Week #6: Markov processes

March 10, 2025

1 Markov Chains

One of the most fundamental and widely studied stochastic processes is the Markov chain, named after the Russian mathematician Andrey Markov.

Definition. Markov Chain

A sequence of random variables $\{X_0, X_1, X_2, \ldots\}$ is said to be a **Markov chain** if, for any $t \geq 0$ and any states $x_0, x_1, \ldots, x_t$, the probability of the next state x_{t+1} depends only on the current state x_t and not on the sequence of states preceding it. Mathematically, this property can be expressed as:

$$P(X_{t+1} = x_{t+1} | X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) = P(X_{t+1} = x_{t+1} | X_t = x_t)$$

This property is often referred to as the **Markov property** or **memorylessness**. In simpler terms, the future state of a Markov chain depends only on its current state and not on its past states.

Example. Random Walk

A **Random walk** is a classic example of a stochastic process and is often used to describe systems or sequences of events where the next state depends only on the current state and some random element. It can be visualized as a path taken by a particle that moves in random directions.

The simplest random walk is the one-dimensional walk. At each step, the walker takes a step either to the right (+1) or to the left (-1) with equal probability.

1. Start at position 0.
2. At each time step, flip a coin:
   - If heads, move +1 step to the right.
   - If tails, move -1 step to the left.
3. Record the position after each step.
4. Repeat for a desired number of steps.

The particle starts at position 0. At each step, it moves either one step to the right (+1) or one step to the left (-1) with equal probability. The path appears quite random, and the particle

Figure 1: 1D Random Walk over 1000 Steps

can drift far from the starting position but can also return close to its starting position at various times.

A 2D random walk can be visualized on a grid or plane. At each step, the walker takes a step either up, down, left, or right with equal probability.

1. Start at position $(0, 0)$.
2. At each time step, randomly choose one of the four directions:
   - Up: $(+0, +1)$
   - Down: $(+0, -1)$
   - Left: $(-1, +0)$
   - Right: $(+1, +0)$
3. Record the position after each step.
4. Repeat for a desired number of steps.

In the 2D random walk, the particle moves in a plane, starting from the origin. It takes steps in one of four possible directions: up, down, left, or right. The resulting path is a series of connected line segments in the plane, illustrating the random journey of the particle over time. The path is visualized using the 'Greys' color map, where the starting point is black and the ending point is white, representing the time progression.

Definition.

For a Markov chain with N possible states, the $N \times N$ matrix $P = [p_{ij}]$ where $p_{ij} = P(X_{t+1} = j | X_t = i)$ is called the **Transition matrix** .

Note that the rows of a transition matrix for a Markov chain must each sum to 1.

