ST3236 Cheatsheet
by Yiyang, AY21/22

Probability Theory, Review

Boole's Inequality

For any events $A_1, A_2, \ldots$,

$$\mathbb{P}\Big(\bigcup_{n\geq 1} A_n\Big) \leq \sum_{n\geq 1} \mathbb{P}(A_n)$$

Conditional Expectation, Properties

- Linearity: $\mathbb{E}(aX_1 + bX_2|Y) = a\mathbb{E}(X_1|Y) + b\mathbb{E}(X_2|Y)$
- Law of Iterated Expectation: $\mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(X)$
- Tower Property: $\mathbb{E}(\mathbb{E}(X|Y,Z)|Y) = \mathbb{E}(X|Y)$
- Independence: $\mathbb{E}(X|Y) = \mathbb{E}(X)$, for X and Y independent.

Markov Chain (MC) Basics

A (discrete-time) **Markov Chain** (MC) is a stochastic process $\{X_n\}_{n=0}^{\infty}$ that satisfies,

$$\begin{aligned}&\mathbb{P}(X_{n+1} = t_{n+1}|X_n = t_n, X_{n-1} = t_{n-1}, \ldots, X_1 = t_1)\\ =&\mathbb{P}(X_{n+1} = t_{n+1}|X_n = t_n)\end{aligned}$$

,whenever $\mathbb{P}(X_n = t_n, X_{n-1} = t_{n-1}, \ldots, X_1 = t_1) > 0$. This property is also called the **Markovian Property**.

A **Time-homogeneous** Markov Chain is one where the conditional probability $\mathbb{P}(X_{n+1} = j|X_n = i)$ does not depend on n. Equivalently, it means for all $n \geq 0$ and $i, j \in S$,

$$\mathbb{P}(X_{n+1} = j|X_n = i) = \mathbb{P}(X_1 = j|X_0 = i)$$

For time-homogeneous MCs, $p_{ij} = \mathbb{P}(X_1 = j|X_0 = i)$ is the **1-step transition probability**, and $P = ((p_{ij}))_{i,j\in S}$ is the **Transition matrix**. Similarly, $p_{ij}^{(k)} = \mathbb{P}(X_k = j|X_0 = i)$ is the **k-Step Transition Probability**.

Chapman-Komogorov Equation

$$p_{ij}^{(k)} = \sum_{i_1,\ldots,i_{k-1}\in S} p_{ii_1}p_{i_1i_2}\ldots p_{i_{k-1}j}$$

Equivalently, it means $P^{(k)} = P^k$.

Accessing States, MC

Accessible States

For states i, j of a MC, j is **Accesible** from i, $i \to j$, if there exists $k \geq 0$ such that,

$$p_{ij} = \mathbb{P}(X_k = j|X_0 = i) > 0$$

States i and j **Communicate**, i.e. $i \leftrightarrow j$ iff. $i \to j \wedge j \to i$.

$\leftrightarrow$ is a **Equivalence Relation**. The equivalent classes of S partitioned by it are called **Communicating classes**.

Notes:

- $i \leftrightarrow i$ as $p_{ii}^{(0)} = 0$ for all $i \in S$
- For $i \in S_i$ and $j \in S_j$, it is possible that $i \to j$.
- For $i \in S_i$ and $j \in S_j$, $i \to j \implies j \not\to i$.

A MC is **Irreducible** if there is only one communicating class for the state space. Otherwise, the MC is **Reducible**.

Essential States

A state i is **Essential** if for every state j it satisfies

$$i \to j \implies j \to i$$

A state that is not essential is called **Inessential**.

Properties of Essential States

- If i essential and $i \to j$ then j is essential.
- All states in one communicating class are either all essential or all inessential.
- A **finite state** MC must have at least one essential state.
- An absorbing state is essential.

