ET73

1

- If A is a stochastic matrix ($A \geq 0$, Row sums $=1$) That is irreducible ($\forall i,j \; \exists N$ with $(A^N)_{ij} > 0$) There is a unique $\vec{p} > 0$ with $\vec{p}A = \vec{p}$ and the sum of its entries equal to one. Define the (p, A) Markov measure on cylinder sets by

$$\mu({}_m[b_0 \ldots b_k]) = p_{b_0} A_{b_0 b_1} A_{b_1 b_2} \cdots A_{b_{k-1} b_k}.$$

It is invariant under the shift $\sigma : \Sigma_n \to \Sigma_n$

- So the probability of the finite sequence $b_0 \ldots b_k$ occuring is the probability of starting at b_0 $(= p_{b_0})$ times the transition probability $b_0 \to b_1$ $(= A_{b_0 b_1})$ times the transition probability $b_1 \to b_2$ $(= A_{b_1 b_2})$ times ... times the transition probability $b_{k-1} \to b_k$

<u>Theorem</u> If A is irreducible, the $(\vec{p}, A)$ Markov measure is ergodic under $\sigma : \Sigma_n \circlearrowleft$. If, in addition, A is primitive ($\exists N$ with $A^N > 0$), then μ is mixing.

<u>Proof</u>: The proof is easy given the following lemmas

<u>Lemma 1</u>. Assume A is stochastic

(a) If A is irreducible with $\vec{p} > 0$ and $\vec{p} A = \vec{p}$

then $\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} A^n \to P := \begin{bmatrix} \vec{p} \\ \vec{p} \\ \vdots \\ \vec{p} \end{bmatrix}$

(b) If A is primitive then

$A^n \to P$.

<u>Lemma 2:</u> Assume $k > m+1+\ell \Rightarrow$

$$\mu\left({}_{\ell}[a_0 \dots a_m] \cap {}_{k}[b_0 \dots b_w]\right)$$
$$= p_{a_0} A_{a_0 a_1} \cdots A_{a_{m-1} a_m} \left(A^{k-m-\ell}\right)_{a_m b_0} A_{b_0 b_1} \cdots A_{b_{w-1} b_w}$$

<u>Proof of theorem</u>: We know that it suffices to check

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu\left(B \cap \sigma^{-i}(C)\right) \to \mu(B)\mu(C)$$

and $\lim_{n\to\infty} \mu\left(B \cap \sigma^{-n}(C)\right) \to \mu(B)\mu(C)$

on cylinder sets to yield ergodicity and mixing respectively.

First note that the formula in Lemma 2 can be rewritten as

$$\mu\left({}_{\ell}[a_0 \ldots a_m] \cap {}_{k}[b_0 \ldots b_w]\right)$$

$$= \mu\left({}_{\ell}[a_0 \ldots a_m]\right) \cdot \frac{\left(A^{k-m-\ell}\right)_{a_m b_0}}{p_{b_0}} \mu\left({}_{k}[b_0 \ldots b_w]\right)$$

Now given cylinder sets $B = {}_{\ell}[b_0 \ldots b_m]$ and $C = {}_{k}[c_0 \ldots c_w]$ since $\sigma^{-n}(C) = {}_{k+n}[c_0 \ldots c_w]$ we may pick n large enough so that $k+n > m+\ell+1$ and then

$$\mu\left(B \cap \sigma^{-n}(C)\right) = \mu(B) \frac{\left(A^{k+n-m-\ell}\right)_{a_m c_0}}{p_{c_0}} \mu\left(\sigma^{-n}(C)\right)$$

$$\underset{\substack{\text{invariance}\\ \text{of measure}}}{=} \mu(B) \frac{\left(A^{k+n-m-\ell}\right)_{b_m c_0}}{p_{c_0}} \mu(C) \to \mu(B)\mu(C)$$

as $n \to \infty$ when A is primitive, by Lemma 1 (b), proving mixing.

Ergodicity is similar using Lemma 1 (a).

- We start with Lemma 2 by first trying to understand what is says via an example

- Here is an example in Σ_3,
  Pictorially ${}_0[10] \cap {}_3[20]$ is $**.10*20**$ with $*$ = "wildcard"

- The key step is to write the intersection as the disjoint union of length 5 cylinder sets

- ${}_0[10] \cap {}_3[20] = {}_0[10020] \amalg {}_0[10120] \amalg {}_0[10220]$

- Thus $\mu({}_0[10] \cap {}_3[20]) = \sum_{i=0}^{2} \mu([10i20])$

$$= \sum_{i=0}^{2} p_1 A_{10} A_{0i} A_{i2} A_{20}$$

$$= p_1 A_{10} \left(\sum_{i=0}^{2} A_{0i} A_{i2} \right) A_{20}$$

but

$$\begin{bmatrix} A_{00} & A_{01} & A_{02} \\ \multicolumn{3}{c}{\text{———}} \\ \multicolumn{3}{c}{\text{———}} \end{bmatrix} \begin{bmatrix} | & | & A_{02} \\ | & | & A_{12} \\ | & | & A_{22} \end{bmatrix}$$

$$= \begin{bmatrix} & & \sum_{i=0}^{2} A_{0i} A_{i2} \\ & & \\ & & \end{bmatrix}$$

which is to say $\sum_{i=0}^{2} A_{0i} A_{i2} = (A^2)_{02}$

So $\mu|_0[10] \wedge_3 [20]) = p_1 A_{10} (A^2)_{02} A_{20}$

Similarly

$$\mu\left({}_0[10] \cap {}_4[20]\right)$$
$$= p_1 A_{10} \left(\sum_{\substack{i=0,1,2 \\ j=0,1,2}} A_{0i} A_{ij} A_{j2} \right) A_{20}$$
$$= p_1 A_{10} \left(A^3\right)_{02} A_{20}$$