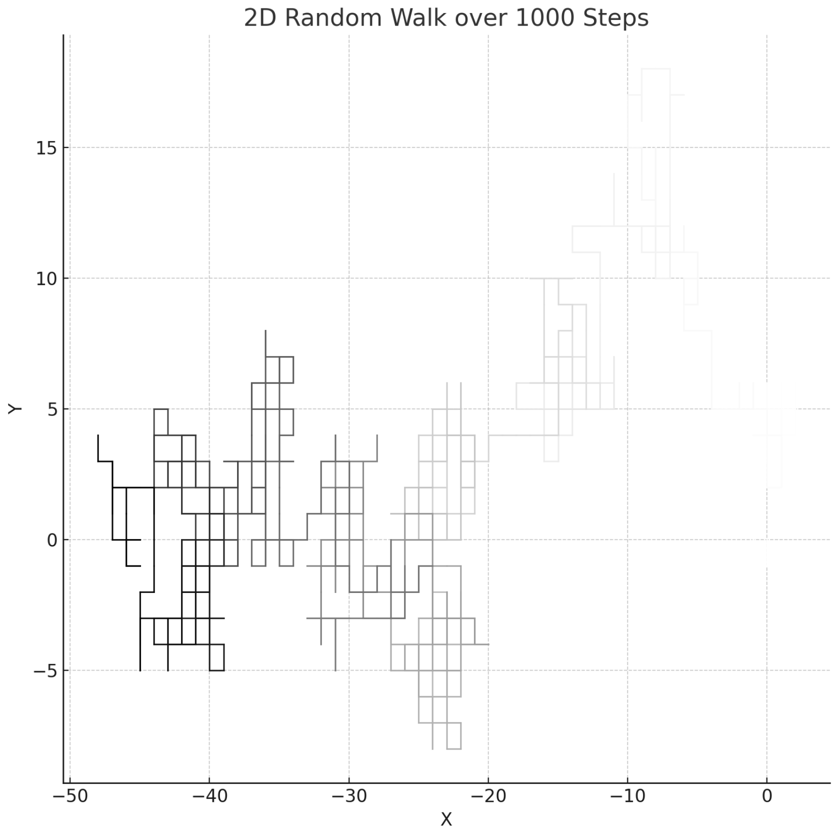

Figure 2: 2D Random Walk over 1000 Steps in Greys

1. **Transition Probabilities**: The probabilities of moving from one state to another are called transition probabilities. They are typically represented in a matrix called the transition matrix.

2. **State Space**: The set of all possible states that the chain can be in. This could be a finite or countably infinite set. In this context, we will focus on a total number of states equal to N.

3. **Time Structure**: Transitions occur at integer time steps. Specifically, you can visualize them as the subscripts of the random variables: the generic time point is the t-th.

Example. Financial Market Dynamics

Financial markets can be modeled as stochastic systems that evolve over time. In this example, we consider a market that can be in one of three states: Bullish, Bearish, and Neutral.

Figure 3 illustrates a corresponding state diagram. The one-step transition probabilities can be summarized in the transition matrix:

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

Example. Infection Dynamics

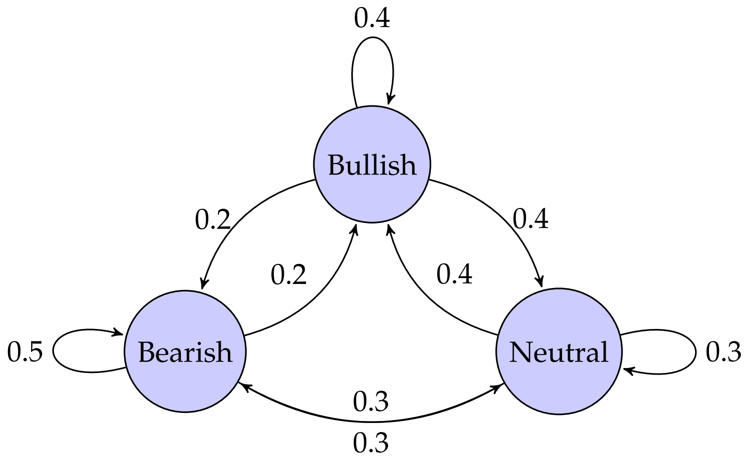

Figure 3: Market State Transition Diagram: Circles represent the states of the system, and the arrows symbolize the potential transitions between these states. The numbers adjacent to the arrows indicate the respective probabilities for these transitions. Given two states i and j, the probability of transition from state i to state j is equal to p_{ij}.

Consider a population that can be in one of three states: Susceptible (S), Infected (I), and Recovered (R). Individuals transition between these states according to the following Markov chain.

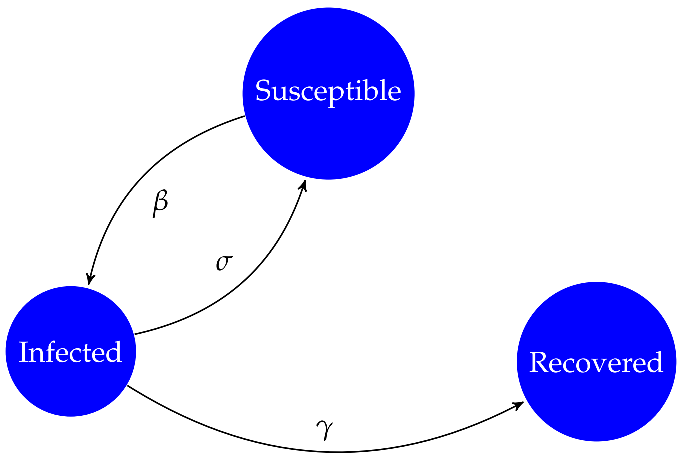

Figure 4: Example of a Markov Chain with three states and transition probabilities.

