Markov Chain and Markov Process

Dr. Collins Odhiambo

Institute of Mathematical sciences, Strathmore University

codhiambo@strathmore.edu

June 22, 2022

Overview

Overview of Markov Processes

- The Markov property means that evolution of the Markov process in the future depends only on the present state and does not depend on past history.
- The Markov process does not remember the past if the present state is given. Hence, the Markov process is called the process with memoryless property.
- Let $X = 1, \cdots, n$ be a discrete state space and let $x(t)$ be a Markov chain (with the Markov property as defined above) on X, where t may be either discrete or continuous. The system "jumps" between the states of X in time. Such a jump is called a transition.

$$P(x(t)|x(t-1), x(t-2), \cdots, x(1)) = P(x(t)|x(t-1))$$

Overview of Markov Processes

- **Higher-order Markov chains**: A Markov chain of order m (or a Markov chain with memory m) where m is finite, is a process for t>m satisfying

$$P(x(t)|x(t-1), x(t-2), \cdots, x(1)) = P(x(t)|x(t-1), x(t-2), \cdots, x(m))$$

- In other words, the future state depends on the past m states. It is possible to construct a chain $(y(t))$ from $(x(t))$ which has the 'classical' Markov property.
- Let $y(t) = (x(t)|x(t-1), x(t-2), \cdots, x(t-1), y(t-m+1))$, the ordered m-tuple of x values. Then $y(t)$ is a Markov chain with state space X^m and has the classical Markov property. This is a result of Taken's embedding theorem.

General Properties of discrete Markov processes

- **Time-homogeneity** Time-homogeneous Markov chains (or stationary Markov chains) are processes where

$$P(x(t+1)=i|x(t)=j)=P(x(s+t+1)=i|x(s+t)=j)$$

  for all s. The probability of the transition is an invariant property of the system, i.e. independent of the time when we evaluate the Markov chain.

- A state j is said to be accessible from a state i (written $i \to j$) if a system started in state i has a non-zero probability of transitioning into state j at some point. Formally, state j is accessible from state i if there exists a time $t \geq 0$ such that

$$P(x(t)=j|x(0)=i)>0.$$

  Allowing t to be zero means that every state is defined to be accessible from itself.

General Properties of discrete Markov processes

- A state i is said to communicate with state j (written $i \leftrightarrow j$) if both $i \to j$ and $j \to i$. A set of states C is a **communicating class** if every pair of states in C communicates with each other, and no state in C communicates with any state not in C.
- A communicating class is **closed** if the probability of leaving the class is zero, namely that if i is in C but j is not, then j is not accessible from i.
- A Markov chain is said to be **irreducible** if its state space is a single communicating class; in other words, if it is possible to get to any state from any state, otherwise it is called **Reducible**. (A Markov chain is said to be **irreducible** if all states communicate with each other.)

General Properties of discrete Markov processes

- **Recurrence** A state i is said to be transient if, given that we start in state i, there is a non-zero probability that we will never return to i. Formally, let the random variable T_i be the first return time to state i:

$$T_i = \inf\{t \geq 1 : x(t) = i | x(0) = i\}.$$

  If a state i is not transient (it has finite hitting time with probability 1), then it is said to be **recurrent** or **persistent**.

- Let M_i be the expected return time,

$$M_i = E[T_i].$$

  Then, state i is **positive recurrent** if M_i is finite; otherwise, state i is **null recurrent** (the terms non-null persistent and null persistent are also used, respectively).

- A state i is called **absorbing** if it is impossible to leave this state.

