Lecture 7. More about discrete Markov chains.

Considering periodicity is simpler than the rest, because it doesn't have anything to do with the limits or infinite series. But first of all we solve the following problem:

$\boxed{7}$ $f_{yy}^{(k)} = 0$ for all k that are not multiples of d if and only if $p_{yy}^{(n)} = 0$ for all n that are not multiples of d.

We defined periodicity of a state through the first-return probabilities $f_{yy}^{(k)}$; this problem shows that we can define it through the transition probabilities $p_{yy}^{(n)}$.

Theorem 7.1. *If one state y belonging to an equivalence class C_i is periodic with period d, then all other states of this class are periodic with period d; if one state is aperiodic, then all other states of the same class are aperiodic too.*

Proof. Suppose we can return from y to y only in a number of steps that is a multiple of d; and suppose that we can return from x to x in a number of steps N that is *not* a multiple of d. We can travel from y to x in K steps: $p_{yx}^{(K)} > 0$, and return to y from x in M steps: $p_{xy}^{(M)} > 0$. Then we can return from y to y in $K + M$ steps: $p_{yy}^{(K+M)} \geq p_{yx}^{(K)} \cdot p_{xy}^{(M)} > 0$. As we assumed, $K + M$ is a multiple of d. But we can also return from y to the same point in $K + N + M$ steps: first going from y to x in K steps, then from x to x in N, and finally from x to y in M steps. But since N is not a multiple of d, the sum $K + N + M$ also isn't one. We get a contradiction, whose source is that we supposed that x and y had different periods. So this is impossible.

Theorem 7.2. *The states belonging to the same equivalence class C_i are either all transient, or all recurrent. If $\lim_{n\to\infty} p_{yy}^{(n)} = 0$ for some $y \in C_i$, then $\lim_{n\to\infty} p_{xx}^{(n)} = 0$ for all states x from the same equivalence class C_i. If the limit $\lim_{n\to\infty} p_{yy}^{(n)} = p_y > 0$, then $\lim_{n\to\infty} p_{xy}^{(n)} = p_y$ for all x of the same equivalence class, and $\lim_{n\to\infty} p_{xx}^{(n)} = p_x$ exists and is positive.*

Proof. Let $x \sim y$. If we have $p_{yx}^{(N)} > 0$, $p_{xy}^{M} > 0$, then for $n \geq N + M$ we have:

$$p_{yy}^{(n)} = P_y\{\xi_n = y\} \geq P_y\{\xi_N = x,\, \xi_n = y\} = p_{yx}^{(N)} \cdot p_{xy}^{(n-N)} = \text{const} \cdot p_{xy}^{(n-N)}, \tag{7.1}$$

$$p_{yy}^{(n)} \geq P_y\{\xi_{n-M} = x,\, \xi_n = y\} = p_{yx}^{(n-M)} \cdot p_{xy}^{(M)} = \text{const} \cdot p_{yx}^{(n-M)}, \tag{7.2}$$

$$p_{yy}^{(n)} \geq p_{yx}^{(N)} \cdot p_{xx}^{(n-N-M)} \cdot p_{xy}^{(M)} = \text{const} \cdot p_{xx}^{(n-N-M)} \tag{7.3}$$

with positive constants. So if the series $\sum_{n=0}^{\infty} p_{yy}^{(n)}$ converges, so do the series $\sum_{n=N}^{\infty} p_{xy}^{(n)}$, $\sum_{n=M}^{\infty} p_{yx}^{(n)}$, $\sum_{n=N+M}^{\infty} p_{xx}^{(n)}$, which is equivalent to convergence of $\sum_{n=0}^{\infty} p_{xy}^{(n)}$, $\sum_{n=0}^{\infty} p_{yx}^{(n)}$, and $\sum_{n=0}^{\infty} p_{xx}^{(n)}$.

Also if $\lim_{n\to\infty} p_{yy}^{(n)} = 0$, also are equal to 0 the limits $\lim_{n\to\infty} p_{xy}^{(n)}$, $\lim_{n\to\infty} p_{yx}^{(n)}$, and $\lim_{n\to\infty} p_{xx}^{(n)} = 0$; and if $\lim_{n\to\infty} p_{yy}^{(n)} > 0$, the limit $\lim_{n\to\infty} p_{xx}^{(n)}$ cannot be equal to 0.

Now suppose $\lim_{n\to\infty} p_{yy}^{(n)} = p_y > 0$. Then the state y (and with it all other states x in the equivalence class C_i) cannot be transient, and it cannot be periodic. By Theorem 6.2 we have $P_x\{\tau_y < \infty\} = 1$, and by Theorem 6.1 we have $\lim_{n\to\infty} p_{xy}^{(n)} = p_y$.

Is our Markov chain, if we consider it on a recurrent aperiodic class C_i, necessarily ergodic? That is: do the limiting probabilities $p_y = \lim_{n\to\infty} p_{xy}^{(n)}$ form a probability distribution: is $\sum_{y\in C_i} p_y = 1$?