The transition probabilities are defined as: β, the probability of a Susceptible individual becoming Infected; γ, the probability of an Infected individual recovering and moving to the Recovered state; σ, the probability of an Infected individual transitioning back to the Susceptible state without recovering.

Given these probabilities, the transition matrix P for this Markov chain can be constructed as:

$$P = \begin{bmatrix} 1-\beta & \beta & 0 \\ \sigma & 1-\sigma-\gamma & \gamma \\ 0 & 0 & 1 \end{bmatrix}$$

Where the rows represent the current state and the columns represent the next state.

This Markov chain captures the essential dynamics of many infectious diseases. The future state of each individual depends only on their current state, satisfying the Markov property. For example, if a person is currently Infected, the probability that they will be Recovered in the next time step is γ, irrespective of their past states.

As a Markov chain evolves over time, we are often interested in predicting the likelihood of reaching a particular state after multiple transitions. Instead of considering only one-step transitions, we can analyze the probability of moving from state i to state j in n steps. This allows us to understand the long-term behavior of the chain and how probabilities propagate over time.

Let $P_{ij}^{(n)}$ denote the probability that the system transitions from state i to state j in n steps. Formally, we define:

$$P_{ij}^{(n)} = P(X_{t+n} = j | X_t = i) = (P^n)_{ij}.$$

Notes:

- $P_{ij}^{(n)}$ can be computed as the (i, j)th entry of the matrix P^n.
- The number of possible paths from i to j in n steps grows exponentially as N^{n-1}, where N is the number of states.
- The matrix power P^n accumulates all transition probabilities over n steps.

A fundamental result that formalizes the relationship between multi-step transitions is the *Chapman-Kolmogorov theorem*, which provides a recursive method to compute transition probabilities over arbitrary time intervals.

Theorem 1. Chapman-Kolmogorov

For any times $s < t < u$, the transition probabilities satisfy the equation:

$$P_{ij}(s, u) = \sum_{k=1}^{N} P_{ik}(s, t) \cdot P_{kj}(t, u). \tag{1}$$

This theorem expresses a key property of Markov chains: the probability of moving from state i at time s to state j at time u can be decomposed into an intermediate step at time t. Intuitively, it states that the probability of a long-term transition can be computed by summing over all possible intermediate states k, multiplying the probability of transitioning from i to k in the first interval $[s, t]$ by the probability of transitioning from k to j in the second interval $[t, u]$.

This result is particularly useful for computing transition probabilities iteratively, understanding how the chain propagates over time, and deriving fundamental properties such as steady-state behavior and mixing times.

Example. Financial Market Dynamics

As in the financial market dynamics example the one-step transition probabilities can be summarized in the transition matrix:

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

Now, suppose we wish to determine the probability that the market transitions from a Bullish state (state 1) to a Bearish state (state 2) in exactly five steps. This probability is given by the $(1, 2)$ entry of P^5, denoted by $(P^5)_{1,2}$. Using the Chapman–Kolmogorov equation, we can express:

$$(P^5)_{1,2} = \sum_{k=1}^{3} (P^2)_{1,k} \cdot (P^3)_{k,2}.$$

Here, P^2 and P^3 are the two- and three-step transition matrices, respectively, computed via standard matrix multiplication.

After performing these computations, we find:

$$(P^5)_{1,2} \approx 0.333.$$

Thus, under this model, there is roughly a 33% chance of transitioning from a Bullish market to a Bearish market over five steps.

This multi-step transition probability is invaluable in understanding longer-term market dynamics, aiding analysts in risk assessment and strategic planning.

One of the most pivotal inquiries in the realm of Markov Chains pertains to their long-term behavior. Specifically, will the system stabilize into a steady state or equilibrium?

Definition. Limit Distribution

A distribution π is called a **limit distribution** for a Markov Chain if

$$\lim_{n \to \infty} P_{ij}(0, n) = \pi_j$$

for every state i. The limit distribution provides insights into the enduring behavior of the system and is inherently linked to the properties discussed below.

The distribution π encapsulates the stable or steady-state probabilities associated with each state as the number of transitions grows indefinitely large. For a finite Markov Chain, the cumulative sum of all elements of π equals 1, emphasizing that π is a probability distribution.