Recurrence Theory

Definition

A state i of a Markov Chain is **Recurrent** if

$$P(X_n = i \text{ for some } n \geq 1|X_0 = i) = 1$$

Otherwise, the state i is **Transient**.

Let $T_i = \min\{n \geq 1 : X_n = i\}$ firt time the chain visits state i, and let $\mu_i = \mathbb{E}[T_i|X_0 = i]$ expected return time to state i, then for a recurrent state i,

- State i is **Positive Recurrent**, if $\mu_i < \infty$
- State i is **Null Recurrent**, if $\mu_i = \infty$

Overall, a state is either 1) transient, 2) null recurrent, or 3) positive recurrent.

Equivalently, for a state i,

- If i is recurrent, $P(X_n = i \text{ infinitely often}|X_0 = i) = 1$
- If i is transient, $P(X_n = i \text{ infinitely often}|X_0 = i) = 0$

Criteria for Determining Recurrence

Number of Visits

Define $f_i = P(X_n = i \text{ for some } n \geq 1|X_0 = i)$ and $N_i = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=i\}}$ as number of visits of state i.
Analysis: Consider the probability of the chain "escaping" from returning to state i, $(N_i + 1)|X_0 = i \sim Geom(1 - f)i)$.

A state i is recurrent, if $P(N_i = \infty|X_0 = i) = 1$. The negation applies for transient state i.

A state i is recurrent, if $\mathbb{E}[N_i|X_0 = i] = \infty$. The negation applies for transient state i.

Properties of Recurrence States

Recurrence is a class property. If i is recurrent and j is in the same communicating class as i, j is recurrent. (Proof Idea: Chapman-Kolmogorov Equation.) **So are Positive and Null Recurrence.**

There always exists a recurrent state in a finite space Markov Chain. (Proof Idea: By contradiction)

First Time Passage

First Passage probability from state i to state j in exactly k steps is defined $f_{i,j}^{(k)} = P(X_k = j, X_{k-1} \neq j, \ldots, X_1 \neq j|X_0 = i)$.

Related Expressions

Renewal Theorem

For $i \neq j$ and $n \geq 1$, we have

$$p_{i,j}^{(n)} = \sum_{k=1}^{n} f_{i,j}^{(k)} p_{j,j}^{(n-k)}$$

Note: Algorithm to compute First Passage Times: To compute $f_{i,j}^{(n)}$, obtain the first n powers of P, then write and solve a linear system involving first n basic renewal equations to get $f_{i,j}^{(1)}, \ldots, f_{i,j}^{(n)}$.

Recursive Relation

Consider first passage times with steps $n \geq 2$, by Markovian Property,

$$f_{i,j}^{(n)} = \sum_{k\neq j} p_{i,k} f_{k,j}^{(n-1)}$$

Note: The recursive relation can also be used to calculate First Passage Times.

Relations to Recurrence

Define $f_{i,j}$ as the probability of chain ever visiting state j starting from state i for $i \neq j$,

$$f_{i,j} = P(X_n = j \text{ for some } n \geq 1|X_0 = i) = \sum_{n=1}^{\infty} f_{i,j}^{(n)}$$

An interesting Identity

For any state $i \neq j$, by using the previous recursive relation,

$$f_{i,j} = p_{i,j} + \sum_{k\neq j} p_{i,k} f_{k,j}$$

Recurrent State Visiting Another

If i is recurrent and $i \to j$, then $f_{i,j} = 1$ and $f_i = f_{i,i} = 1$.

For any state i and a recurrent state j,

$$f_{i,j} = \sum_{k \in S} p_{i,k} f_{k,j}$$

It is valid for both cases of $i \neq j$ and $i = j$.

Stationary Distribution

Overview

A **Stationary Distribution**, π, of a Markov Chain is one such that $\pi^T P = \pi^T$.

