Lecture 11:

Proof of the Convergence Theorem I, Uniqueness

Assumptions:
I: Irreducible
A: Aperiodic
R: Recurrent
S: existence of a stationary distribution $\vec{\pi}$

Lemma 11.1. If S holds, then all states y that have $\vec{\pi}_y > 0$ are recurrent.

Proof. From Lemma 5 & Lemma 6 in Lecture 7, we know two representations of $\mathbb{E}_x N(y)$:

$$\mathbb{E}_x N(y) = \sum_{n=1}^{\infty} [P^n]_{xy} = \begin{cases} 0, & \rho_{xy} = 0; \\ \dfrac{\rho_{xy}}{1-\rho_{yy}}, & \rho_{xy} > 0. \end{cases}$$

On the one hand, since $\pi_y > 0$, one has

$$\sum_{x \in \mathcal{X}} \vec{\pi}_x \mathbb{E}_x N(y) = \sum_{x \in \mathcal{X}} \sum_{n=1}^{\infty} \vec{\pi}_x [P^n]_{xy} = \sum_{n=1}^{\infty} \sum_{x \in \mathcal{X}} \vec{\pi}_x [P^n]_{xy}$$

$$= \sum_{n=1}^{\infty} [\vec{\pi} P^n]_y = \sum_{n=1}^{\infty} \vec{\pi}_y = \infty \cdot \vec{\pi}_y = \infty.$$

On the other hand, we have

$$\sum_{x \in \mathcal{X}} \vec{\pi}_x \mathbb{E}_x N(y) \le \sum_{x \in \mathcal{X}} \vec{\pi}_x \cdot \frac{1}{1-\rho_{yy}} = \frac{1}{1-\rho_{yy}}.$$

Therefore, $\dfrac{1}{1-\rho_{yy}} = \infty$ and thus $\rho_{yy} = 1$ ■.

Remark 11.1. If I & S hold, then R also holds.

Proof. Since $\sum_{x\in\mathcal{X}} \vec{\pi}_x = 1$, there exists $y \in \mathcal{X}$ with $\vec{\pi}_y > 0$.

why? Lemma 11.1 says y is recurrent. Thus, R holds.

Theorem 9.3. If I & S hold, then $\vec{\pi}_y = \dfrac{1}{\mathbb{E}_y \tau_y}$, $\forall y \in \mathcal{X}$.

Proof. Suppose $X_0 \sim \vec{\pi}$. From Theorem 9.2, it follows that

$$\frac{N_n(y)}{n} \xrightarrow{a.s.} \frac{1}{\mathbb{E}_y \tau_y}, \quad \forall y \in \mathcal{X}.$$

Taking expected values at both sides, one has

$$\mathbb{E}_\pi \mathbb{E}_x \frac{N_n(y)}{n} \xrightarrow{n\to\infty} \frac{1}{\mathbb{E}_y \tau_y}, \quad \forall y \in \mathcal{X}.$$

Notice that $N_n(y) = \sum_{i=1}^{n} \mathbb{1}_{\{X_i = y\}}$, and

$$\mathbb{E}_x N_n(y) = \sum_{i=1}^{n} \mathbb{E}_x \mathbb{1}_{\{X_i=y\}} = \sum_{i=1}^{n} \mathbb{P}_x(X_i = y) = \sum_{i=1}^{n} [P^i]_{xy}.$$

Thus,

$$\begin{aligned}\mathbb{E}_\pi \mathbb{E}_x N_n(y) &= \sum_{x\in\mathcal{X}} \mathbb{E}_x N_n(y)\, \vec{\pi}_x \\ &= \sum_{x\in\mathcal{X}} \sum_{i=1}^{n} [P^i]_{xy} \cdot \vec{\pi}_x \\ &= \sum_{i=1}^{n} [\vec{\pi} \cdot P^i]_y = \sum_{i=1}^{n} \vec{\pi}_y = n\vec{\pi}_y.\end{aligned}$$

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

Therefore, $\vec{\pi}_y = \dfrac{1}{\mathbb{E}_y \tau_y}$, $\forall y \in \mathcal{X}$. ∎

Recall: *Theorem 8.1*. If I & R hold and $\vec{\mu}^x_y := \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, \tau_x > n)$, then $\vec{\mu}^x$ is a stationary measure with $0 < \vec{\mu}^x_y < \infty$, $\forall y \in \mathcal{X}$.