Example. Convergence of P to π

For the given transition matrix P, let's inspect its powers:
For P^5:

$$\begin{bmatrix} 0.34296 & 0.35264 & 0.3044 \\ 0.34664 & 0.33984 & 0.31352 \\ 0.33752 & 0.3648 & 0.29768 \end{bmatrix}$$

For P^{10}:

$$\begin{bmatrix} 0.34260 & 0.35183 & 0.30557 \\ 0.34251 & 0.35210 & 0.30539 \\ 0.34268 & 0.35159 & 0.30573 \end{bmatrix}$$

For P^{20}:

$$\begin{bmatrix} 0.34259 & 0.35185 & 0.30556 \\ 0.34259 & 0.35185 & 0.30556 \\ 0.34259 & 0.35185 & 0.30556 \end{bmatrix}$$

By the time we examine P^{50} and P^{100}, the matrix has stabilized to:

$$\begin{bmatrix} 0.34259 & 0.35185 & 0.30556 \\ 0.34259 & 0.35185 & 0.30556 \\ 0.34259 & 0.35185 & 0.30556 \end{bmatrix}$$

From the matrices above, we discern a clear trend: as we raise P to higher powers, the rows of the matrix are converging to the limit distribution π. This showcases the theoretical underpinning that, given certain conditions, the Markov Chain will stabilize to a unique long-term distribution.

Definition.

A Markov Chain is **irreducible** if it is possible to traverse from any state to any other state within a finite number of steps. Formally, for any states $i, j \in \{1, ..., N\}$, there exists $n \geq 1$ such that $P_{ij}^{(n)} > 0$.

Irreducibility plays a paramount role in systems like social networks, ensuring the flow of information across the entire network.

Definition.

A state of a Markov Chain exhibits **aperiodicity** if it doesn't revisit itself in a fixed pattern. Formally, a state i is aperiodic if the greatest common divisor of the set of steps n at which it returns to itself is one: $\gcd\{n : P_{ii}^{(n)} > 0\} = 1$.

Aperiodicity is crucial in financial models to avoid deterministic cyclical behaviors, ensuring the model captures the nuances of real-world dynamics.

Definition.

A Markov Chain is termed **ergodic** if it embodies both irreducibility and aperiodicity.

Ergodicity ensures the existence of a unique limit distribution π, which characterizes the long-term behavior of a Markov Chain.

Theorem 2.

A Markov Chain that is irreducible and aperiodic has a unique limit distribution π, which satisfies:

$$\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j, \quad \forall i.$$

To illustrate these concepts, we consider an *influence game* in which a group or society must reach an agreement. The game follows these rules:

- One person speaks first and then chooses another person to speak next.
- The process continues, with each speaker selecting the next person to speak.

Different selection rules generate different Markov chains. We analyze the following cases:

1. **Periodicity** Consider a situation where two people continuously pass the word to each other. The transition matrix is:
$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
This chain has period 2 since each person can only speak every second step. Therefore, it is not aperiodic and does not converge to a unique stationary distribution.

2. **Irreducibility** A Markov Chain is irreducible if every state is reachable from any other state. Consider a case where some members of the group never receive the word. For instance, if three groups exist but never communicate, the system is reducible:
$$P = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$
Here, the first two people can never reach the last two, making the chain reducible.

Now, consider a hierarchical model where:

- A **leader** is at the top and receives the word with probability 0.5.
- **Second-level leaders** receive the word with probability 0.25.
- **Lower-level members** receive it with decreasing probabilities.

We can model this as an *infinite-state* Markov chain, where person X_1 (the leader) speaks first, and at each step, the probability of passing the word down follows a structured pattern:

$$P = \begin{bmatrix} 0.5 & 0.25 & 0.25 & 0 & 0 & 0 & 0 & \dots \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 & \dots \\ 0.5 & 0 & 0 & 0 & 0.5 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0.5 & 0 & 0 & 0 & 0 & 0.5 & 0 & \dots \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0.5 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Is This Chain Ergodic?

Irreducibility: Every person can be reached from any other in a finite number of steps. For example, from C, we always return to L, and L can reach anyone.

Aperiodicity: The system has no strict cycle since any transition probabilities allow for variation.

Since the chain is both irreducible and aperiodic, it is **ergodic**, meaning it has a unique stationary distribution.

Thus, the hierarchical influence game reaches a stable equilibrium, where the influence of each member is captured in the limit distribution π.