Lecture 6

We continue to consider some more examples of ergodic transformations.

Example 1. (Bernoulli shift on finite state space).

Let $l \geq 2$ be an integer. Consider $\Sigma^{\mathbb{N}} = \{(x_i)_{i=1}^{\infty} : x_i \in \{1, 2, \cdots, l\}\}$ and $\sigma : \Sigma^{\mathbb{N}} \to \Sigma^{\mathbb{N}}$ defined by $\sigma((x_i)_{i=1}^{\infty}) = (x_{i+1})_{i=1}^{\infty}$. Let $(p_1, p_2, \cdots, p_l)$ be a probability vector, i.e. $p_i > 0$ for each i and $\sum_{i=1}^{l} p_i = 1$. Define μ on $\Sigma^{\mathbb{N}}$ by $\mu([i_1 i_2 \cdots i_k]) = p_{i_1} p_{i_2} \cdots p_{i_k}$ for any $i_i i_2 \cdots i_k \in \{1, 2, \cdots, l\}^k$, where $[i_i i_2 \cdots i_k] := \{x \in \Sigma^{\mathbb{N}} : x_1 = i_1, x_2 = i_2, \cdots, x_k = i_k\}$ is called a cylinder. Let $\mathscr{G}$ be the collection of all cylinders, then $\mathscr{G}$ is a semi-algebra generating $\mathscr{B}(\Sigma^N)$. Since μ is countably additive on $\mathscr{G}$, by Kolmogorov consistency theorem, μ extends uniquely to a probability measure on $\mathscr{B}$, still denoted by μ. We claim that σ is ergodic w.r.t μ. To see this, let $A = [i_1 i_2 \cdots i_k]$ and $B = [j_1 j_2 \cdots j_m]$ be two cylinders, then for $i > m$, $\mu(\sigma^{-i} A \cap B) = \mu(A)\mu(B)$. Hence

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i} A \cap B) = \mu(A)\mu(B), \text{ for } A, B \in \mathscr{G}.$$

By Theorem 3.10, σ is ergodic.

The same argument also works in the following general setting.

Example 2. (Bernoulli shift on general state spaces).

Let $(Y, \mathscr{F}, \mu)$ be a probability space. Let $(X, \mathscr{B}, m) = \prod_{i=0}^{\infty} (Y, \mathscr{F}, \mu)$. Define $T : X \to X$ by $(y_i)_{i=0}^{\infty} \mapsto (y_{i+1})_{i=0}^{\infty}$. Then T is ergodic w.r.t m.

Example 3. (Markov shift).

Let $l \geq 2$ be an integer. Let $A = (a_{ij})_{l\times l}$ with $0, 1$ entries. Define

$$\Sigma_A^{\mathbb{N}} := \{(x_i)_{i=1}^{\infty} : x_i \in \{1, 2, \cdots, l\} \text{ and } a_{x_i x_{i+1}} = 1 \text{ for all } i\}.$$

Define $\sigma : \Sigma_A^{\mathbb{N}} \to \Sigma_A^{\mathbb{N}}$ by $\sigma((x_i)_{i=1}^{\infty}) = (x_{i+1})_{i=1}^{\infty}$. $(\Sigma_A^{\mathbb{N}}, \sigma)$ is called a subshift of finite type. Let $P = (p_{ij})_{l\times l}$ be a stochastic matrix in the sense that $p_{ij} \geq 0$ and $\sum_{j=1}^{l} p_{ij} = 1$ for each i. We assume that $p_{ij} > 0$ iff $a_{ij} = 1$. Suppose $\vec{p} = (p_1, p_2, \cdots, p_l)$ is a probability vector with $p_i > 0$ for each i and $\vec{p}P = \vec{p}$. Then define μ on $\Sigma_A^{\mathbb{N}}$ by

$$\mu([i_1 i_2 \cdots i_n]) = p_{i_1} p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_{n-1} i_n},$$

for any $i_1 i_2 \cdots i_n \in \{1, 2, \cdots, l\}^n$ with $a_{i_k i_{k+1}} = 1$ for $k = 1, 2, \cdots, n-1$. μ is called a $(\vec{p}, P)$ Markov measure. μ is σ-invariant. Moreover μ is ergodic iff A is irreducible in the sense that there exists N, such that $A + A^2 + \cdots + A^N$ is strictly positive, equivalently for any pair $1 \leq i \leq j \leq l$, there exist $i_1, i_2, \cdots, i_k \in \{1, 2, \cdots, l\}$ such that $a_{ii_1} = a_{i_1 i_2} = \cdots = a_{i_k j} = 1$.

Example 4. (Continued fraction transformation).

