Stats 150, Fall 2019
Lecture 3, Thursday, 9/5/2019

Administrative Book-keeping: Homeworks are due weekly on Thursday at midnight. The assignment for the following week will be released at worst (latest) Friday morning.
Pitman opens 4-6pm and 6-8pm for Office Hours.

Pitman notes that we've only really made it through to §1.2, and we need to catch up to our advertised speed. We will handle in-class questions but will omit most proofs. For this lecture and forward, we will assume the audience has read the relevant textbook sections.

1 §1.3 : Classification of States

Consider a Markov chain with transition matrix P. With the state space S, consider states $x, y, i, j \in S$. Take P, S to be fixed.
We are interested in 'hitting times':

$$T_B := \min\{n \geq 1 : X_n \in B\}$$

Given boundary state B, we consider one path X_n to get from x to T_B. We call these **first passage (hitting) times**.
The immediate (pedantic) issue with this is if $X_n \notin B$ for all n. In this case, we need to define the convention:

$$\min\{\emptyset\} = \inf\{\emptyset\} := \infty$$

Theorem 1.1. Strong Markov Property (SMP)

Given $T_y = n < \infty$ and $X_n = y$, then

$$(X_n, X_{n+1}, X_{n+1}, \ldots)$$

is the original MC given that it starts at y.

To see this, we start with $X_0, X_1, X_2, \ldots$ which is a Markov chain. Here, X_0 has any initial distribution.

The distribution of $X_{n+1} \mid X_n = y$ is $\delta_y P(\cdot) = \mathbb{P}(y, \cdot)$, which is:

$$\mathbb{P}(X_{n+1} = z \mid T_y = n, X_n = y) = \mathbb{P}(y, z), \text{ so}$$
$$\mathbb{P}(X_{n+1} = z, X_{n+2} = w \mid T_y = n, X_n = y) = \mathbb{P}(y, z)\mathbb{P}(z, w),$$

and so on, giving rise to a family of equations.

Proof. See Durett page 14. □

Remark: In discrete time (even with general state space), **all** Markov chains have the Strong Markov Property.

Now, we can use the SMP to discover and prove things about Markov chains.

2 Iterating:

$$T^k := \min\{n > T_y^{k-1} : X_n = y\}.$$

Suppose we have a path that hits the state y a finite number of times, say four times: T_y, T_y^2, T_y^3, T_y^4. Then via our convention, we say that $T_y^5 = \infty$.
Random variable:
Consider the number of hits at y (not counting time 0):

$$N_y := \sum_{n=1}^{\infty} 1\left(X_y = y\right)$$

to be the total number of hits to y at n after time 1. We consider that the possible values of this is $\{0, 1, 2, \ldots, \infty\}$, an infinite time horizon.
Equivalently, by definition or logic, we have:

$$(N_y = 0) = (T_y = \infty).$$

As another example, consider $(N_y \geq 1)$ is the complement of $(N_y = 0)$ because we include ∞ as a part of $(N_y \geq 1)$. Hence:

$$(N_y \geq 1) = (T_y < \infty).$$

Recall that $N_y := \sum_{n=1}^{\infty} 1(X_n = y)$, simply counting the number of hits on y.
Pitman asks the audience to explicitly find:

$$\begin{aligned}
(N_y \geq 3) &= (T_y^3 < \infty) \\
(N_y = 3) &= (T_y^3 < \infty, T_y^4 = \infty) \\
(N_y \geq k) &= (T_y^k < \infty) \\
(N_y = k) &= (T_y^k < \infty, T_y^{k+1} = \infty).
\end{aligned}$$

Now let's discuss the probabilities. Let

$$\mathbb{P}_y(T_y^k < \infty) = \rho_y^k,$$

where k on the RHS is a power, and k on the LHS is an index. Now taking $k = 1$, we have the definition of ρ_y:

$$\mathbb{P}_y(T_y < \infty) = \rho_y.$$

Now why is this true, as to get from ρ_y to ρ_y^2? Basically, this is by the SMP (Strong Markov Property).
Notice:

$$(T_y^k < \infty) = (N_y \geq k)$$

which tells us that the probability of hitting y at least k times is:

$$\mathbb{P}_y(N_y \geq k) = \rho_y^k \text{ for } k = 0, 1, 2, 3, \cdots = \mathbb{N}.$$

If we want to find the point probability that $N_y = k$, we take:

$$\begin{aligned}
\mathbb{P}_y(N_y = k) &= \mathbb{P}_y(N_y \geq k) - \mathbb{P}_y(N_y \geq k + 1) \\
&= \rho_y^k - \rho_y^{k+1} \\
&= \boxed{\rho_y^k(1 - \rho_y)},
\end{aligned}$$

which tells us that the probability distribution (starting at y) of $N_y := \sum_{n=1}^{\infty} 1(X_n = y)$ is Geometric with

$$p = 1 - \rho_y = \mathbb{P}_y(T_y = \infty) = \mathbb{P}_y(N_y = 0).$$

Pitman wants us to notice, using tail-sums:

$$\begin{aligned} \mathbb{E}_y N_y &= \sum_{k=1}^{\infty} \mathbb{P}_y(N_y \geq k) \\ &= \sum_{k=1}^{\infty} \rho_y^k = \frac{\rho_y}{1-\rho_y} = \frac{q}{p}, \end{aligned}$$

