Lecture 6 notes

klin@math.arizona.edu

February 4, 2020

Today we continued proving the results in Sect. 1.3. Here is the list from last time:

1) ~~(Lemma 1.10)~~ $P_x(T_y^k < \infty) = \rho_{xy}\rho_{yy}^{k-1}$

2) (Lemma 1.3) Suppose there is $y \in S$, integer $k > 0$, and real $\alpha > 0$ such that for all $x \in S$

$$P_x(T_y \leqslant k) \geqslant \alpha. \tag{1}$$

   Then $P_x(T_y > nk) \leqslant (1-\alpha)^n$.

3) (Lemma 1.4) If $x \to y$ and $y \to z$, then $x \to z$.

4) (Theorem 1.5) If $x \to y$ and $\rho_{yx} < 1$, then x is transient.

5) (Lemma 1.6) If x is recurrent and $x \to y$, then $\rho_{yx} = 1$.

Last time we proved 1, 3–5. I left 2 as an exercise; I'll say more about it below.

Today our main goal was to prove

6) (Theorem 1.7) If C is closed, finite, and irreducible, then all states in C are recurrent.

7) (Theorem 1.8) If S is finite, then S can be written (uniquely) as a disjoint union of the form $T \cup R_1 \cup \cdots \cup R_k$, where T is the set of all transient states and the R_i are closed irreducible sets consisting of recurrent states.

As the text explains, 7 readily follows from 6. So we really need to prove 6. We introduced

8) (Lemma 1.9) If x is recurrent and $x \to y$, then y is recurrent.

9) (Lemma 1.10) Every finite closed set contains at least one recurrent state.

Using 8 and 9 we easily proved 6. So it remains to prove 8 and 9. To do so, we need

A) (Lemma 1.11) For all $x, y \in S$,

$$E_x N(y) = \frac{\rho_{xy}}{1-\rho_{yy}} \tag{2}$$

   where $N(y) = \#\{n > 0 \mid X_n = y\}$, and the above is interpreted as 0 if $\rho_{xy} = 0$ and ∞ if $\rho_{yy} = 1$.

B) (Lemma 1.12)

$$E_x N(y) = \sum_{n=1}^{\infty} p^n(x,y). \tag{3}$$

C) (Theorem 1.13) If $x \to y$, then y is recurrent if and only if

$$\sum_{n=1}^{\infty} p^n(x,y) = \infty. \tag{4}$$

Statements A and B imply C. So for those of you keeping score, this means we need to

- Prove A
- Prove B
- Use these to prove 8 and 9.

I'm doing this largely by following the book. A few notes:

1) Eq. (1.6) in the text can be proved more simply as follows: consider the sum

$$\begin{array}{llll} P(X=1) & +P(X=2) & +P(X=3) & +\cdots \\ & +P(X=2) & +P(X=3) & +\cdots \\ & & +P(X=3) & +\cdots \\ & & & +\cdots \end{array} \tag{5}$$

If we sum *vertically* first, i.e., along each column, this gives

$$\sum_{n=1}^{\infty} nP(X=n) = EX. \tag{6}$$

If we sum *horizontally* first, i.e., along each row, this gives

$$\sum_{n=1}^{\infty} P(X \geqslant n). \tag{7}$$

So

$$EX = \sum_{n=1}^{\infty} P(X \geqslant n). \tag{8}$$

(If any of you are worried about convergence issues and whether it's really legit to exchange the sums, this can be justified on the basis that the series – if it converges at all – converges absolutely. It is a theorem of real analysis that every rearrangement of an absolutely convergent series yields the same result.)

2) To really understand Theorem 1.8 (item 7 above), think about the 7-state chain in Example 1.14. What happens if we set $p(1,2) = \varepsilon > 0$? How many closed irreducible components does the chain have now? Are there any transient states?

3) About Lemma 1.3: I had said the proof copies that of Example 1.13, which I discussed in class. Here is how you can show Lemma 1.3, for general k and $n=2$:

$$P_x(T_y > 2k) = P_x(X_1 \neq y, \cdots, X_{2k} \neq y) \tag{9a}$$

$$= \sum_{i \neq y} P_x(\underbrace{X_1 \neq y, \cdots, X_{k-1} \neq y, X_k = i}_{A}, \underbrace{X_{k+1} \neq y, \cdots, X_{2k} \neq y}_{B}) \tag{9b}$$

$$= \sum_{i \neq y} P_x(\underbrace{X_{k+1} \neq y, \cdots, X_{2k} \neq y}_{B} \mid \underbrace{X_1 \neq y, \cdots, X_{k-1} \neq y, X_k = i}_{A}) \cdot P_x(\underbrace{X_1 \neq y, \cdots, X_{k-1} \neq y, X_k = i}_{A}) \tag{9c}$$

$$= \sum_{i \neq y} P_i(X_{k+1} \neq y, \cdots, X_{2k} \neq y) \cdot P_x(X_1 \neq y, \cdots, X_{k-1} \neq y, X_k = i) \tag{9d}$$

$$\leqslant (1-\alpha) \sum_{i \neq y} P_x(X_1 \neq y, \cdots, X_{k-1} \neq y, X_k = i) \tag{9e}$$

$$\leqslant (1-\alpha)^2. \tag{9f}$$

From line (a) to line (b), we just expressed the probability as a sum over possible values of X_k. Line (c) is obtained by re-expressing each term using conditional probabilities, i.e., writing $P(A \cap B) = P(B|A) \cdot P(A)$. Line (d) is obtained from line (c) by applying the Markov property to the factor A. Line (e) is obtained from (d) by recognizing that the first factor in (d) is a probability of the form $P_x(T_y > k)$, which is bounded above by $1-\alpha$. And again from line (e) to (f).