Corollary

Let $(S^{\mathbb{N}_0}, \mathcal{B}(S^{\mathbb{N}_0}), \sigma)$ be a $(q,P)$-Markov chain with $P$ irreducible. Let $p \in [0,1]^S$ stationary, i.e., $pP = p$. Then

$$\forall f \in \mathbb{C}^S \quad \frac{1}{N}\sum_{n=0}^{N-1} f\circ\pi_n \longrightarrow \sum_{s\in S} p(s)f(s) \quad \sigma\text{-a.s. and in } L^1(\sigma)$$

Proof: $f\circ\pi_n = f\circ\pi_0\circ\sigma^n$, thus

$$\frac{1}{N}\sum_{n=0}^{N-1} f\circ\pi_n \longrightarrow \int f\circ\pi_0 \, d\mu = \sum_{s\in S} f(s)\mu(\mathcal{C}(s)) = \sum_{s\in S} p(s)f(s)$$

$\boxtimes$

Corollary

Let $(S^{\mathbb{N}_0}, \mathcal{B}(S^{\mathbb{N}_0}), \sigma)$ be an irreducible $(q,P)$-Markov chain, $pP = p$ the stationary distribution. Let $s \in S$ and, given $N \in \mathbb{N}$, let $V_s(N) = \sum_{n=0}^{N-1} 1\!\!1_{\{\pi_n = s\}}$. Then

$$\frac{V_s(N)}{N} \longrightarrow \frac{1}{\mathbb{E}_\mu(\tau_s|\pi_0=s)} \quad \sigma\text{-a.s. and in } L^1(\sigma).$$

Proof: $V_s(N) = \sum_{n=0}^{N-1} 1\!\!1_{\{s\}}\circ\pi_n$ and $\frac{1}{\mathbb{E}_\mu(\tau_s|\pi_0=s)} = p(s)$.

$\boxtimes$

Corollary:

Let $P$ be an irreducible matrix with probability vectors as rows. Let $p$ be the stationary distribution. Then

$$\frac{1}{N}\sum_{n=0}^{N-1} P_{st}^{(n)} \longrightarrow p(t)$$

as $N\to\infty$.

Proof: Let $(S^{\mathbb{N}_0}, \mathcal{B}(S^{\mathbb{N}_0}), \sigma)$ be the irreducible Markov chain induced by $(e_s, P)$. Then

$$\frac{1}{N}\sum_{n=0}^{N-1} P_{st}^{(n)} = \mathbb{E}_\sigma\left(\frac{1}{N}\sum_{n=0}^{N-1} 1\!\!1_{\{t\}}\circ\pi_n\right) \longrightarrow p(t).$$

$\boxtimes$

Theorem

Let $(S^{\mathbb{N}_0}, \mathcal{B}(S^{\mathbb{N}_0}), \mu)$ be a fully supported stationary $(p,P)$-Markov chain. Then

$$Q = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} P^n$$

exists, is stochastic (i.e., rows are probability vectors), and $QP = PQ = Q$.

If $p \in \mathbb{R}^S$ is an eigenvector for eigenvalue $1$ (of $P$), then $p$ is an eigenvector of $Q$ for eigenvalue $1$.

Proof: Let $s \in S$ and $f_s = 1\!\!1_{\{s\}}\circ\pi_0$. By Birkhoff's ergodic theorem

$$A_N f_s = \frac{1}{N}\sum_{n=0}^{N-1} 1\!\!1_{\{s\}}\circ\pi_n \xrightarrow{N\to\infty} \mathbb{E}(f_s|\Sigma_{\sigma,\mu}) =: f_s^* \quad \mu\text{-a.s. and in } L^1(\mu).$$

Recall that $f_s^*$ is essentially $\sigma$-invariant.

Let $t \in S$. By Lebesgue's dominated convergence

$$\int f_t \cdot A_N f_s \, d\mu \xrightarrow{N\to\infty} \int f_t f_s^* \, d\mu.$$

(In particular, if $\mu$ is ergodic, then $\int f_t \cdot A_N f_s \, d\mu \longrightarrow p(t)\cdot p(s)$.)