In the case of a *finite* C_i, yes: $\sum_{y\in C_i} p_y = \sum_{y\in C_i} \lim_{n\to\infty} p_{xy}^{(n)} = \lim_{n\to\infty} \sum_{y\in C_i} p_{xy}^{(n)}$ $= \lim_{n\to\infty} 1 = 1$. In the case of (countably) infinite class, it is possible that $p_y = \lim_{n\to\infty} p_{xy}^{(n)} = 0$ for all $y \in C_i$ (as in the case of the simplest random walk – which is, though, a periodic chain with period 2). But if all $p_y > 0$, is their sum necessarily equal to 1? can it be equal, say, to 2/3 or to 3/2?

Theorem 7.3. *For the limits* $p_y = \lim_{n\to\infty} p_{xy}^{(n)}$ (supposing these limits exist), *we have*: $\sum_{y\in C_i} p_y \le 1$.

Proof. Let us number the points of the class C_i: $x_1, x_2, \ldots, x_j, \ldots$ (we don't have to consider the case of C_i being finite). Let x_0 be a fixed state belonging to the class C_i (x_0 is equal to one of $x_1, x_2, \ldots, x_j, \ldots$). We have:

$$\sum_{j=1}^{m} p_{x_j} = \sum_{j=1}^{m} \lim_{n\to\infty} p_{x_0 x_j}^{(n)} = \lim_{n\to\infty} \sum_{j=1}^{m} p_{x_0 x_j}^{(n)} \le \lim_{n\to\infty} \sum_{y\in C_i} p_{x_0 y}^{(n)} = 1, \tag{7.4}$$

$$\sum_{y\in C_i} p_y = \lim_{m\to\infty} \sum_{j=1}^{m} p_{x_j} \le 1. \tag{7.5}$$

Theorem 7.4. *For all* $y \in C_i$

$$\sum_{x\in C_i} p_x \cdot p_{xy} = p_y \tag{7.6}$$

(in matrix notations: $\boldsymbol{p} \cdot P = \boldsymbol{p}$, where $\boldsymbol{p}$ is the row vector with components p_y, $y \in C_i$, and $P = (p_{xy})_{x,\, y\in C_i}$ is the one-step transition matrix restricted to the class C_i).

Proof. Take a finite sum first:

$$\begin{aligned}\sum_{j=1}^{m} p_{x_j} \cdot p_{x_j y} = \sum_{j=1}^{m} \lim_{n\to\infty} p_{x_0 x_j}^{(n)} \cdot p_{x_j y} &= \lim_{n\to\infty} \sum_{j=1}^{m} p_{x_0 x_j}^{(n)} \cdot p_{x_j y} \\ &\le \lim_{n\to\infty} \sum_{x\in C_i} p_{x_0 x}^{(n)} \cdot p_{xy} = \lim_{n\to\infty} p_{x_0 y}^{(n+1)} = p_y.\end{aligned} \tag{7.7}$$

Taking $m \to \infty$, we get:

$$\sum_{x\in C_i} p_x \cdot p_{xy} = \lim_{m\to\infty} \sum_{j=1}^{m} p_x \cdot p_{xy} \le p_y. \tag{7.8}$$

Now, suppose that a strict inequality holds in (7.8) for some $y = y_0$:

$$\sum_{x \in C_i} p_x \cdot p_{xy_0} < p_{y_0}. \tag{7.9}$$

Let us take the sum of both sides in (7.8) over all $y \in C_i$. Since all left-hand sides are not greater than the corresponding right-hand sides, and at least one is smaller, we have:

$$\sum_{y \in C_i} \sum_{x \in C_i} p_x \cdot p_{xy} < \sum_{y \in C_i} p_y. \tag{7.10}$$

Of course, when we see an iterated (a double) sum (and the summands are nonnegative), we should reverse the order of summation:

$$\sum_{x \in C_i} \sum_{y \in C_i} p_x \cdot p_{xy} < \sum_{y \in C_i} p_y. \tag{7.11}$$

The sum in the left-hand side is equal to

$$\sum_{x \in C_i} p_x \sum_{y \in C_i} p_{xy} = \sum_{x \in C_i} p_x \cdot 1 = \sum_{x \in C_i} p_x. \tag{7.12}$$

So we get $\sum_{x \in C_i} p_x < \sum_{y \in C_i} p_y$: a contradiction. Its source is our assuming (7.9); so it's impossible, and equality (7.6) holds for all y.

Theorem 7.5. *If* $p_y = \lim_{n\to\infty} p_{xy}^{(n)} > 0$, *then* $\sum_{y \in C_i} p_y = 1$: *ergodicity.*

