Lecture 19 Mon Apr/01/2024

Last time
- Stationary measures
- Existence
- Uniqueness

Today
- Aperiodicity
- Convergence Theorem

Aperiodicity

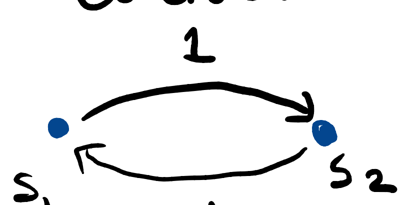
Our goal today is to understand the asymtotic distribution of X_n,
$$\lim_{n\to\infty} \mathbb{P}_x(X_n = y)$$

If this yields a probability dist, we can run Markov Chains for a while to sample!

Notice that if y is transcient $\Rightarrow$ this limit is zero.

A natural question is when is that the limit exists?

Example: Consider the chain



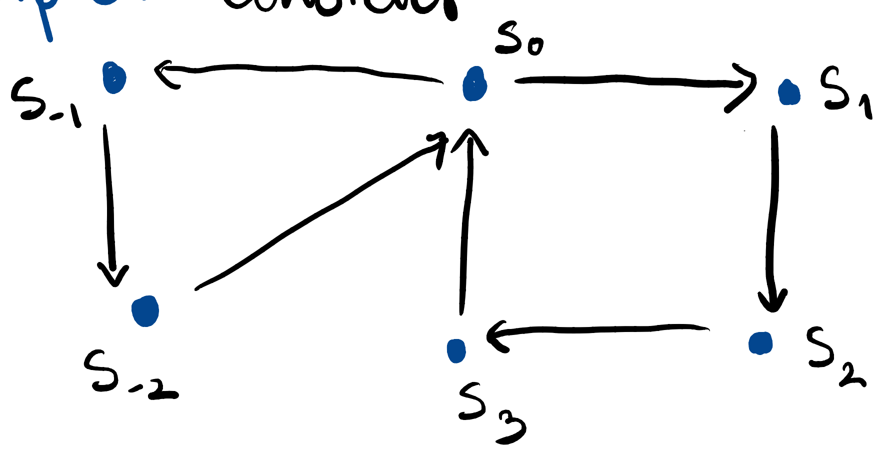
Then,
$$1 = \mathbb{P}_{s_1}(X_{2n} = s_1) \neq \mathbb{P}(X_{2n+1} = s_1) = 0.$$
$\dashv$

We shall see that this periodic behavior is the only thing preventing convergence.

Def: For any recurrent $x \in S$, the period of x, called d_x, is the greatest common divisor of $I_x = \{ n \geq 1 : p^{(n)}(x,x) > 0 \}$. $\dashv$

The previous example has a period of two.

Example: Consider



There are two cycles including s_0:

$$s_0 \to s_1 \to s_2 \to s_3 \to s_0$$

$$s_0 \to s_{-1} \to s_{-2} \to s_0$$

Thus, $I_{s_0} = \{ 3n, 4n \}$, therefore $d_{s_0} = 1$. $\dashv$

We say that a chain is aperidic if $d_x = 1 \; \forall x$. In turn, aperiodicity

Holds for everyone in an irreducible class.

is a "class property."

Lemma (⊠): If $P_{xy} > 0 \Rightarrow d_y = d_x$.

Left as exercise.

Lemma (ϟ): If $d_x = 1$, then, $p^{(m)}(x,x) > 0$ for $m \geq m_0$.

Proof: We will use two claims:

Claim: If $\exists m$ s.t. $m \in I_x$, $m+1 \in I_x$, then the result follows.

Fact: If $\gcd(I_x) = 1$, then $\exists i_1, \ldots, i_k \in I_x$ and $c_1, \ldots, c_k \in \mathbb{Z}$ s.t. $\sum_{\ell=1}^{k} c_\ell i_\ell = 1$.

Fact from number theory that we will not prove.