Define $T : (0,1) \to (0,1)$ by $Tx = \frac{1}{x} - [\frac{1}{x}]$, where $[x]$ denotes the integral part of x. T is called the continued fraction transformation. Consider the continued fraction of a real number x,

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}, \ a_1, a_2, \cdots \in \mathbb{N}.$$

Notice that $a_1 = [\frac{1}{x}]$, $a_2 = [\frac{1}{\frac{1}{x} - [\frac{1}{x}]}] = [\frac{1}{Tx}]$, inductively $a_n = [\frac{1}{T^{n-1}x}]$. Now define a measure μ on $(0,1)$ by $\mu(B) = \frac{1}{\log 2} \int_B \frac{1}{x+1} dx$ for Borel set $B \subset (0,1)$. μ is called the Gaussian measure. μ is T-invariant and ergodic. See Pollicott and Yuri's book for a proof.

3.5 Mixing

Recall that a measure-preserving transformation T is ergodic if and only if for any $A, B \in \mathscr{B}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} m(T^{-k}A \cap B) = m(A)m(B).$$

We can change the way that the limit converges to give the following notions.

Definition 3.4. *Let $(X, \mathscr{B}, m, T)$ be a MPS.*

(i) We say T is weak-mixing if for any $A, B \in \mathscr{B}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |m(T^{-k}A \cap B) - m(A)m(B)| = 0.$$

(ii) Say that T is mixing (or strong-mixing) if for any $A, B \in \mathscr{B}$,

$$\lim_{n\to\infty} m(T^{-n}A \cap B) = m(A)m(B).$$

Remark: (1). In probability view, T is ergodic $\Leftrightarrow$ for any $A, B \in \mathscr{B}$, $T^{-n}A$ is independent form B on average. T is mixing $\Leftrightarrow$ for any $A, B \in \mathscr{B}$, $T^{-n}A$ is asymptotically independent form B.

(2). It is clear that mixing $\Rightarrow$ weak-mixing $\Rightarrow$ ergodicity.

Example 1. Let α be an irrational number. Let m be the Haar measure on $\mathbb{R}/\mathbb{Z}$. Define $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ by $Tx = x + \alpha (\bmod 1)$. Then T is ergodic but not weak-mixing. To see this, let $A = [0, \frac{1}{8}]$ and $B = [\frac{7}{8}, 1)$. Notice that for each k, $T^{-k}A = A - k\alpha(\mathrm{mod}1)$. Since $\{k\alpha(\mathrm{mod}1) : k \in \mathbb{N}\}$ is uniformly distributed on $[0,1)$, there are half of k such that $-k\alpha(\mathrm{mod}1) \in [0, \frac{1}{2})$, for such k, we have $A - k\alpha(\mathrm{mod}1) \subset [0, \frac{1}{8} + \frac{1}{2}]$ disjoint with B, therefore

$$\varliminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |m(T^{-k}A \cap B) - m(A)m(B)| \geq \frac{1}{2} m(A)m(B) > 0.$$

Hence T is not mixing.

Remark: There are examples of weak-mixing MPSs which are not mixing.

Just like the case of ergodicity, to check mixing property it is enough to consider a subcollection of $\mathscr{B}$ that generates $\mathscr{B}$. The following theorem can be proved in the same way as Theorem 3.10.

Theorem 3.11. *Let $(X, \mathscr{B}, m, T)$ be a MPS. Let $\mathscr{G}$ be a semi-algebra generating $\mathscr{B}$. Then*

(i) T is ergodic $\Leftrightarrow$ for any $A, B \in \mathscr{G}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} m(T^{-k}A \cap B) = m(A)m(B).$$

(ii) T is weak-mixing $\Leftrightarrow$ for any $A, B \in \mathscr{G}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |m(T^{-k}A \cap B) - m(A)m(B)| = 0.$$

(iii) T is mixing $\Leftrightarrow$ for any $A, B \in \mathscr{G}$,

$$\lim_{n\to\infty} m(T^{-n}A \cap B) = m(A)m(B).$$

Example 2. (Bernoulli shift on finite state space).

Let $\mathscr{G} = \{[i_1 i_2 \cdots i_k] : i_1 i_2 \cdots i_k \in \{1, 2, \cdots, l\}^k, k \in \mathbb{N}\}$, then $\mathscr{G}$ is a semi-algebra generating $\mathscr{B}$. Recall we have shown that for any $A, B \in \mathscr{G}$, $\mu(\sigma^{-n}A \cap B) = \mu(A)\mu(B)$ when n is large, hence σ is mixing.

Example 3. (Markov shift).

Let $(\vec{p}, P)$ be a Markov measure on $\Sigma_A^{\mathbb{N}}$. Then T is mixing $\Leftrightarrow$ T is weak-mixing $\Leftrightarrow$ P is primitive in the sense that there exists N such that P^N is strictly positive.

