Stats 150, Fall 2019

Lecture 5, Thurs, 9/12/2019

Topics Today:

- §1.4 - 1.8 of text
- The remaining of readings are assigned outside of lecture.

1 Key Points for Homework

Pitman gives a few key pointers (which are from the textbook) that may help with finishing the homework due tonight.
(1) Recall the definition of an irreducible chain. That is,

$$\forall x, y \in S, \exists_n : P^n(x, y) > 0.$$

This forbids a random walk on a graph with 2 or more components (closed classes). Most of the chains we commonly deal with (and in our homework) are irreducible.
(2) Fact: (See text). If P is irreducible and if there is a stationary probability vector π for P (that is, we can solve $\pi P = \pi$ where $\sum_x \pi_x = 1, \pi_x \geq 0$), then all the states are recurrent (the chain is recurrent). (Theorem 1.7 in Durett).

Definition: Positive Recurrent -

We say that the chain is **positive recurrent** when, for some or for all x:

$$\mathbb{E}_x T_x < \infty$$

which is:

$$\mathbb{E}_x T_x = \sum_{n=1}^{\infty} \mathbb{P}_x(T_x \geq n).$$

We should check that if $E_x T_x < \infty$ for some x and P is irreducible, then

$$E_x T_x < \infty, \quad \forall_x.$$

Definition: Null recurrent -

If a state is recurrent but not positive recurrent (for example $P_x(T_x < \infty) = 1$ but $E_x T_x = \infty$), then we say that x is **null recurrent**.

2 Review

Pitman reminds us that there is a formula relating the mean return time and the stationary probability (Theorem 1.21 Durett):

$$\pi_x = \frac{1}{\mathbb{E}_x T_x}$$

As a simple corollary, this formula directly implies that π is unique. There is no doubt about this for a stationary measure in terms of the mean recurrence time. If we discuss a system of countably infinite space, our traditional linear algebra may fail. This result provides an interpretation beyond a system of finitely many equations and unknowns.

Remark: Conversely, if P is irreducible and positive recurrent, then there exists this π. This is almost trivial, but of course we have to check that π is a stationary probability.

3 Example of Null Recurrence

Example: Consider a simple (symmetric) random walk with equal probability of going either direction on $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$. We take the usual notation S_n for the walk.
Start at $x = 0$, so that $S_n := \Delta_1 + \Delta_2 + \cdots + \Delta_n$, where Δ_k has the value $+1$ with probability $\frac{1}{2}$ and -1 with probability $\frac{1}{2}$. Now this gives:

$$P^n(0,0) = \begin{cases} 0, & \text{if } n \text{ is odd} \\ \binom{2m}{m}\left(\frac{1}{2}\right)^{2m}, & \text{if } n = 2m \text{ is even} \end{cases}.$$

Now Pitman notes we can tell recurrence or transience by looking at the fact that the total number of visits to 0 follows a geometric distribution with $(1 - \rho_0)$:

$$\mathbb{E}_0(\text{total \# visits to } 0) = \sum_{n=1}^{\infty} P^n(0,0)$$

But we know that $\binom{2m}{m}\left(\frac{1}{2}\right)^{2m}$ is the same as the probability of m heads and m tails in $2m$ tosses. Increasing tosses gives a very 'flat' normal curve because the mean of $\mathbb{E}_0 S_{2m} = 0$. Now because the variance of each summed term is 1, the mean square is:

$$\mathbb{E}_0 S_{2m}^2 = \underbrace{1 + 1 + \cdots + 1}_{2m} = 2m.$$

We call this "diffusion", in that on average the center of our distribution goes no where, but the distribution spreads out and flattens.
Using Stirling's formula (or the Normal Approximation), that is,

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n},$$

and apply this to our earlier expression to show that:

$$P^{2m}(0,0) \sim \frac{C}{\sqrt{m}},$$

where C is some constant.
To see recurrence versus transience, we look at, from earlier,

$$\sum_{n=1}^{\infty} P^n(0,0) = \sum_{m=1}^{\infty} P^{2m}(0,0) \sim \sum_{n=1}^{\infty} \frac{C}{\sqrt{m}} = \infty$$

A rather paradoxical fact: This implies that the expected return time to 0 is infinite:

$$\mathbb{E}_0 T_0 = \infty,$$

although we are sure to (eventually) return with probability 1.
Recall that the definition of recurrent gives:

$$\mathbb{P}_x(T_x < \infty) = 1 \iff \mathbb{P}_x(T_x \geq n) \downarrow 0 \text{ as } n \uparrow \infty.$$

Also, we should know that positive recurrence implies recurrence, but not necessarily conversely.

Remark: Pitman summarizes that on our homework, we can quote the result that:
If we have a stationary measure, then the chain is **positive recurrent**.

4 Notion of x-blocks of a Markov chain

Start at x (for simplicity) or wait until we hit x. Then look the successive return times $T_x^{(i)}$ which is the ith copy of T_x. Now recall this has the Strong Markov Property, which gives us two things:
(1) Every $T_x^{(i)}$ has the same distribution as T_x.
(2) Further, they are independent copies. That is, $T_x^{(1)}, T_x^{(2)}, \ldots$ are independent.