Number of Visits Analysis

Define $\phi_i(j)$ as the expected number of timers the Markov Chain visits j when starting from i before returning to i for $i \neq j$,

$$\phi_i(j) = \mathbb{E}[\sum_{n=0}^{T_i - 1} \mathbb{1}_{\{X_n = j\}} | X_0 = i]$$

, then by definition, $\sum_{j \in S} \phi_i(j) = \mathbb{E}(T_i | X_0 = i) = \mu_i$.

Of Positive Recurrent States

For a **Positive Recurrent** state i, the vector $\pi_i j$ forms a stationary distribution for all states $j \in S$,

$$\pi_i(j) = \frac{\phi_i(j)}{\sum_k \phi_i(k)}$$

Then if state i is in an irreducible chain, $\pi = \pi_i$ is the unique stationary distribution.

Of Transient & Null Recurrent States

Suppose a stationary distribution pi exists for a general Markov Chain, then $\pi(i) = 0$ for all transient or null recurrent states i.
Proof Idea: Consider $\phi_i(j)$ of such states.

Of Irreducible Markov Chains

Properties

For an irreducible Markov Chain, below are equivalent:

1. At least one state is positive recurrent
2. A stationary distribution exists
3. A stationary distribution exists and is unique
4. All states are positive recurrent

Note: The four statements above are just equivalent. They may not always be true for an irreducible chain!

Related Expressions

For an **irreducible** Markov Chain with a stationary distribution π, and for all states $i \in S$,

$$\pi(i) = \frac{1}{\mu_i} = \frac{1}{\mathbb{E}[T_i | X_0 = i]}$$

Special Cases

Any finite state space Markov Chain always have a stationary distribution.

Of General Markov Chains

Algorithm for Constructing a Stationary Distribution

For any Markov Chain, procedures for constructing a stationary distribution if there exists one:

1. For each positive recurrent class $C_j \subset S$, let π_j^* denotes the unique stationary distribution of the Markov Chain with transition matrix reduced from P to $P_{C_j \times C_j}$.
2. Let $\bar{\pi}_j^*$ denotes the vector of size equal to $|S|$ by appending 0 to π_j^* for all states $i \in S \quad C_j$.
3. Then probability vector $\pi = \sum_{j: C_j \text{ +ve recurrent}} \alpha_j \bar{\pi}_j^*$ forms a stationary distribution for all non-negative constants $a_1, \ldots,$ that sum to 1.

Significance of the Algorithm

A Markov Chain has a stationary distribution as long as it has **a positive recurrent class**. A finite state Markov Chain always has a stationary distribution. Proof idea: as it cannot have all of its finite number of states transient or null recurrent.

Long-Term Behaviours

Periodicity

For a **recurrent** state i of a Markov Chain, its **Period** is defined as the maximum integer such that the chain returns from that sate to itself only in step sizes which are multiples of the integer.

Mathematically, for a recurrent state i, we define $Z_i = \{n \geq 1 : p_{ii}^{(n)} > 0\}$, then the period of i is defined as $d_i = \gcd Z_i$, where $\gcd$ is the greatest common divisor.

A state i is **Aperiodic** if $d_i = 1$.

Properties of Periodicity

Periodicity is a Class Property. For two recurrent states i and j in the same communicating class, $d_i = d_j$.
Therefore, a class is **aperiodic** iff. any of its states is **aperiodic**.

An aperiodic state can lead to itself in steps of all sizes large enough. Mathematically, it means if $d_i = 1$ then there exists $K \in \mathbb{Z}^+$ such that $p_{ii}^{(k)} > 0$ for all $k \geq K$.

The Convergence Theorem

For an **irreducible**, **aperiodic** and **positive recurrent** chain, let π be its unique stationary distribution, then as $n \to \infty$, we have

$$p_{i,j}^{(n)} \to \pi(j)$$

, for all states i, j.
It means that for such a chain, the stationary distribution is its limiting distribution.

Intermediate Results & Lemmas

Homework5 Qn1 In a finite state space Markov Chain, essentiality is equivalent to recurrence.