Note that

$$\int f_t \cdot A_N f_s \, d\mu = \frac{1}{N}\sum_{n=0}^{N-1} \int 1\!\!1_{\{\pi_0 = t\}} 1\!\!1_{\{\pi_n = s\}} \, d\mu$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \mu(\pi_n = s, \pi_0 = t) = \frac{p(t)}{N}\sum_{n=0}^{N-1} \mu(\pi_n = s|\pi_0 = t)$$

$$= p(t)\cdot\frac{1}{N}\sum_{n=0}^{N-1} (P^n)_{ts}$$

$$\Rightarrow \frac{1}{N}\sum_{n=0}^{N-1} (P^n)_{ts} \longrightarrow \frac{1}{p(t)}\int f_t f_s^* \, d\mu, \text{ i.e.,}$$

$$Q_{st} = \frac{1}{p(s)}\int f_s f_t^* \, d\mu.$$

Verification of the properties is straight-forward:

- $QP = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} P^n = \lim_{N\to\infty} \frac{N+1}{N}\left(\frac{1}{N+1}\sum_{n=0}^{N} P^n - \frac{1}{N+1}P^0\right) = Q$
- $Q^2 = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \underbrace{QP^n}_{=Q} = Q$
- $\sum_{t\in S} Q_{st} = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \overbrace{\sum_{t\in S}(P^n)_{st}}^{=1} = 1$
- $vP = v$
  $\Rightarrow vQ = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \underbrace{vP^n}_{=v} = v$

**Theorem** (Characterization of ergodicity)

Let $(S^{\mathbb{N}_0}, \mathcal{B}(S^{\mathbb{N}_0}), \mu)$ be a stationary fully supported $(p,P)$-Markov chain. Let $Q = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} P^n$. The following are equivalent.

(i) $\sigma$ is ergodic

(ii) The rows of $Q$ are pairwise identical.

(iii) $Q>0$

(iv) $P$ is irreducible

(v) $1$ is a simple eigenvalue of $P$.

We will prove (i)$\Leftrightarrow$(ii)$\Leftrightarrow$(iii)$\Leftrightarrow$(iv) only. We use the following lemma.

**Lemma**

Let $(X,\mathcal{B},\mu,T)$ a pmps. Then $\mu$ is ergodic iff there exists a dense subspace $V\subseteq L^2(\mu)$ s.t.

$$\forall f,g\in V \quad \frac{1}{N}\sum_{n=0}^{N-1}\langle f\circ T^n, g\rangle_2 \longrightarrow \langle f,1\rangle_2\langle 1,g\rangle_2 .$$

**Sketch** As shown in exercises, for $V=L^2(\mu)$, this characterizes ergodicity. Let $f,g\in L^2(\mu)$ and $f_\varepsilon, g_\varepsilon\in L^2(\mu)$ s.t. $\|f-f_\varepsilon\|_2<\varepsilon$, $\|g-g_\varepsilon\|_2<\varepsilon$, then

$$\frac{1}{N}\sum_{n=0}^{N-1}\langle f\circ T^n, g\rangle_2 = O(\varepsilon(\|g\|_2+\|f_\varepsilon\|_2)) + \frac{1}{N}\sum_{n=0}^{N-1}\langle f_\varepsilon\circ T^n, g_\varepsilon\rangle_2,$$

since $U_T$ is an isometry. ∎

**Proof:** (i)$\Rightarrow$(ii): We have seen before that for ergodic $\mu$ we have

$$Q_{st} = p(t).$$

(ii)$\Rightarrow$(iii): Suppose $Q_{st}=0 \Rightarrow Q_{\cdot t}\equiv 0 \Rightarrow p(t) = (pQ)_t = pQ_{\cdot t} = 0$ ↯

(iii)$\Rightarrow$(iv): Let $s,t\in S$. Since

$$0 < Q_{st} = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} P^n_{st},$$

there is $n\geq 0$ (in fact, there are $\infty$-many $n$) s.t. $P^n_{st}>0$, i.e., $P$ is irreducible.

(iv)$\Rightarrow$(iii): Let $s\in S$ and $S_s := \{t\in S : Q_{st}>0\}$. Since $Q$ is a stochastic matrix, we know that $S_s\neq\emptyset$ for all $s\in S$.

