Markov Chains - Summary

Lectured by Prof Geoffrey Grimmett

Michaelmas 2014

Definition 1 (Markov Chain)**.** Let S be a set (the *state space*). Let $\mathbf{X} = (X_n : n = 0, 1, 2, \ldots)$ be a sequence of random variables taking values is S. Then X is a *Markov chain* if it satisfies the *Markov property*

$$\mathbb{P}(X_{n+1} = i_{n+1} \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n) \qquad \forall n \in \mathbb{N},\ i_0, i_1, \ldots, i_{n+1} \in S$$

Definition 2 (Homogeneous)**.** A Markov chain X is *homogeneous* if

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

does not depend on n.

Definition 3 (Initial distribution)**.** The *initial distribution* is $\lambda = (\lambda_i : i \in S)$ where $\lambda_i = \mathbb{P}(X_0 = i)$.

Definition 4 (Transition probabilities)**.** The (1-step) *transition probabilities* are $p_{i,j} = \mathbb{P}(X_{n+1} = j \mid X_n = i)$ and the *transition matrix* is $P = (p_{i,j})$ for $i, j \in S$.

Theorem 1. Let λ be a distribution on a state space S and P a transition matrix. Then the sequence $\mathbf{X} = (X_n : n = 0, 1, 2, \ldots)$ is a Markov chain with initial distribution λ and transition matrix P if and only if

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \lambda_{i_0} p_{i_0,i_1} p_{i_1,i_2} \ldots p_{i_{n-1},i_n} \quad \forall n \in \mathbb{N},\ i_0, \ldots, i_n \in S$$

Theorem 2 (Extended Markov Property)**.** Let X be a Markov chain. For $n \geq 0$, let H be an event given in terms of the history $(X_0, X_1, \ldots, X_{n-1})$ and let F be an event given in terms of the future $(X_{n+1}, X_{n+1}, \ldots)$. Then

$$\mathbb{P}(F \mid X_n = i, H) = \mathbb{P}(F \mid X_n = i) \quad \forall n \in \mathbb{N},\ i \in S$$

Definition 5 (n-step transition probabilities)**.** Let X be a Markov chain. Then the *n-step transition probability* is $p_{i,j}(n) = \mathbb{P}(X_n = j \mid X_0 = i)$.

Theorem 3 (Chapman-Kolmogorov)**.** Let X be a Markov chain over a state space S. Then

$$p_{i,j}(m+n) = \sum_{k \in S} p_{i,k}(m) p_{k,j}(n)$$

Definition 6 (Communication)**.** Let X be a homogeneous Markov chain over a state space S. Let $i, j \in S$.

We say i *leads to* j if $p_{i,j}(n) > 0$ for some $n \geq 0$ and write $i \to j$.

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

An equivalence class of $(S, \leftrightarrow)$ is called a *communicative class.*

Definition 7 (Reducibility)**.** Let S be a state space with an associated Markov chain. The chain is *irreducible* if S is the single equivalence (communicative) class. It is *reducible* otherwise.

Definition 8 (Closure and absorption)**.** A subset C of a state space S is *closed* if $i \in C, i \to j \implies j \in C$.

If $C = \{i\}$ is closed, the state i is called *absorbing.*

Notation. We write $\mathbb{P}(\cdot \mid X_0 = i) = \mathbb{P}_i(\cdot)$ and $\mathbb{E}(\cdot \mid X_0 = i) = \mathbb{E}_i(\cdot)$.

Definition 9 (First-passage time)**.** The *first-passage time* to a state j is

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

Definition 10 (First-passage probabilities)**.** The *first-passage probabilities* are $f_{i,j}(n) = \mathbb{P}_i(T_j = n)$. We also write $f_{i,j} = \sum_{n \geq 1} f_{i,j}(n)$.

Definition 11 (Recurrence and transience)**.** A state i is called *recurrent* (or *persistent*) if $f_{i,i} = \sum_{n \geq 1} f_{i,i}(n)$ satisfies $f_{i,i} = 1$. A non-recurrent state is called *transient.*

Theorem 4. A state i is recurrent if and only if $\sum_n p_{i,i}(n) = \infty$.

Theorem 5. Let C be a communicating class. Then

(i) either every state in C is recurrent or every state in C is transient;

(ii) if C contains some recurrent state then C is closed.

Theorem 6. Suppose the state space S is finite. Then

(i) there exists at least one recurrent state.

(ii) if the chain is irreducible, every state is recurrent.

Theorem 7 (Polya)**.** A random walk X on $\mathbb{Z}^d$ is recurrent if $d = 1$ or 2, and transient if $d \geq 3$.

Definition 12 (Hitting times and probabilities)**.** The hitting time for an event $A \subseteq S$ is

$$H^A = \inf \{n \geq 0 : X_n \in A\}.$$

(Note that if $X_0 \in A$ then $H^A = 0$.)