Proposition 11.1. If I(, R) & S hold and $\vec{\mu}^x_y := \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, \tau_x > n)$,

then
$$\vec{\mu}^x_y = \frac{\vec{\pi}_y}{\vec{\pi}_x}, \quad \forall x, y \in \mathcal{X}.$$

Proof. For any fixed $x \in \mathcal{X}$, Theorem 8.1 implies that $\vec{\pi}_x \cdot \vec{\mu}^x$ is a stationary measure. Notice that

$$\begin{aligned}
\sum_{y \in \mathcal{X}} \vec{\mu}^x_y &= \sum_{y \in \mathcal{X}} \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, \tau_x > n) \\
&= \sum_{n=0}^{\infty} \sum_{y \in \mathcal{X}} \mathbb{P}_x(X_n = y, \tau_x > n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}_x(\tau_x > n) \\
&= \mathbb{E}_x \tau_x \\
&= \frac{1}{\vec{\pi}_x}.
\end{aligned}$$

why?

By Thm 9.3

Thus, $\sum_{y \in \mathcal{X}} [\vec{\pi}_x \cdot \vec{\mu}^x]_y = \vec{\pi}_x \cdot \sum_{y \in \mathcal{X}} \vec{\mu}^x_y = 1$.

Therefore, $\vec{\pi}_x \cdot \vec{\mu}^x$ is also a stationary distribution.

Then Corollary 9.1 implies $\vec{\pi}_x \cdot \vec{\mu}^x = \vec{\pi}$.

That is, $\forall y \in \mathcal{X}$, $\vec{\pi}_y = [\vec{\pi}_x \cdot \mu^x]_y = \vec{\pi}_x \cdot \vec{\mu}^x_y$. ∎

Recall Theorem 9.4. If I & S hold, and $f : \mathcal{X} \longrightarrow \mathbb{R}$ has $\sum_{x \in \mathcal{X}} |f(x)| \vec{\pi}_x < \infty$, then

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k) \xrightarrow{a.s.} \sum_{x \in \mathcal{X}} f(x) \cdot \vec{\pi}_x = \mathbb{E}_{X \sim \pi}[f(X)].$$

Sketch of the Proof. Suppose the chain starts at x. Let $\tau_x^0 := 0$, and $\tau_x^k := \min\{n > \tau_x^{k-1} : X_n = x\}$ be the time of the k-th return to x. Define

$$Y_k := \sum_{m=\tau_x^{k-1}+1}^{\tau_x^k} f(X_m), \quad \forall k = 1, 2, \cdots.$$

By the Strong Markov Property, $Y_1, Y_2, Y_3, \cdots$ are i.i.d. Then the Strong Law of Large Numbers tells

$$\frac{1}{L}\sum_{k=1}^{L} Y_k \xrightarrow{a.s.} \mathbb{E}_x Y_1 . \qquad (*)$$

On the other hand,

$$\mathbb{E}_x Y_1 = \mathbb{E}_x\left[\sum_{m=1}^{\tau_x^1} f(X_m)\right]$$

$$= \mathbb{E}_x\left[\sum_{m=1}^{\infty} f(X_m)\cdot \mathbb{1}_{\{\tau_x^1 \geq m\}}\right]$$

$$= \sum_{m=1}^{\infty} \mathbb{E}_x\left[f(X_m)\cdot \mathbb{1}_{\{\tau_x^1 \geq m\}}\right]$$

$$= \sum_{m=1}^{\infty} \sum_{y\in\mathcal{X}} f(y)\, \mathbb{P}_x(X_m = y,\ \tau_x^1 \geq m)$$

$$= \sum_{y\in\mathcal{X}} f(y) \sum_{m=1}^{\infty} \mathbb{P}_x(X_m = y,\ \tau_x^1 \geq m)$$

why?

$$= \sum_{y\in\mathcal{X}} f(y) \sum_{m=0}^{\infty} \mathbb{P}_x(X_m = y,\ \tau_x^1 > m)$$

$$= \sum_{y\in\mathcal{X}} f(y)\, \vec{\mu}^{\,x}_y$$

$$= \sum_{y\in\mathcal{X}} f(y)\cdot \frac{\vec{\pi}_y}{\vec{\pi}_x} .$$

Notice that $\displaystyle\sum_{k=1}^{N_n(x)} Y_k \leq \sum_{m=1}^{n} f(X_m) < \sum_{k=1}^{N_n(x)+1} Y_k$.