Given $s,r,t\in S$, we have

$$Q_{st} = \sum_{g\in S} Q_{sg}P_{gt} \geq Q_{sr}P_{rt}$$

$Q=QP$

In particular, for all $s\in S$ and for all $r\in S_s$ we have

$$\forall t\in S \quad P_{rt}>0 \Rightarrow Q_{st}>0, \text{ i.e.,}$$

$$\forall t\in S \quad P_{rt}>0 \Rightarrow t\in S_s .$$

Thus $\mathrm{supp}(P_{r\cdot})\subseteq S_s$. It follows that

$$\forall r\in S_s \quad \sum_{t\in S} P_{rt} = \sum_{t\in S_s} P_{rt},$$

i.e.,

$$\forall r\in S_s\, \forall t\notin S_s \quad P_{rt} = 0.$$

If $t\in S - S_s$, it follows that $\forall \ell\geq 0\, \forall s=s_0,s_1,\dots,s_\ell\in S^{\ell+1}$

$$P_{s_\ell t}\cdot\prod_{k=0}^{\ell-1} P_{s_k s_{k+1}} = 0,$$

i.e., $s \not\to t$.

Since $P$ is irreducible, $s\to t$ for all $s,t\in S$, and thus $S_s = S$.

(iii)$\Rightarrow$(ii): Let $t\in S$, $q_t = \max_{s\in S} Q_{st}$. Suppose (for sake of contradiction) that $\exists \mathfrak{s}\in S$ s.t.

$Q_{\mathfrak{s}t} < q_t$. Then for all $s\in S$

$$Q_{st} = (Q^2)_{st} = \sum_{r\in S} Q_{sr}Q_{rt} \overset{(iii)}{<} q_t\sum_{r\in S} Q_{sr} = q_t \quad ↯$$

(ii)$\Rightarrow$(i): It suffices to show that

$$\forall f,g\in C(X) \quad \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\langle f\circ T^n, g\rangle_2 = \langle f,1\rangle_2\langle 1,g\rangle_2 .$$

Since $f,g$ are uniformly continuous, there exist for every $\varepsilon>0$ some $m(\varepsilon)\in\mathbb{N}$ s.t.

$$\|\mathbb{E}(f|\mathcal{F}_m) - f\|_\infty < \varepsilon \quad\text{and}\quad \|\mathbb{E}(g|\mathcal{F}_m) - g\|_\infty < \varepsilon \qquad (m\geq m(\varepsilon)),$$

where for $\varphi\in C(X)$, we let

$$\mathbb{E}(\varphi|\mathcal{F}_m) = \sum_{\underline{s}\in S^{m+1}} \frac{\mathbb{E}(\varphi\mathbb{1}_{\mathcal{C}(\underline{s})})}{\mu(\mathcal{C}(\underline{s}))}\mathbb{1}_{\mathcal{C}(\underline{s})}$$

pointwise. As in the proof of the lemma, it follows that

$$\frac{1}{N}\sum_{n=0}^{N-1}\langle f\circ T^n, g\rangle_2 = \frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{E}(f|\mathcal{F}_m)\circ T^n, \mathbb{E}(g|\mathcal{F}_m)\rangle_2 + O(\varepsilon(\|g\|_\infty + \|f\|_\infty))$$

Hence, it suffices to show that

$$\frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{E}(f|\mathcal{F}_m)\circ T^n, \mathbb{E}(g|\mathcal{F}_m)\rangle_2 \longrightarrow \langle \mathbb{E}(f|\mathcal{F}_m), 1\rangle_2\langle 1, \mathbb{E}(g|\mathcal{F}_m)\rangle_2$$

$\langle f,1\rangle_2\langle 1,g\rangle_2$

Since $m$ is fixed, it suffices to show that

$$\forall \underline{s},\underline{t}\in S^{m+1} \quad \frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 \longrightarrow \mu(\mathcal{C}(\underline{s}))\mu(\mathcal{C}(\underline{t}))$$