Let's show that these two imply the result. Let $a_\ell = c_\ell^+$ and $b_\ell = c_\ell^-$ |max(0,c)|

then

$$\underbrace{a_1 i_1 + \ldots + a_k i_k}_{m+1 \in I_x} = \underbrace{b_1 i_1 + \ldots + b_k i_k}_{m \in I_x} + 1.$$

Then, the result follows by the claim. □

Convergence Theorem

We are now ready to prove the main result today

Theorem: Suppose that a MC with transition prob. p is irreducible and aperiodic, and has a stationary distribution π. Then, for any $x \in S$

$$\mathbb{P}_x(X_n = \cdot) \xrightarrow{w} \pi.$$

Proof: Let $S^2 = S \times S$ and define the chain given by

$$\bar{p}((x_1, y_1), (x_2, y_2)) = p(x_1, x_2)\, p(y_1, y_2).$$

▷ First, we note that $\bar{p}$ is irreducible. To see this, note that since p is irreducible $\Rightarrow \exists$ K and L s.t. $p^{(K)}(x_1, x_2) > 0$ and $p^{(L)}(y_1, y_2) > 0$. By Lemma (3), $\exists M$ s.t. $p^{(L+M)}(x_2, x_2) > 0$ and $p^{(K+M)}(y_2, y_2) > 0$

$$\Rightarrow \bar{p}^{(K+L+M)}((x_1, y_1), (x_2, y_2)) \geq p^{(K)}(x_1, x_2)\, p^{(L+M)}(x_2, x_2)\, p^{(L)}(y_1, y_2)\, p^{(K+M)}(y_2, y_2) > 0.$$

▷ Second, we note that $\bar{\pi}((a,b)) = \pi(a)\pi(b)$ defines a stationary distribution (since both components are ind.) and moreover $\bar{p}$ makes all states S^2 recurrent. This follows from the following

Lemma: If there is a stationary distribution, then all states y s.t. $\pi(y) > 0$ are recurrent.

Proof of Lemma: Note that

$$\infty = \sum_{n=1}^{\infty} \pi(y) \overset{\text{stationarity}}{=} \sum_{n=1}^{\infty} \sum_{x} \pi(x) p^{(n)}(x,y)$$

Fubini's $\rightarrow$
$$= \sum_{x} \pi(x) \sum_{n=0}^{\infty} p^{(n)}(x,y)$$

Formula for $\mathbb{E}_x N(y)$
$$= \sum_{x} \pi(x) \frac{\rho_{xy}}{1-\rho_{yy}}$$

π is a dist and $\rho_{xy} \leq 1$.
$$\leq \frac{1}{1-\rho_{yy}}$$

So we conclude that $\rho_{yy} = 1$. □

▷ Let $(X_n, Y_n) \sim \bar{p}$, let $T = \inf\{n \geq 1 \mid X_n = Y_n\}$. Note that for any fixed x we have

$T_x = \inf\{ n \geq 1 \mid X_n = Y_n = x\} < \infty$ a.s. since (x,x) is recurrent.

Claim: On $\{T \leq n\}$, X_n and Y_n have the same distribution.

Proof of the claim:

$$\mathbb{P}(X_n = y, T \leq n) = \sum_{m=1}^{n} \sum_{x} \mathbb{P}(T = m, X_m = x, X_n = y)$$

$$\overset{\text{Markov}}{=} \sum_{m=1}^{n} \sum_{x} \mathbb{P}(T = m, X_m = x)\, \mathbb{P}(X_n = y \mid X_m = x)$$

$$= \sum_{m=1}^{n} \sum_{x} \mathbb{P}(T = m, Y_m = x)\, \mathbb{P}(Y_n = y \mid Y_n = x)$$

$$= \mathbb{P}(Y_n = y, T \leq n). \qquad \square$$

Observe that

$$\mathbb{P}(X_n = y) = \mathbb{P}(Y_n = y, T \leq n) + \mathbb{P}(X_n = y, T > n)$$
$$\leq \mathbb{P}(Y_n = y) + \mathbb{P}(X_n = y, T > n)$$

Inverting the role of X_n and Y_n gives

$$|\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| \leq \mathbb{P}(X_n = y, T > n) + \mathbb{P}(Y_n = y, T > n)$$

Summing over y we get

(b) $$\sum_{y} |\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| \leq 2\, \mathbb{P}(T > n).$$

This holds regardless of how we initialize X_0 and Y_0. Assume $X_0 = x$ and $Y_0 \sim \pi$. Then, $(\heartsuit)$ gives

$$\| \mathbb{P}_x(X_n = \cdot) - \pi(\cdot) \|_{TV} = \mathbb{P}(T > n) \to 0.$$

← Recall the TV-dist from Lecture 4

Then, the result follow from the Lemma in page 4 of Lecture 4. $\square$