Lemma 3: $\sum A_{a i_1} A_{i_1 i_2} \cdots A_{i_k b}$ (A is $n \times n$)

with the sum $i_j \in \{0, \ldots, n-1\}$ for all j.

is equal to $\left(A^{k+1}\right)_{ab}$

Proof: Induction

What is the meaning in terms of transition probability. Back to the Σ_2 example

$$\begin{tikzcd} & 0 \arrow[dr, "A_{02}"] & \\ 0 \arrow[ur, "A_{00}"] \arrow[r, "A_{01}"] \arrow[dr, "A_{02}"'] & 1 \arrow[r, "A_{12}"] & 2 \\ & 2 \arrow[ur, "A_{22}"'] & \end{tikzcd}$$

So $\sum_{i=0}^{2} A_{0i} A_{i2}$ is the transition probability from $0 \to 2$ in $\underline{2}$ steps.

Using Lemma 3, in general, $(A^k)_{ab}$ is the transition probability $a \to b$ in k-steps.

Thus, for example,

$\mu({}_0[10] \wedge {}_4[20])$ is the probability of the event start at 1, transition to 0, transition to 2 in 3 steps, transition to 0 or

$$p_1 A_{10} (A^3)_{02} A_{20}$$

Proof of Lemma 2: Let $\ell = 0$ for simplicity

$${}_0[a_0 .. a_m] \wedge {}_k[b_0 \dots b_w]$$

$$= \coprod_{i_{m+1}, .., i_{k-1} \in \{0, .., n-1\}} [a_0 .. a_m i_{m+1} \cdots i_{k-1} b_0 \dots b_w]$$

So

$$\mu\left({}_0[a_0 \ldots a_m] \cap {}_k[b_0 \ldots b_w]\right)$$

$$= \sum p_{a_0} A_{a_0 a_1} \cdots A_{a_{m-1} a_m} A_{a_m i_{m+1}} \cdots A_{i_{k-2} i_{k-1}} \cdot A_{i_{k-1} b_0} A_{b_0 b_1} \cdots A_{b_{w-1} b_w}$$

$$\underset{\substack{\uparrow \\ \text{Lemma 3}}}{=} p_{a_0} A_{a_0 a_1} \cdots A_{a_{m-1} a_m} \left(A^{k-m}\right)_{a_m b_0} A_{b_0 b_1} \cdots A_{b_{w-1} b_w} \quad \blacksquare$$

It remains to prove Lemma 1, Part (b) is from the P.F. Theorem.

Proof of Lemma 1 (a): We first show that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} A^n \text{ exists}$$

Let χ_j be the indicator function of the cylinder set ${}_0[j]$. By Pointwise ergodic Theorem

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_j(\sigma^n(x)) \to \chi_j^*(x) \quad \text{a.e.}$$

multiply by χ_i and use bounded convergence theorem as before

this is an i

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int \chi_j(\sigma^n(x))\, \chi_i(x)\, d\mu$$

$$\to \int \chi_j^* \chi_i \, d\mu$$

but note that

$$\int \chi_j(\sigma^n(x))\,\chi_i(x)\,d\mu$$

$$= \int \chi_{\sigma^{-n}[j] \cap [i]}\,d\mu$$

$$= \mu(\sigma^{-n}[j] \cap [i])$$

$$= p_i (A^n)_{ij} \quad \text{by Lemma 2}$$

Thus if we define $Q = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} A^n$

we have $Q_{ij} = \dfrac{\int \chi_j^* \chi_i \, d\mu}{p_i}$.

It remains to show that $Q = \begin{bmatrix} \vec{p} \\ \vec{p} \\ \vdots \\ \vec{p} \end{bmatrix}$ under the assumption that A is irreducible and thus there exists a unique $\vec{p}$ with $\vec{p}A = \vec{p}$, $\vec{p} > 0$ using P.F.

To show this, first note $QA = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} A^{n+1}$

$= \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} A^n = Q$ by the same trick we used in the proof of the pointwise ergodic theorem. (ET9, page 10). Thus if $\vec{R}_i$ is the i^{th} row of Q, $\vec{R}_i A = \vec{R}_i$. This implies that $\vec{R}_i$ is a multiple of $\vec{p}$ since it is a left eigenvector of the eigenvalue 1. To finish, we need that the

The sum of the entries of $\vec{R}_L$ is one.

Now A is stochastic and thus so is A^n for any $n > 0$, thus

$$\sum_{n=0}^{N-1} A^n \text{ has row sums equal to } N$$

and so $\frac{1}{N} \sum_{n=0}^{N-1} A^n$ is stochastic

and thus $\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} A^n = Q$ is also,

so Q has row sums equal to one

and thus $Q = \begin{bmatrix} \vec{P} \\ \vdots \\ \vec{P} \end{bmatrix}$ $\blacksquare$