The hitting probability for an event A starting in a state i is

$$h_i^A = \mathbb{P}_i \left(H^A < \infty \right)$$

i.e. the probability that the Markov chain ever reaches A given that it starts in state i.

The expected hitting time is

$$k_i^A = \mathbb{E}_i \left(H^A \right) \qquad [= \infty \text{ if } h_i^A < 1].$$

Theorem 8. The vector $h^A = \left(h_i^A : i \in S\right)$ of hitting probabilities is the *minimal* non-negative solution of the equations

$$h_i^A = \begin{cases} 1 & \text{if } i \in A \\ \sum_{j \in S} p_{ij} h_j^A & \text{if } i \notin A \end{cases}$$

where *minimal* means that if $x = (x_i : i \in S)$ is any non-negative solution to the above then $h_i^A \leq x_i$ for all $i \in S$.

Definition 13 (Stopping time)**.** A random variable $T : \Omega \to \{0, 1, 2, \ldots\} \cup \{\infty\}$ is a *stopping time* if for all $n \geq 0$ the event $\{T = n\}$ is given in terms of $X_0, X_1, \ldots, X_n$.

Theorem 9 (Strong Markov Property)**.** Let X be a Markov chain with transition matrix P and let T be a stopping time for X. Given that $T < \infty$ and that $X_T = i$, the sequence $Y_k = X_{T+k}$ $(k \geq 0)$ is a Markov chain with transition matrix P. In addition, given $T < \infty$, $Y = (Y_k)$ is independent of the $X_0, X, \ldots, X_{T-1}$.

Definition 14 (Return visits)**.** Suppose $X_0 = i$, then $V = |\{n \geq 1 : X_n = i\}|$ is the number of return visits of X to state i.

Note that $\mathbb{P}_i(V \geq 1) = f_{i,i}$ and so inductively by the Strong Markov Property, $\mathbb{P}_i(V \geq k) = f_{i,i}^k$.

Definition 15 (Mean recurrence time)**.** The *mean recurrence time* of a state $i \in S$ is

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

Definition 16 (Null and positive recurrence)**.** Suppose state i is recurrent. Then i is *null* recurrent if $\mu_i = \infty$. Conversely, i is *positive* (or *non-null*) recurrent if $\mu_i < \infty$.

Definition 17 (Period)**.** The *period* d_i of a state $i \in S$ is

$$d_i = \gcd\{n \geq 1 : p_{i,i}(n) > 0\}.$$

We say that i is *aperiodic* if $d_i = 1$.

Definition 18 (Ergodic)**.** A state i is called *ergodic* if it is positive, recurrent and aperiodic.

Definition 19 (Invariant distribution)**.** Let X be a Markov chain with transition matrix P. The vector $\pi = (\pi_i : i \in S)$ is an *invariant distribution* if

(i) $\pi_i \geq 0$ for all i

(ii) $\sum_i \pi_i = 1$

(iii) $\pi = \pi P$

Theorem 10. Let X be an *irreducible* Markov chain. Then

(i) There exists an invariant distribution if some state is positive recurrent.

(ii) If there exists an invariant distribution π then every state is positive recurrent, and $\pi_i = \frac{1}{\mu_i}$ for each $i \in S$. In particular, there is a unique invariant distribution.

Theorem 11. Let $X = (X_n)$ be a Markov chain which is irreducible, positive recurrent and aperiodic. Then $p_{i,j}(n) \to \pi_j$ as $n \to \infty$ where $\pi = (\pi_j)$ is the unique invariant distribution.

Theorem 12. Denote $V_i(n) = |\{1 \leq m \leq n : X_m = i\}|$ (number of visits to i up to time n).

For an irreducible, positive recurrent Markov chain, $\frac{V_i(n)}{n} \Rightarrow \pi_i$ as $n \to \infty$ where $Z_n \Rightarrow c$ as $n \to \infty$ means 'convergence in probability', $\mathbb{P}(|Z_n - c| > \epsilon) \to \infty\ \forall \epsilon > 0$.

Theorem 13. Let $X = (X_0, X_1, \ldots, X_n)$ be an irreducible Markov chain with invariant distribution π. Suppose X_0 has distribution π. Define $Y_k = X_{n-k}$ for $0 \leq k \leq N$. Then Y is a Markov chain with transition matrix $\hat{p}_{ij} = \frac{\pi_j}{\pi_i} p_{ji}$ and invariant distribution π.

Definition 20 (Reversible Markov chain)**.** Let X and Y be defined as in the theorem above. Then X is reversible if Y and X have the same transition matrix, i.e. if

$$\pi_i p_{ij} = \pi_k p_{ji} \quad \text{for } i, j \in S$$

These are known as the *detailed balance equations.*

Theorem 14. Let X be an irreducible Markov chain with transition matrix P. Let π be a distribution satisfying the detailed balance equations $\pi_i p_{ij} = \pi_j p_{ji}$ for $i, j \in S$. Then π is the unique invariant distribution.