3 States of y

Now there are two cases Pitman wants us to consider.
(1) y is **transient** : $0 \leq \rho_y < 1$. This implies that our expected number of visits is:

$$\mathbb{E}_y N_y = \frac{\rho_y}{1-\rho_y} < \infty$$

which implies

$$\mathbb{P}_y(N_y < \infty) = 1,$$

which says that if we have a transient state, then we only return to y a finite number of times. In other words, after some point, the Markov chain never visits y again.

(2) y is **recurrent** : $\rho_y = 1$. In other words, $\mathbb{P}_y(N_y = \infty) = 1$ in that given any number of hits, we are sure to hit y again.

Break time.

Pitman gives us a formula if we really want to have a more concrete understanding of ρ_y:

$$\begin{aligned} \rho_y &= P(y,y) + \sum_{y_1 \neq y} P(y,y_1)P(y_1,y) + \sum_{y_1 \neq y} \sum_{y_2 \neq y} P(y,y_1)P(y_1,y_2)P(y_2,y) + \cdots \\ &= \mathbb{P}_y(T_y = 1) + \mathbb{P}_y(T_y = 2) + \mathbb{P}_y(T_y = 3) + \cdots \end{aligned}$$

4 Lemma 1.3

Take B to be a set of states.

4.1 Hypothesis:

Suppose the probability starting at x that $T_y \leq k$ is at least $\alpha > 0$ for some k and all x. In other words,

$$\mathbb{P}_x(T_y \leq k) \geq \alpha > 0$$

for some k and all x.
As an example of this hypothesis, consider the Gambler's ruin chain with state 0 and n absorbing states. That is, $B = \{0, N\}$, with $P(i, i+1) = p$ and $P(i, i-1) = q$. Then this condition holds with $k = \lceil \frac{N}{2} \rceil$. Then

$$\alpha = \max\{p^k, q^k\}.$$

4.2 Conclusion:

Suppose the hypothesis is satisfied. Then

$$\mathbb{P}_x(T_B > nk) \leq (1-\alpha)^n.$$

Proof. Consider the subset $U \subset \{0, 1, 2, \dots\}$. Surely, this is trivial for $n = 1$ in that $\mathbb{P}_y(T_B > k) \leq 1 - \alpha$ for all y.
Now,

$$\begin{aligned}\mathbb{P}_x(T_B > (n+1)k) &= \mathbb{P}_x(T_B > nk \text{ and after time } nk \text{ before time } (n+1)k \text{ still dont hit } B) \\ &= \sum_{y \notin B} \mathbb{P}_x(T_B > nk, X_{nk} = y, \text{ do not hit } B \text{ before time } (n+1)k) \\ &\leq \sum_{y \notin B} \mathbb{P}_x(T_B > nk, X_{nk} = y)(1-\alpha),\end{aligned}$$

as an upper bound.

□

Pitman mentions Kai Lai Chung of Stanford who paints this concept as the idea of a pedestrian crossing the road. This gives a geometric bound . If there is a certain chance of reaching a boundary state, eventually a Markov chain will hit such a boundary state.
Observe that in standard Real numbers,

$$0 \leq \mathbb{P}_x(T_B = \infty) \leq (1-\alpha)^n, \ \forall_n$$

implies

$$\mathbb{P}_x(T_B = \infty) = 0.$$

Further, notice:

$$(T_B = \alpha) \subseteq (T_B > nk),$$

$$\implies \mathbb{P}_x(T_B = \infty) \leq \mathbb{P}_x(T_B > nk) \leq (1-\alpha)^n.$$

Conclusion of Gambler's Ruin Example: Suppose $0 < p < 1$. In terms of transient and recurrent states, every state $x \notin \{0, N\}$ is **transient**! Moreover, $x \in \{0, N\}$ is recurrent.

Definition: Irreducible Matrix -

We say that a matrix P is **irreducible** if

$$\forall_{x,y \in S}, \exists_n : P^n(x,y) > 0$$

In words, for every pair of states x, y, it is possible to get from x to y in some number n of steps.
Note that n is a function of x, y. That is, $n = n(x, y)$.

Theorem 4.1. If matrix P is irreducible, then either:
- all states are recurrent
- all states are transient

With sloppy language, we then say either the matrix P is recurrent or transient, respectively

Notice that the Gambler's Ruin problem exhibits a matrix that is NOT irreducible, which can be seen via the definition above and the requirement that there exists some n where $P^n(x, y) > 0$.
An easy fact:
Suppose S is finite and P is irreducible. Then in this language, P is 'recurrent'.

Lecture ends here.

On Tuesday we will cover stationary distributions and stationary measures.