Lecture 8

David Aldous

14 September 2015

Ideas used in Lecture 7.

- First-step analysis of simple asymmetric random walk.
- Analysis of "success runs" chain using special structure.
- Analysis of "death and immigration" chain using special structure.

Recall that the transition matrix $\mathbf{P}$ of a Markov chain can be represented as a weighted directed graph. In this lecture we first look at some "structure theory". This will show us that some qualitative aspects of the chain's behavior do not depend on actual numerical transition probabilities but only on the graph of possible transitions.

This material can be found in Chapter 4 of [KP].

Here are some definitions that depend only on the graph of possible transitions.

- j is **accessible** from i if there is a (directed) path from i to j (or $i = j$).
- i and j **communicate** if each is accessible from the other.
- Because "communicate" is an equivalence relation, the state space **States** can be partitioned into **communicating classes** (CCs), say $C_1, C_2, \ldots$, such that i and j **communicate** if and only if they are in the same CC.
- A class C is **open** if it is possible to leave; that is if $p_{ij} > 0$ for some $i \in C$ and $j \notin C$. Otherwise it is **closed**.
- The graph is **strongly connected** if there is only one CC, that is if all states communicate. In the Markov chain context this property is called **irreducible**.

I will give some theory results but just outline proofs.

Lemma. If the state space is finite then there exists at least one closed CC.

If the hitting time $T_C = \min\{t : X_t \in C\}$ on a **closed** class C is finite, then the chain must remain in C for all times $t \geq T_C$.

Lemma. If the state space is finite then, for any initial i,

$$\mathbb{P}_i(T_C < \infty \text{ for some closed } C) = 1.$$

The probabilities $p_{i,C} = \mathbb{P}_i(T_C < \infty)$ can be calculated using first-step analysis.

Stationary distributions

Recall the distribution $\mu(t)$ of X_t evolves as $\mu(t) = \mu(t-1)\mathbf{P}$ in vector-matrix notation. So suppose a probability distribution $\pi = (\pi_i, i \in \mathbf{States})$ satisfies

$$\pi = \pi\mathbf{P}; \quad \text{that is } \sum_i \pi_i p_{ij} = \pi_j \ \forall j. \tag{1}$$

If the chain has initial (time-0) distribution $\mu(0) = \pi$ then $\mu(t) = \pi$ for every time t. A distribution π satisfying (1) is called **stationary**.

This language is a bit confusing, when we imagine a Markov chain as a particle jumping between states. The particle continues to move even when we have a stationary distribution; **stationary** refers to the fact that the **probabilities** (of where the particle is at time t) do not change with time t.

We will soon see theory relating long-term behavior of a Markov chain to its stationary distribution. First we look at some examples where we can easily find the stationary distribution.

Two remarks.

(1) The notion of stationary distribution is also useful when the number of states is infinite, though some of our theorems assume a finite number of states.

(2) Usually we find a stationary distribution by first finding numbers $w_i > 0$ such that

$$\sum_i w_i p_{ij} = w_j \ \forall j$$

and then *normalizing* by setting

$$\pi_i = w_i / w, \quad \text{where } w = \sum_i w_i.$$

If the number of states is infinite, this only works if $w < \infty$.

Special setting: Doubly stochastic chains

By definition a transition matrix **P** has the property $p_{ij} \geq 0 \ \forall i, j$ and the *stochastic matrix* property

$$\sum_j p_{ij} = 1 \ \forall i.$$

A matrix that has the extra property

$$\sum_i p_{ij} = 1 \ \forall j$$

is called **doubly stochastic**. Given a doubly stochastic transition matrix on n states, it is clear that the uniform distribution $\pi_i = 1/n \ \forall i$ is a stationary distribution.

Example: asymmetric RW on *n*-cycle. [board]

Example: card-shuffling models. [board]

Special setting: success runs.
Here the states are $\{0, 1, 2, \ldots\}$ and the transition probabilities are of the form

$$p_{i,i+1} = q_i, \;\; p_{i,0} = 1 - q_i$$

where $0 < q_i < 1$.

Here we calculate [board]

$$\pi_i = \frac{s_i}{\sum_{j=0}^{\infty} s_j} \text{ where } s_i = \prod_{j=0}^{i-1} q_j .$$

Here we are assuming $\sum_{j=0}^{\infty} s_j < \infty$.

Special setting: detailed balance.

Suppose we can find numbers $w_i > 0$ such that

$$w_i p_{ij} = w_j p_{ji} \quad \forall i, j. \tag{2}$$

Then

$$\pi_i = w_i / \sum_j w_j$$

is a stationary distribution.

[board]

The condition (2) is called **detailed balance**. It is stronger than the **balance** condition

$$\sum_i w_i p_{ij} = w_j \ \forall j$$

which is the requirement for $(w_i / \sum_j w_j)$ to be a stationary distribution.

Example: random walk on a weighted undirected graph.

Suppose we are given an undirected graph, and suppose there is a "weight" $a_{ij} = a_{ji} > 0$ on each edge (i, j). Define $a_i = \sum_j a_{ij}$. Then

$$p_{ij} = a_{ij}/a_i$$

defines a transition matrix.

Easy to check [board] that the detailed balance condition (2) always holds for $w_i = a_i$. So

$$\pi_i = a_i/a, \quad a = \sum_j a_j$$

is a stationary distribution.