**, i.e., a transient state i satisfies $f_i < 1$.

Starting from a transient state i,

- The process will never again revisit the state with a positive probability $1 - f_i$;
- The process will revisit the state just once with a probability $f_i(1 - f_i)$;
- The process will revisit the state just twice with a probability $f_i^2(1 - f_i)$;
- ...

Recurrence and Transience (Continued)

From the above two definitions, we can easily see the following conclusions:

- The number of times that the process will visit a transient state has a geometric distribution.
- A transient state will only be visited a finite number of times.
- In a finite-state Markov chain not all states can be transient.

One Commonly Used Criterion of Recurrence

Theorem

A state i is recurrent if and only if

$$\sum_{n=1}^{+\infty} P_{ii}^{n} = \infty.$$

Theorem and Implications

Theorem

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

Two conclusions can be drawn from the theorem:

1. Transience is also a class property.
2. All states of a finite irreducible Markov chain are recurrent.

Example

Let the Markov chain consist of the states 0, 1, 2, 3, and have transition probability matrix

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

Determine which states are transient and which are recurrent.

Example

- Discuss the recurrent property of a one-dimensional random walk.
- Conclusion:
  - Symmetric random walk is recurrent;
  - Asymmetric random walk is not.

Long-Run Proportions and Limiting Probabilities

Long-Run Proportions of MC

- Consider a discrete-time Markov chain $(X_1, X_2, \ldots, X_n, \ldots)$.
- Let π_j denote the long-run proportion of time that the Markov chain is in state j, i.e.,

$$\pi_j = \lim_{n \to \infty} \frac{\#\{1 \leq i \leq n : X_i = j\}}{n}$$

Long-Run Proportions of MC (Continued)

- A simple fact is that, if a state j is transient, the corresponding $\pi_j = 0$.
- Therefore we only consider recurrent states in this subsection:
- Let $N_j = \min\{k > 0 : X_k = j\}$, the number of transitions until the Markov chain makes a transition into the state j. Denote m_j to be its expectation, i.e. $m_j = E[N_j | X_0 = j]$.

Long-Run Proportions of MC (Continued)

Theorem

If the Markov chain is irreducible and recurrent, then for any initial state

$$\pi_j = \frac{1}{m_j}.$$

Positive Recurrent and Null Recurrent

- Definition: we say a state is positive recurrent if $m_j < +\infty$; and say that it is null recurrent if $m_j = +\infty$.
- It is obvious from the previous theorem, if state j is positive recurrent, we have $\pi_j > 0$
- If state i is positive recurrent and i communicates with j, then state j is also positive recurrent!

How to Determine π_j?

Theorem

Consider an irreducible Markov chain. If the chain is also positive recurrent, then the long-run proportions of each state are the unique solution of the equations:

$$\pi_j = \sum_i \pi_i p_{ij}, \forall j \qquad \sum_j \pi_j = 1.$$

If there is no solution of the preceding linear equations, then the chain is either transient or null recurrent and all $\pi_j = 0$.

Example: Rainy Days in Singapore

- Assume that in Singapore, if it rains today, then it will rain tomorrow with prob. 60%; and if it does not rain today, then it will rain tomorrow with prob. 40%. What is the average proportion of rainy days in Singapore?

Example: A Model of Class Mobility

- A problem of interest to sociologists is to determine the proportion of a society that has an upper-, middle-, and lower-class occupations.
- Let us consider the transitions between social classes of the successive generations in a family. Assume that the occupation of a child depends only on his or her parent's occupation.

Example: A Model of Class Mobility (Continued)

- The transition matrix of this social mobility is given by

$$P = \begin{pmatrix} 0.45 & 0.48 & 0.07 \\ 0.05 & 0.70 & 0.25 \\ 0.01 & 0.50 & 0.49 \end{pmatrix}$$

- That is, for instance, the child of a middle-class worker will attain an upper-class occupation with prob. 5%, will move down to a lower-class occupation with prob. 25%.
- What is the long-run proportion for each class?