Proof. By Theorem 7.4, we have $\boldsymbol{p} \cdot P = \boldsymbol{p}$. Multiplying both parts by P on the right, we get: $\boldsymbol{p} \cdot P^2 = \boldsymbol{p} \cdot P = \boldsymbol{p}$. Multiplying again and again by P, we get $\boldsymbol{p} \cdot P^n = \boldsymbol{p}$ for every n. In components, it is written as

$$\sum_{x \in C_i} p_x \cdot p_{xy}^{(n)} = p_y. \tag{7.13}$$

Now we take $n \to \infty$. The x-th term of the sum converges to $p_x \cdot p_y$ as $n \to \infty$; but does the sum converge to $\sum_{x \in C_i} p_x \cdot p_y$? (This is an *infinite* sum, and the limit passage needs justification.) We are using the dominated-convergence theorem (remember, a sum is a Lebesgue integral with respect to the counting measure #): the summands $p_x \cdot p_{xy}^{(n)}$ are, for every n, dominated by the function p_x, and $\sum_{x \in C_i} p_x \le 1$ by Theorem 7.3. So we have

$$p_y = \lim_{n\to\infty} \sum_{x \in C_i} p_x \cdot p_{xy}^{(n)} = \sum_{x \in C_i} \lim_{n\to\infty} p_x \cdot p_{xy}^{(n)} = \sum_{x \in C_i} p_x \cdot p_y. \tag{7.14}$$

Since $p_y > 0$, we can divide both sides by p_y and get: $\sum_{x \in C_i} p_x = 1$.

For an aperiodic class C_i with $p_y > 0$, $y \in C_i$, we have:

$$\sum_{y \in C_i} \frac{1}{E_y \tau_y} = 1; \tag{7.15}$$

and $\lim_{n\to\infty} p_{xy}^{(n)} = p_y$ for every $x \sim y$.

What can we say about $\lim_{n\to\infty} p_{xy}^{(n)}$ for $x \not\sim y$? This means that either we can go from y to x, but not back; or that $x \to\to y$, but we cannot go from y to x; or that we can neither do from y to x, nor from x to y.

In the first and in the last case clearly $p_{xy}^{(n)} = 0$, $\lim_{n\to\infty} p_{xy}^{(n)} = 0$; but the second case is possible – if the state x is transient. In this case, as we know (Theorem 6.1), $\lim_{n\to\infty} p_{xy}^{(n)} = P_x\{\tau_y < \infty\} \cdot p_y$; but how to find $P_x\{\tau_y < \infty\}$??

But it seems we have come closer to the end of our considering discrete Markov chains; let me drop this question for the time being.

I want to say a couple of words about the *periodic* case.

In this case we can consider on the equivalence class C_i the aperiodic Markov chain $\tilde{\xi}_n = \xi_{nd}$ with one-step transition matrix $\tilde{P} = P^d$ and $\tilde{p}_{xy}^{(n)} = p_{xy}^{(nd)}$. We saw (Theorem 5.3) that $\lim_{n\to\infty} p_{yy}^{(nd)} = 1/E_y\tilde{\tau}_y = d/E_y\tau_y$, and $p_{yy}^{(k)} = 0$ for k not being multiples of d.

We can subdivide the equivalence class C_i into d subclasses in the following way. We introduce a new equivalence relation $\sim_d$: $x \sim_d y$ if we can go from x to y and from y to x in some number of steps that is a multiple of d: $x \sim_d y$ if there exists an integer n such that $p_{xy}^{(nd)} > 0$, and another integer m such that $p_{yx}^{(md)} > 0$. For some $x_0 \in C_i$ we introduce the class of $\sim_d$-equivalence $C_{i0} \subset C_i$ containing x_0: $C_{i0} = \{x\colon x \sim_d x_0\}$.

Then we consider the subclass C_{i1} consisting of all states $z \in C_i$ such that the one-step transition probability $p_{xz} > 0$ for some $x \in C_{i0}$. We can reach a point $z \in C_{i1}$ from the point x_0 (and from every other point $x \in C_{i0}$) only in number of steps n that has the form $n = 1$, or $n = d + 1$, or $n = 2d + 1$, ...: all n with remainder 1 after division by d. Then we introduce $C_{i2} = \{z \in C_i\colon p_{x_0 z}^{(n)} = 0$ for $n = 2, d + 2, 2d + 2, 3d + 2, ...\}$, etc., and finally $C_{i,\,d-1} = \{z \in C_i\colon p_{x_0 z}^{(n)} = 0$ except for $n = d - 1, 2d - 1, 3d - 1, ...\}$.

It is easy to check that $C_i = \bigcup_{j=0}^{d-1} C_{ij}$, that the subclasses C_{ij} are disjoint for different j, and that every x and y belonging to the same subclass are $\sim_d$-equivalent. If we start from a state ξ_0 belonging to a subclass C_{ij}, our next state ξ_1 is almost surely in the subclass $C_{i,\,j+1}$ (if $j \ne d - 1$; if $j = d - 1$, it is in the subclass C_{i0}); the next state ξ_2 is in $C_{i,\,j+2}$ or $C_{i,\,j+2-d}$; etc.

For $x \in C_{ij}$, $y \in C_{ik}$ we have $\lim_{n\to\infty} p_{xy}^{(nd+m)} = d/E_y\tau_d$ if $m + j - k$ is divisible by d, and $p_{xy}^{(nd+m)} = 0$ if it is not.

The sum $\sum_{y\in C_{ij}} d/E_y\tau_y = 1$, so $\sum_{y\in C_{ij}} 1/E_y\tau_y = 1/d$, and $\sum_{y\in C_i} 1/E_y\tau_y = 1$.