Lecture 12:

Proof of the Convergence Theorem II

Assumptions:

I: Irreducible

A: Aperiodic

R: Recurrent

S: existence of a stationary distribution $\vec{\pi}$

Recall Lemma 8.3. If x is aperiodic, then there is an n_0 such that all $n \geq n_0$ are in I_x.

Theorem 9.1. (Convergence Theorem) If I, A & S hold, then $\lim_{n\to\infty} [P^n]_{xy} = \pi_y$, $\forall x, y \in \mathcal{X}$.

Proof Idea: $X_0 = x$, $\mathbb{P}(X_n = y \mid X_0 = x) = [P^n]_{xy}$

$Y_0 \sim \vec{\pi}$, $\mathbb{P}(Y_n = y \mid Y_0 \sim \vec{\pi}) = \pi_y$.

If $|\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)| \xrightarrow{n\to\infty} 0$, the proof is done.

Proof: Let $\mathcal{X}$ be the state space. Define a new transition matrix Q on $\mathcal{X} \times \mathcal{X}$ by

$$Q_{(x_1,y_1),(x_2,y_2)} = P_{x_1,x_2} P_{y_1,y_2}, \quad \forall (x_1,y_1), (x_2,y_2) \in \mathcal{X} \times \mathcal{X}.$$

This implies. $[Q^n]_{(x_1,y_1),(x_2,y_2)} = [P^n]_{x_1,x_2} \cdot [P^n]_{y_1,y_2}$.

① Claim 1. Q is irreducible.

Pf. $\forall (x_1,y_1),(x_2,y_2)\in \mathcal{X}\times\mathcal{X}$, $\exists\, K, L \geq 1$, s.t. $[P^K]_{x_1,x_2} > 0$ and $[P^L]_{y_1,y_2} > 0$. From Lemma 8.3, $\exists\, M \in \mathbb{N}$, s.t. $[P^{L+M}]_{x_2,x_2} > 0$ and $[P^{K+M}]_{y_2,y_2} > 0$.

Thus, $[Q^{K+L+M}]_{(x_1,y_1),(x_2,y_2)} = [P^{K+L+M}]_{x_1,x_2} \cdot [P^{K+L+M}]_{y_1,y_2}$

$$\geq [P^K]_{x_1,x_2} [P^{L+M}]_{x_2,x_2} \cdot [P^L]_{y_1,y_2} [P^{K+M}]_{y_2,y_2} > 0. \quad \square$$

② Claim 2. Q is recurrent.

Pf. Define $\vec{\vec{\pi}}_{(x,y)} := \vec{\pi}_x \cdot \vec{\pi}_y$. Then $\forall (x_1,y_1),(x_2,y_2)\in\mathcal{X}\times\mathcal{X}$,

$$\sum_{(x_1,y_1)\in\mathcal{X}\times\mathcal{X}} \vec{\vec{\pi}}_{(x_1,y_1)} Q_{(x_1,y_1),(x_2,y_2)} = \sum_{x_1\in\mathcal{X}}\sum_{y_1\in\mathcal{X}} \vec{\pi}_{x_1}\cdot\vec{\pi}_{y_1}\cdot P_{x_1,x_2} P_{y_1,y_2}$$

$$= \Big(\sum_{x_1\in\mathcal{X}} \vec{\pi}_{x_1}\cdot P_{x_1,x_2}\Big)\cdot\Big(\sum_{y_1\in\mathcal{X}} \vec{\pi}_{y_1}\cdot P_{y_1,y_2}\Big) = [\vec{\pi}P]_{x_2}\cdot[\vec{\pi}P]_{y_2}$$

$$= \vec{\pi}_{x_2}\cdot\vec{\pi}_{y_2} = \vec{\vec{\pi}}_{(x_2,y_2)}.$$

Thus, $\vec{\vec{\pi}}$ is a stationary distribution of Q.

why? Lemma 11.1 implies that there exists at least one recurrent state. Since Q is irreducible, it is recurrent. $\square$

time X_n first meets Y_n

time of first visit to (x,x)

$\boxed{\text{Cor 1. If } x \to y \text{ and } x \text{ is recurrent, then } \rho_{yx} = 1.}$

③. Let $(X_n, Y_n)_{n\in\mathbb{N}}$ denote the chain on $\mathcal{X}\times\mathcal{X}$ with transition matrix Q and let

$$T := \min\{n \geq 0 : X_n = Y_n\}$$

and

$$V_{xx} := \min\{n \geq 0 : X_n = Y_n = x\}.$$

Since Q is irreducible, $\rho_{(x,x),(x',y')} > 0$, $\forall x', y' \in \mathcal{X}$.

Since Q is recurrent, (x,x) is recurrent.

By Corollary 1 in Lecture 7, $\rho_{(x',y'),(x,x)} = 1$.

That is, $\mathbb{P}_{(x',y')}(V_{xx} < \infty) = 1$, $\forall x', y' \in \mathcal{X}$.

Since $T \leq V_{xx}$, one has $\mathbb{P}(T < \infty) = 1$. $\square$

④. For any $y \in \mathcal{X}$ and $n \geq 1$,

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

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

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

$$= \sum_{m=1}^{n} \sum_{x\in\mathcal{X}} \mathbb{P}(X_n = y \mid T = m, X_m = x) \cdot \mathbb{P}(T = m, X_m = x)$$

$$= \sum_{m=1}^{n} \sum_{x \in \mathcal{X}} \mathbb{P}(X_n = y \mid X_m = x) \cdot \mathbb{P}(T = m, X_m = x)$$

$$= \sum_{m=1}^{n} \sum_{x \in \mathcal{X}} [P^{n-m}]_{xy} \, \mathbb{P}(T = m, Y_m = x)$$

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

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

Thus, $|\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)|$

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

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

$$\leq \mathbb{P}(X_n = y, \ T > n) + \mathbb{P}(Y_n = y, \ T > n), \ \forall y \in \mathcal{X}, n \geq 1.$$

This implies, $\sum_{y \in \mathcal{X}} |\mathbb{P}(X_n = y) - \mathbb{P}(Y_n = y)|$

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

$$= 2\,\mathbb{P}(T > n) .$$

Let $X_0 = x$ and $Y_0 \sim \vec{\pi}$, then $\forall y \in \mathcal{X}$,

$$0 \leq |[P^n]_{xy} - \vec{\pi}_y| \leq \sum_{z \in \mathcal{X}} |[P^n]_{xz} - \vec{\pi}_z|$$

$$= \sum_{z \in \mathcal{X}} |\mathbb{P}(X_n = z) - \mathbb{P}(Y_n = z)|$$

$$\le 2\,\mathbb{P}(T>n).$$

Since $\lim_{n\to\infty} 2\,\mathbb{P}(T>n) = 2\,\mathbb{P}(T=\infty) = 2\cdot(1-\mathbb{P}(T<\infty)) = 0$, the squeeze theorem implies

$$\lim_{n\to\infty} \left|[P^n]_{xy} - \vec{\pi}_y\right| = 0.$$

That is, $[P^n]_{xy} \xrightarrow{n\to\infty} \vec{\pi}_y$, $\forall\, x,y \in \mathcal{X}$. ■

This is the end of this lecture!

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)