Note that

$$\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 = \mu(T^{-n}\mathcal{C}(\underline{s})\cap\mathcal{C}(\underline{t}))$$

$$\overset{n>m}{=} p(t_0)\cdot P_{t_0t_1}\cdot\ldots\cdot P_{t_{m-1}t_m}\cdot P^{(n-m)}_{t_ms_0}\cdot P_{s_0s_1}\cdot\ldots\cdot P_{s_{m-1}s_m},$$

hence

$$\frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 = \frac{1}{N}\sum_{n=0}^{m}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2$$

$$+ p(t_0)P_{t_0t_1}\cdot\ldots\cdot P_{t_{m-1}t_m}\cdot P_{s_0s_1}\cdot\ldots\cdot P_{s_{m-1}s_m}\cdot \underbrace{\frac{1}{N}\sum_{n=m+1}^{N-1} P^{(n-m)}_{t_ms_0}}_{\to Q_{t_ms_0}}$$

Now $pQ = p \Rightarrow p(s_0) = \sum_{t\in S} p(t)Q_{ts_0} \overset{(ii)}{=} Q_{\cdot s_0}\sum_{t\in S} p(t) = Q_{\cdot s_0}$

$$\therefore \frac{1}{N}\sum_{n=0}^{N-1}\langle \mathbb{1}_{\mathcal{C}(\underline{s})}\circ T^n, \mathbb{1}_{\mathcal{C}(\underline{t})}\rangle_2 \xrightarrow{N\to\infty} p(t_0)\prod_{k=0}^{m-1}P_{t_kt_{k+1}}\cdot p(s_0)\prod_{\ell=0}^{m-1}P_{t_\ell t_{\ell+1}}$$

$$= \mu(\mathcal{C}(\underline{t}))\mu(\mathcal{C}(\underline{s})).$$

Exercise: Show that (ii) $\Leftrightarrow$ (v).

End of lecture 20



(Notions of) Mixing

Recall: $(X,\mathcal{B},\mu,T)$ is ergodic $\Leftrightarrow \forall f,g\in L^2(\mu)\ \frac{1}{N}\sum_{n=0}^{N-1}\langle f\circ T^n, g\rangle_2 \longrightarrow \langle f,1\rangle_2\langle 1,g\rangle_2$.

$$\Leftrightarrow \forall A,B\in\mathcal{B} \quad \frac{1}{N}\sum_{n=0}^{N-1}\mu(T^{-n}A\cap B)\longrightarrow \mu(A)\mu(B).$$

Definition ((Strong) mixing)

Let $(X,\mathcal{B},\mu,T)$ a pmps. $T$ (or $\mu$) is called (strong) mixing if

$$\forall A,B\in\mathcal{B} \qquad \mu(T^{-n}A\cap B)\xrightarrow{n\to\infty}\mu(A)\mu(B).$$

Lemma

Let $(X,\mathcal{B},\mu,T)$ a pmps. $T$ is mixing if

$$\forall f,g\in L^2(X,\mu) \quad \langle U_T^n f, g\rangle_2 \longrightarrow \langle f,1\rangle_2\langle 1,g\rangle_2$$

Proof: Skipped; cf. characterization of ergodicity in the exercise sheets.

This lemma, for invertible $T$, gives an important connection between dynamical systems and unitary representation theory, which is about (strongly continuous) group homomorphisms

$$\pi: G\longrightarrow \mathcal{U}(H).$$

For example, if $\Gamma = SL_2(\mathbb{Z})$, $G = SL_2(\mathbb{R})$, then $\Gamma\backslash G$ admits a $G$-invariant probability measure and every element $g\in G$ induces an (effectively) mixing transformation, i.e.,

$$\forall \varphi,\psi\in C_c^\infty(\Gamma\backslash G) \qquad |\langle \varphi\circ g,\psi\rangle_2 - \langle\varphi,1\rangle_2\langle 1,\psi\rangle_2| \ll \xi(g)S(\varphi)S(\psi)$$

Unfortunately, this is basically the whole theory for this notion.

Definition (Weak mixing)

Let $(X,\mathcal{B},\mu,T)$ a pmps. $T$ is weak mixing if

$$\forall A,B\in\mathcal{B} \quad \frac{1}{N}\sum_{n=0}^{N-1}|\mu(T^{-n}A\cap B)-\mu(A)\mu(B)| \xrightarrow{N\to\infty} 0.$$

Weak mixing is surprisingly interesting/important which is due to the many equivalent characterisations. We need a few more notions:

Definition (Density)

A subset $J\subseteq\mathbb{N}_0$ has density zero if

$$0 = \lim_{N\to\infty}\frac{1}{N}|J\cap[0,N)|$$