Review Sheet

EE 226A

Ahmed Shakil

Contents

1 Discrete-Time Markov Chains **2**
1.1 Definition of a Markov Chain 2
1.1.1 The Chapman-Kolmogorov equations 2
1.2 Markov Limit Theorems 3
1.2.1 Stationary distributions and the issue of convergence 4
1.3 4

1 Discrete-Time Markov Chains

1.1 Definition of a Markov Chain

Definition 1.1: Markov chain

A Markov chain is a process $(X_n)_{n \geqslant 0}$ satisfying

$$\Pr\{X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\} = \Pr\{X_{n+1} = j \mid X_n = i\}$$

for all $n \geqslant 1$ and $j, i, i_{n-1}, i_0 \in \mathcal{S}$. A Markov chain is said to be **temporally homogeneous** if there are numbers $(p_{ij})_{i,j \in \mathcal{S}}$ such that

$$\Pr\{X_{n+1} = j \mid X_n = i\} = p_{ij}$$

for all $n \geqslant 0$ and all states $i, j \in \mathcal{S}$. The numbers $(p_{ij})_{i,j \in \mathcal{S}}$ are generically referred to as the **transition probabilities** of the Markov chain.

Definition 1.2: Transition matrix

$$P = \begin{bmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

The matrix P is called the transition matrix. It is a stochastic matrix, which means it is square with non-negative entries, whose rows sum to one. A Markov chain and transition matrix are equivalent representations of each other.

1.1.1 The Chapman-Kolmogorov equations

Proposition 1.3: The Chapman-Kolmogorov Equations

We define **multi-step transition probabilities**

$$P_{ij}^n \coloneqq \Pr\{X_{n+m} = j \mid X_m = i\}, \quad n, m \geqslant 0.$$

The Chapman-Kolmogorov equations give a recursive formula for computing the n-step transition probabilities.
For all $m, n \geqslant 0$ and states i, j,

$$P_{ij}^{m+n} = \sum_k P_{ik}^m P_{kj}^n.$$

In particular, we have

$$P^n = \begin{bmatrix} P_{00}^n & P_{01}^n & \cdots \\ P_{10}^n & P_{11}^n & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

where P^n is the transition matrix P raised to the n^{th} power. Also note $P^{m+n} = P^m P^n$.

Definition 1.4: Irreducible

A **class** of states is a nonempty set of states such that every state in the set can communicate with one another. These classes form an equivalence relation and partition the state space. We say a Markov chain is irreducible if there is only one class.

Definition 1.5: Periodicity

For a state i, define its **period**

$$d(i) := \gcd\{n \geqslant 1 : P_{ii}^n > 0\}.$$

States with period 1 are called **aperiodic**. Periodicity is a class property.

Proposition 1.6

Define the **first return time** for state $j \in \mathcal{S}$ as

$$T_j := \inf\{n \geqslant 1 : X_n = j\}.$$

Let $(X_n)_{n \geqslant 0}$ be a Markov chain with transition matrix P. Conditioned on the event $\{T_j < \infty\}$, the process $(X_{n+T_j})_{n \geqslant 0}$ is a Markov chain with transition matrix P and starting state j and is independent of $X_0, \ldots, X_{T_j}$.

Proposition 1.7

A state j is **recurrent** if $\Pr\{T_j < \infty \mid X_0 = j\} = 1$, and it is called **transient** if $\Pr\{T_j < \infty \mid X_0 = j\} < 1$. Recurrence and transience are class properties.

Lemma 1.8

State i is recurrent if and only if $\sum_{n=1}^{\infty} P_{ii}^n < \infty$.

1.2 Markov Limit Theorems

Theorem 1.9: Strong Law of Large Numbers for Markov Chains

Define $N_j(n)$, $n \geqslant 1$, to be the number of transitions into state j, up to and including time n. More precisely,

$$N_j(n) := \#\{1 \leqslant k \leqslant n : X_k = j\}.$$

Also define the expected first return time to be

$$\mu_{jj} := \mathbb{E}[T_j | X_0 = j].$$

Let $(X_n)_{n \geqslant 0}$ be a Markov chain starting in state $X_0 = i$. If $i \leftrightarrow j$, then

$$\frac{N_j(n)}{n} \to \frac{1}{\mu_{jj}} \quad \text{a.s.}$$

Corollary 1.10

For an irreducible Markov chain $(X_n)_{n \geqslant 0}$, we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} P_{ij}^k = \frac{1}{\mu_{jj}}.$$

1.2.1 Stationary distributions and the issue of convergence

Definition 1.11: Sationary Distribution

A probability distribution $(\pi_j)_{j\in\mathcal{S}}$ is a stationary distribution for a Markov chain with transition probability matrix P if $\pi_j = \sum_i \pi_i p_{ij}$ for each $j \in \mathcal{S}$. Equivalently in matrix notation, $\pi = \pi P$, when π is considered as a row vector.

Definition 1.12: Positive and Null recurrence

A recurrent state j is positive recurrent if $\mu_{jj} < \infty$, or null recurrent if $\mu_{jj} = \infty$. Positive and null recurrence are class properties.

It is important to note stationary distributions aren't necessarily unique. It is important to note that stationary distribution does not always exist.

Theorem 1.13

An irreducible Markov chain satisfies exactly one of the following:

1. All states are transient, or all states are null recurrent. In this case, $\frac{1}{n}\sum_{k=1}^{n} P_{ij}^k \to 0$ as $n \to \infty$ for all states i, j, and no stationary distribution exists.

2. All states are positive recurrent. In this case, a unique stationary distribution exists and is given by $\pi_j = \frac{1}{\mu_{jj}} = \lim_{n\to\infty} \frac{1}{n}\sum_{k}^{n} P_{ij}^k$.

Theorem 1.14

Let $(X_n)_{n\geqslant 0}$ be an irreducible, aperiodic, and positive recurrent Markov chain with stationary distribution π. Then

$$\lim_{n\to\infty} \sum_j \left|P_{ij}^n - \pi_j\right| = 0 \text{ for all } i \in \mathcal{S}.$$

In particular, $P_{ij}^n = \pi_j$ for all i, j.

1.3 Reversibility and Spectral Gap