- The **period** of a state i is the largest integer d satisfying the following property: $p^(n)_i i = 0$, whenever n is not divisible by d. The period of i is shown by $d(i)$. If $p^(n)_i i = 0$, for all $n > 0$, then we let $d(i) = \infty$.
  - If d(i)¿1, we say that state i is periodic.
  - If d(i)=1, we say that state i is aperiodic.
- You can show that all states in the same communicating class have the same period.
- A class is said to be periodic if its states are periodic. Similarly, a class is said to be aperiodic if its states are aperiodic. Finally, a Markov chain is said to be aperiodic if all of its states are aperiodic. If

$$i \leftrightarrow j, \Rightarrow d(i) = d(j).$$

Ergodic Markov Chains

- A Markov chain is called an ergodic chain if it is possible to go from every state to every state (not necessarily in one move).
- A Markov chain is said to be ergodic if there exists a positive integer T_0 such that for all pairs of states i, j in the Markov chain, if it is started at time 0 in state i then for all $t > T_0$, the probability of being in state j at time t is greater than 0.
- For a Markov chain to be ergodic, two technical conditions are required of its states and the non-zero transition probabilities; these conditions are known as **irreducibility** and **aperiodicity**. Informally, the first ensures that there is a sequence of transitions of non-zero probability from any state to any other, while the latter ensures that the states are not partitioned into sets such that all state transitions occur cyclically from one set to another.

Other Prepositions

- For any communication class C, all states in C are either recurrent or all states in C are transient. Thus if i and j communicate and i is recurrent, then so is j. Equivalenly if i and j communicate and i is transient, then so is j. In particular, for an irreducible Markov chain, either all states are recurrent or all states are transient.
- An irreducible Markov chain with a finite state space is always recurrent: all states are recurrent.
- Suppose $i \neq j$ are both recurrent. If i and j communicate and if j is positive recurrent ($E(\tau_{jj}) < \infty$), then i is positive recurrent ($E(\tau_{ii}) < \infty$) and als($E(\tau_{ij}) < \infty$). In particular, all states in a recurrent communication class are either all together positive recurrent or all together null recurrent.

Steady-state analysis and limiting distributions

- A Markov process has a unique stationary distribution π if and only if it is irreducible and all of its states are positive recurrent.
- Note that this in particular includes ergodic processes, because ergodicity is a stronger requirement. In that case, π is related to the expected return time:
$$\pi = \frac{1}{M_i}.$$
- Further, if the chain is both irreducible and aperiodic, then for any i and j,
$$\lim_{t\to\infty} P(x(t) = j|x(0) = i) = \frac{1}{M_i}.$$
- Note that there is no assumption on the starting distribution; the chain converges to the stationary distribution regardless of where it begins. Such π is called the equilibrium distribution of the chain.

Steady-state analysis and limiting distributions

Limiting distribution of Markov chain

For an irreducible and aperodic Markov chain, we have $P_{ij}(n) \to \frac{1}{\mu_j}$ as $n \to \infty$ for all i and j

Results:

- When the chain is transient or null persistent, $\mu = \infty$, therefore $p_{ij}(n)$ as $n \to \infty, \forall i, j$
- If the chain is non-null persistent, $P_{ij}(n) \to \frac{1}{\mu_j} = \pi_j$, That is, $P_{ij}(n)$ converges to the probability under the stationary distribution of going to state j
- $\lim_{n\to\infty} P_{ij}(n)$ does not depend on the starting point i, which is technically stated as the chain forget its origin.

$$P(X_n = j) \sum_i P(X_0 = i) p_{ij}(n) \to \frac{1}{\mu_j} \sum_i P(X_0 = i) = \frac{1}{\mu_j}$$

Example: Limiting distributions

Consider a Markov chain with two possible states, $S = \{0, 1\}$. In particular, suppose that the transition matrix is given by

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

where a and b are two real numbers in the interval $[0, 1]$ such that $0 < a + b < 2$. Suppose that the system is in state 0 at time n=0 with probability α, i.e.,

$$\pi^{(0)} = \begin{bmatrix} P(X_0 = 0) & P(X_0 = 1) \end{bmatrix} = \begin{bmatrix} \alpha & 1-\alpha \end{bmatrix},$$

where $\alpha \in [0, 1]$
Using induction (or any other method), show that

From previous slide-Example: Limiting distributions

1. Show that

$$P^n = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b)^n}{a+b}\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}.$$

2. Show that

$$\lim_{n\to\infty} P^n = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix}.$$

3. Show that

$$\lim_{n\to\infty} \pi^{(n)} = \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}.$$

R Scripts for Markov Chain

R Scripts for Markov Chain

The End