Stationary Distribution of MC

For a Markov chain, any set of $\{\pi_i\}$ satisfying $\pi_j = \sum \pi_i p_{ij}, \forall j$, and $\sum \pi_j = 1$, is called a stationary probability distribution of the Markov chain.

Theorem

If the Markov chain starts with an initial distribution $\{\pi_i\}$ i.e., $P(X_0 = i) = \pi_i$, then $P(X_t = i) = \pi_i$ for all state i and $t \geq 0$.

Limiting Probabilities

- Reconsider the example of rainy days in Singapore. Suppose that the transition probability matrix is given by
$$P = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}$$
- Please calculate what are the probabilities that it will rain in 10 days, in 100 days, and in 1,000 days, given it does not rain today.

Limiting Probabilities (Continued)

With the help of MATLAB, we have

$$P^{10} = \begin{pmatrix} 0.5715 & 0.4285 \\ 0.5714 & 0.4286 \end{pmatrix}$$

$$P^{100} = \begin{pmatrix} 0.5714 & 0.4286 \\ 0.5714 & 0.4286 \end{pmatrix}$$

$$P^{1000} = \begin{pmatrix} 0.5714 & 0.4286 \\ 0.5714 & 0.4286 \end{pmatrix}$$

Limiting Probabilities (Continued)

Observations:

- As $n \to \infty$, P_{ij}^n converges;
- The limit of P_{ij}^n does not depend on the initial state i.
- The limits coincide with the stationary distribution of the Markov chain

$$\begin{aligned}
\pi_0 &= 0.7\pi_0 + 0.4\pi_1; \\
\pi_1 &= 0.3\pi_0 + 0.6\pi_1; \\
\pi_0 &= \frac{4}{7} \approx 0.5714; \\
\pi_1 &= \frac{3}{7} \approx 0.4286.
\end{aligned}$$

Limiting Probabilities and Stationary Distribution

Theorem

In a positive recurrent (aperiodic) Markov chain, we have

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

where $\{\pi_j\}$ is the stationary distribution of the chain.

A positive recurrent Markov chain will reach an equilibrium/a steady state after long-term transition.

Periodic vs. Aperiodic

- The requirement of aperiodicity in the last theorem turns out to be very essential.
- We say that a Markov chain is periodic if, starting from any state, it can only return to the state in a multiple of $d > 1$ steps. Otherwise, we say that it is aperiodic.
- Example: A Markov chain with period 3.

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

Summary

- In a positive recurrent aperiodic Markov chain, the following three concepts are equivalent:
    1. Long-run proportion;
    2. Stationary probability distribution;
    3. Long-term limits of transition probabilities.

The Ergodic Theorem

Theorem

Let $\{X_n, n \geq 1\}$ be an irreducible Markov chain with stationary distribution $\{\pi_i\}$ and let $r(\cdot)$ be a bounded function on the state space. Then,

$$\lim_{n \to +\infty} \frac{\sum_{i=1}^{n} r(X_i)}{n} = \sum_j r(j)\pi_j.$$

Example: Bonus-Malus Automobile Insurance System

- In most countries of Europe and Asia, automobile insurance premium is determined by use of a Bonus-Malus (Latin for Good-Bad) system.
- Each policyholder is assigned a positive integer valued state and the annual premium is a function of this state.
- Lower numbered states correspond to lower number of claims in the past, and result in lower annual premiums.

Example (Continued)

- A policyholder's state changes from year to year in response to the number of claims that he/she made in the past year.
- Consider a hypothetical Bonus-Malus system having 4 states.

State	Annual Premium ($)
1	200
2	250
3	400
4	600

Example (Continued)

- According to historical data, suppose that the transition probabilities between states are given by

$$P = \begin{pmatrix} 0.6065 & 0.3033 & 0.0758 & 0.0144 \\ 0.6065 & 0 & 0.3033 & 0.0902 \\ 0 & 0.6065 & 0 & 0.3935 \\ 0 & 1 & 0.6065 & 0.3935 \end{pmatrix}$$

- Find the average annual premium paid by a typical policyholder.