Now Pitman mentions a variation on this theme of x-blocks, which explains many things:

Example: Let $N_{xy}^{(2)} :=$ the # of visits to y in the ith block of length T_x. In our previous in-class example, this gives a sequence:

$$2, 0, 6, 0, 4, 2, \ldots$$

Now for some book keeping, consider what happens if we sum over all states y. Of course, this just gives the length of $T_x^{(i)}$ by 'Accounting 101'.

$$\sum_{y \in S} N_{xy}^{(i)} = T_x^{(i)}.$$

Now this implies that there is a formula involving expectations. Take $\mathbb{E}_x$, the expectation starting at x:

$$\sum_{y \in S} \mathbb{E}_x N_{xy}^{(i)} = \mathbb{E}_x T_x^{(i)},$$

where this is really the same equation for all i by the Strong Markov Property. Fix x, y and look at $N_{xy}^{(1)}, N_{xy}^{(2)}, \ldots$, each of which:
(1) $N_{xy}^{(i)}$ has the same distribution as $N_{xy} := N_{xy}^{(1)}$.
(2) Further, the $N_{xy}^{(i)}$ are independent and identically distributed (iid).

Pitman reminds us that as we return to x, via the Strong Markov Property, nothing of the past changes our expectations or distributions going forward.

Break time.

5 Positive Recurrent Chains (P irreducible)

Notice that if $\mathbb{E}_x T_x < \infty$, and we define N_{xy} as we have earlier, then we can let:

$$\begin{aligned} \mu(x,y) &:= \mathbb{E}_x(N_{xy}) \\ \mu(x) &:= \mathbb{E}_x T_x = \text{ mean length of } x\text{-block} \end{aligned}$$

Correspondingly to our Accounting 101 from earlier, we write:

$$\sum_{y\in S} \mu(x,y) = \mu(x) < \infty.$$

Further, we can show (see text for details) that if we sum:

$$\sum_y \mu(x,y)P(y,z) = \mu(x,z),$$

or in other words, $\mu(x,\cdot)$ is a stationary measure (not a stationary probability, as it is an unnormalized measure). That is,

$$\mu(x,\cdot)P = \mu(x,\cdot).$$

This is important because it gives us a simple explicit construction of a stationary measure $\mu(x,\cdot)$ for every state x in state space S of a positive recurrent (PR) irreducible chain with matrix P. Notice that this is not just any measure. By convention, we say that the number of times we visit x in the duration of T_x is 1 (this is necessary to satisfy our constructions today). That is, we must not count a visit twice, and we must set:

$$\mu(x,x) := 1,$$

in order to get:

$$\sum_{y\in S} \mu(x,y) = \mu(x) < \infty.$$

Now to get a stationary probability measure, we take:

$$\pi(y) = \frac{\mu(x,y)}{\sum_z \mu(x,z)} = \frac{\mu(x,y)}{\mu(x)},$$

and this does NOT depend on x. We can take any reference state and we get the same thing when we look at these ratios.

5.1 Explanation of the Key Formula

We may ask why we have:

$$\sum_y \mu(x,y)P(y,z) = \mu(x,z). \tag{1}$$

Recall that $\mu(x,y)$ is the expected number of hits on y before T_x. That is,

$$\mu(x,y) = \mathbb{E}_x[\# \text{ of hits on } y \text{ before } T_x]$$

Now, every time we hit y, then $P(y,z)$ is the probability that the next step is to state z. Therefore (at least intuitively), $\mu(x,y)P(y,z)$ has a particular meaning. That is,

$$\mu(x,y)P(y,z) = \mathbb{E}_x\text{ (\# of transitions } y \to z \text{ before } (\leq) T_x \text{)}$$

The distribution of a single x-block gives the following formulas for the invariant probability measure π:

$$\pi(x) = \frac{1}{\mathbb{E}_x T_x}, \quad \frac{\pi(y)}{\pi(x)} = \mu(x,y)$$

6 Limit Theorems

If we let $N_n(y) := \sum_{k=1}^{n} 1(X_k = y) = \#$ of hits on y in first n steps , then:

$$\begin{aligned}\mathbb{E}_x \frac{N_n(y)}{n} &= \text{ mean \# hits on } y \text{ per unit time up to } n \\ &= \frac{1}{n}\sum_{k=1}^{n} P^k(x,y) \to \pi(y)\end{aligned}$$

We have this Cesàro mean convergence always for irreducible positive recurrent chains (these themselves do not converge, but their average converges). Now if we additionally impose aperiodicity, we have:

$$P^n(x,y) \to \pi(y),$$

always for irreducible and positive recurrent and aperiodic.

6.1 Review and Audience Questions:

A null recurrent chain has a stationary measure with reference state x assigned as measure 1. If we do this on a simple random walk, we find that the expectated probability of every state is 1, which explains why we expect to spend so much time to return back to x.
If we have a stationary probability measure:

$$\mathbb{E}_\pi \frac{1}{n}\sum_{k=1}^{n} P^k(x,y) = \pi(y),$$

we can argue that the stationary measure must be approached in the limit (and hence is unique as a limit must be unique).

Lecture ends here.