Taking $L = N_n(x) = \max\{k : \tau_x^k \leq n\}$, one has

$$\frac{N_n(x)}{n}\cdot\frac{1}{L}\sum_{k=1}^{L} Y_k \leq \frac{1}{n}\sum_{m=1}^{n} f(X_m) < \frac{N_n(x)+1}{n}\cdot\frac{1}{L+1}\sum_{k=1}^{L+1} Y_k .$$

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

From $(*)$ and Corollary 9.2, we have

$$\frac{1}{n}\sum_{m=1}^{n} f(X_m) \xrightarrow{a.s.} \vec{\pi}_x \cdot \sum_{y\in\mathcal{X}} f(y)\cdot \frac{\vec{\pi}_y}{\vec{\pi}_x} = \sum_{y\in\mathcal{X}} f(y)\cdot \vec{\pi}_y .$$

∎

Theorem 11.1 If I holds, and $\vec{\nu}$ satisfies $\begin{cases} \vec{\nu}P = \vec{\nu} \\ \vec{\nu}_x = 1 \quad \text{for some } x \end{cases}$,
then $\vec{\nu} = \vec{\mu}^x$. Here $\vec{\mu}^x$ is defined in Theorem 8.1 such that $\vec{\mu}^x_y = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, \tau_x > n) \quad \forall y \in \mathcal{X}$.

Proof. For any $y \neq x$, $\vec{\nu}_y = \sum_{y_1\in\mathcal{X}} \vec{\nu}_{y_1} P_{y_1 y} = P_{xy} + \sum_{y_1\neq x} \vec{\nu}_{y_1} P_{y_1 y}$.
Using the above equation recursively, one has

$$\begin{aligned}
\vec{\nu}_y &= P_{xy} + \sum_{y_1\neq x}\Big(P_{xy_1} + \sum_{y_2\neq x}\vec{\nu}_{y_2}P_{y_2y_1}\Big)P_{y_1y} \\
&= P_{xy} + \sum_{y_1\neq x} P_{xy_1}P_{y_1y} + \sum_{y_1,y_2\neq x}\Big(P_{xy_2} + \sum_{y_3\neq x}\vec{\nu}_{y_3}P_{y_3y_2}\Big)P_{y_2y_1}P_{y_1y} \\
&= \cdots \\
&\geq P_{xy} + \sum_{y_1\neq x} P_{xy_1}P_{y_1y} + \sum_{y_1,y_2\neq x} P_{xy_2}P_{y_2y_1}P_{y_1y} + \cdots \\
&= \mathbb{P}_x(X_1 = y, \tau_x > 1) + \mathbb{P}_x(X_2 = y, \tau_x > 2) + \mathbb{P}_x(X_3 = y, \tau_x > 3) + \cdots \\
&= \vec{\mu}^x_y , \qquad \forall\ y \neq x .
\end{aligned}$$

Also, $\vec{\nu}_x = 1 = \vec{\mu}^x_x$. Thus, $\vec{\nu} \geq \vec{\mu}^x$.

Define $\vec{w} = \vec{\nu} - \vec{\mu}^x$.

Then, $\vec{w} \geq 0$ and $\vec{w} P = \vec{\nu} P - \vec{\mu}^x P = \vec{\nu} - \vec{\mu}_x = \vec{w}$.

For any y, since P is irreducible, $\exists M_y$, s.t. $[P^{M_y}]_{yx} > 0$.

Thus, $0 = \vec{w}_x = [\vec{w} \cdot P^{M_y}]_x = \sum_{z \in \mathcal{X}} \vec{w}_z [P^{M_y}]_{zx} \geq \vec{w}_y \cdot [P^{M_y}]_{yx}$

≥ 0

This implies, $\vec{w}_y \cdot [P^{M_y}]_{yx} = 0$.

And thus, $\vec{w}_y = 0$, $\forall y \in \mathcal{X}$.

Therefore, $\vec{\nu} = \vec{\mu}_x$. $\boxtimes$

Theorem 11.2. If I & R hold, all the stationary measures are unique up to a constant.

Proof. Suppose $\vec{\mu}$ is a stationary measure. There exists $x \in \mathcal{X}$, s.t. $\vec{\mu}_x > 0$. Let $\vec{\nu} := \frac{\vec{\mu}}{\vec{\mu}_x}$. Then $\vec{\nu} = \vec{\nu} P$, and $\vec{\nu}_x = 1$. Theorem 11.1 tells $\vec{\nu} = \vec{\mu}^x$.

Thus, $\vec{\mu} = \vec{\mu}_x \cdot \vec{\mu}^x$. This means, any stationary measure equals to an element from $\{\vec{\mu}^x : x \in \mathcal{X}\}$, up to some positive constant. For any $y \in \mathcal{X}$, Theorem 8.1 says $0 < \vec{\mu}_x^y < \infty$. Thus, $\vec{\mu}^y = \vec{\mu}_x^y \cdot \vec{\mu}^x$. Therefore, any stationary measure equals to $\vec{\mu}^x$, up to some positive constant. $\boxed{}$

This is the end of this lecture!