Part IB — Markov Chains

Definitions

Based on lectures by G. R. Grimmett

Notes taken by Dexter Chua

Michaelmas 2015

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Discrete-time chains

Definition and basic properties, the transition matrix. Calculation of n-step transition probabilities. Communicating classes, closed classes, absorption, irreducibility. Calculation of hitting probabilities and mean hitting times; survival probability for birth and death chains. Stopping times and statement of the strong Markov property. [5]

Recurrence and transience; equivalence of transience and summability of n-step transition probabilities; equivalence of recurrence and certainty of return. Recurrence as a class property, relation with closed classes. Simple random walks in dimensions one, two and three. [3]

Invariant distributions, statement of existence and uniqueness up to constant multiples. Mean return time, positive recurrence; equivalence of positive recurrence and the existence of an invariant distribution. Convergence to equilibrium for irreducible, positive recurrent, aperiodic chains *and proof by coupling*. Long-run proportion of time spent in a given state. [3]

Time reversal, detailed balance, reversibility, random walk on a graph. [1]

Contents

0 Introduction

1 Markov chains

1.1 The Markov property

Definition (Markov chain)**.** Let $X = (X_0, X_1, \cdots)$ be a sequence of random variables taking values in some set S, the *state space*. We assume that S is countable (which could be finite).

We say X has the *Markov property* if for all $n \geq 0, i_0, \cdots, i_{n+1} \in S$, we have

$$\mathbb{P}(X_{n+1} = i_{n+1} \mid X_0 = i_0, \cdots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n).$$

If X has the Markov property, we call it a *Markov chain*.

We say that a Markov chain X is *homogeneous* if the conditional probabilities $\mathbb{P}(X_{n+1} = j \mid X_n = i)$ do not depend on n.

1.2 Transition probability

Definition (n-step transition probability)**.** The n-step transition probability from i to j is

$$p_{i,j}(n) = \mathbb{P}(X_n = j \mid X_0 = i).$$

Notation. Write $P(m) = (p_{i,j}(m))_{i,j \in S}$.

2 Classification of chains and states

2.1 Communicating classes

Definition (Leading to and communicate)**.** Suppose we have two states $i, j \in S$. We write $i \to j$ (i *leads to* j) if there is some $n \geq 0$ such that $p_{i,j}(n) > 0$, i.e. it is possible for us to get from i to j (in multiple steps). Note that we allow $n = 0$. So we always have $i \to i$.

We write $i \leftrightarrow j$ if $i \to j$ and $j \to i$. If $i \leftrightarrow j$, we say i and j *communicate*.

Definition (Communicating classes)**.** The equivalence classes of $\leftrightarrow$ are *communicating classes*.

Definition (Irreducible chain)**.** A Markov chain is *irreducible* if there is a unique communication class.

Definition (Closed)**.** A subset $C \subseteq S$ is *closed* if $p_{i,j} = 0$ for all $i \in C, j \notin C$.

2.2 Recurrence or transience

Notation. For convenience, we will introduce some notations. We write

$$\mathbb{P}_i(A) = \mathbb{P}(A \mid X_0 = i),$$

and

$$\mathbb{E}_i(Z) = \mathbb{E}(Z \mid X_0 = i).$$

Definition (First passage time and probability)**.** The *first passage time* of $j \in S$ starting from i is

$$T_j = \min\{n \geq 1 : X_n = j\}.$$

Note that this implicitly depends on i. Here we require $n \geq 1$. Otherwise T_i would always be 0.

The *first passage probability* is

$$f_{ij}(n) = \mathbb{P}_i(T_j = n).$$

Definition (Recurrent state)**.** A state $i \in S$ is *recurrent* (or *persistent*) if

$$\mathbb{P}_i(T_i < \infty) = 1,$$

i.e. we will eventually get back to the state. Otherwise, we call the state *transient*.

2.3 Hitting probabilities

Definition (Hitting time)**.** The *hitting time* of $A \subseteq S$ is the random variable $H^A = \min\{n \geq 0 : X_n \in A\}$. In particular, if we start in A, then $H^A = 0$. We also have

$$h_i^A = \mathbb{P}_i(H^A < \infty) = \mathbb{P}_i(\text{ever reach } A).$$

2.4 The strong Markov property and applications

Definition (Stopping time)**.** Let X be a Markov chain. A random variable T (which is a function $\Omega \to \mathbb{N} \cup \{\infty\}$) is a *stopping time* for the chain $X = (X_n)$ if for $n \geq 0$, the event $\{T = n\}$ is given in terms of $X_0, \cdots, X_n$.

2.5 Further classification of states

Definition (Mean recurrence time)**.** Let T_i be the returning time to a state i. Then the *mean recurrence time* of i is

$$\mu_i = \mathbb{E}_i(T_i) = \begin{cases} \infty & i \text{ transient} \\ \sum_{n=1}^\infty n f_{i,i}(n) & i \text{ recurrent} \end{cases}$$

Definition (Null and positive state)**.** If i is recurrent, we call i a *null state* if $\mu_i = \infty$. Otherwise i is *non-null* or *positive*.

Definition (Period)**.** The *period* of a state i is $d_i = \gcd\{n \geq 1 : p_{i,i}(n) > 0\}$.

A state is *aperiodic* if $d_i = 1$.

Definition (Ergodic state)**.** A state i is *ergodic* if it is aperiodic and positive recurrent.

3 Long-run behaviour

3.1 Invariant distributions

Definition (Invariant distriubtion)**.** Let X_j be a Markov chain with transition probabilities P. The distribution $\pi = (\pi_k : k \in S)$ is an *invariant distribution* if

(i) $\pi_k \geq 0$, $\sum_k \pi_k = 1$.

(ii) $\pi = \pi P$.

The first condition just ensures that this is a genuine distribution.

An invariant distribution is also known as an invariant measure, equilibrium distribution or steady-state distribution.

3.2 Convergence to equilibrium

4 Time reversal

Definition (Reversible chain)**.** An irreducible Markov chain $X = (X_0, \cdots, X_N)$ in its invariant distribution π is *reversible* if its reversal has the same transition probabilities as does X, ie

$$\pi_i p_{i,j} = \pi_j p_{j,i}$$

for all $i, j \in S$.

This equation is known as the *detailed balance equation.* In general, if λ is a distribution that satisfies

$$\lambda_i p_{i,j} = \lambda_j p_{j,i},$$

we say (P, λ) is in *detailed balance.*