We can further characterize weak-mixing as follows.

Definition 3.5. *A subset J of $\mathbb{N}$ is said to have zero density in $\mathbb{N}$ if*

$$\lim_{n\to\infty} \frac{1}{n} \sharp(J \cap [0, n-1]) = 0.$$

For example $\{1, 2^2, 3^2, \cdots\}$ has zero density, the set of all primes has zero density.

Theorem 3.12. *Let $(X, \mathscr{B}, m, T)$ be a MPS. The following are equivalent.*

(i) T is weak-mixing.

(ii) For any $A, B \in \mathscr{B}$, there exists a subset $J = J(A, B)$ of $\mathbb{N}$ of density 0, such that

$$\lim_{J \not\ni n \to\infty} m(T^{-n}A \cap B) = m(A)m(B).$$

(iii) For any $A, B \in \mathscr{B}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |m(T^{-k}A \cap B) - m(A)m(B)|^2 = 0.$$

This theorem follows from the following lemma immediately.

Lemma 3.13. *Let $\{a_n\}$ be a bounded sequence of real numbers. The following are equivalent.*

(i) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |a_k| = 0.$

(ii) There exists a subset J of $\mathbb{N}$ of density 0 such that

$$\lim_{J\not\ni n\to\infty} a_n = 0.$$

(iii) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} a_k^2 = 0.$

Proof. (i) ⇒ (ii). For $k \in \mathbb{N}_+$, define $J_k = \{n \in \mathbb{N} : |a_n| \geq \frac{1}{k}\}$, clearly $J_1 \subseteq J_2 \subseteq \cdots$, we claim that each J_k is of density 0. Notice that

$$\sum_{j=0}^{n-1} |a_j| \geq \sum_{\substack{0\leq j\leq n-1\\ j\in J_k}} |a_j| \geq \sum_{\substack{0\leq j\leq n-1\\ j\in J_k}} \frac{1}{k} = \frac{1}{k}\sharp(J_k \cap [0, n-1]),$$

hence

$$0 = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} |a_j| \geq \frac{1}{k} \overline{\lim_{n\to\infty}} \frac{1}{n}\sharp(J_k \cap [0, n-1]).$$

Hence each J_k is of density 0. Therefore we can find a sequence of integers $0 = l_0 < l_1 < l_2 < \cdots$, such that

$$\frac{1}{n}\sharp(J_{k+1} \cap [0, n-1]) \leq \frac{1}{k+1} \text{ for any } n \geq l_k.$$

Now define

$$J = \bigcup_{k=0}^{\infty} (J_{k+1} \cap [l_k, l_{k+1})).$$

We claim that J has zero density. Let n be given, pick k such that $l_k \leq n < l_{k+1}$. Since $J_1 \subseteq J_2 \subseteq \cdots$, we have

$$J \cap [0, n-1] \subseteq \bigcup_{i=0}^{k} \big(J_{k+1} \cap [l_k, l_{k+1}) \cap [0, n-1]\big) \subseteq J_{k+1} \cap [0, n-1],$$

then

$$\frac{1}{n}\sharp\big(J \cap [0, n-1]\big) \leq \frac{1}{n}\sharp(J_{k+1} \cap [0, n-1]\big) \leq \frac{1}{k+1},$$

since as $n \to \infty$, $k \to \infty$, we see that J is of density 0. Let $n \notin J$ and $l_k \leq n < l_{k+1}$, then $n \notin J_{k+1}$, so $|a_n| < \frac{1}{k+1}$, hence $\lim_{J\not\ni n\to\infty} a_n = 0$. Notice that a similar argument yields (iii) ⇒ (ii). (ii) ⇒ (i) and (ii) ⇒ (iii) are straightforward. This completes the proof. □

One way to obtain new MPSs from old ones is to consider their product.

Definition 3.6. *Let* $(X_1, \mathscr{B}_1, m_1, T_1)$ *and* $(X_2, \mathscr{B}_2, m_2, T_2)$ *be two MPSs. Their product is denoted by* $(X_1 \times X_2, \mathscr{B}_1 \times \mathscr{B}_2, m_1 \times m_2, T_1 \times T_2)$, *where*

(i) $\mathscr{B}_1 \times \mathscr{B}_2$ *is the smallest* σ*-algebra containing all rectangles* $B_1 \times B_2$ *with* $B_1 \in \mathscr{B}_1, B_2 \in \mathscr{B}_2$.

(ii) $m_1 \times m_1$ *is the product probability measure.*

(iii) $T_1 \times T_2$ *is defined by* $(T_1 \times T_2)(x, y) := (T_1 x, T_2 y)$, *for* $(x, y) \in X_1 \times X_2$.

The fact $T_1 \times T_2$ is a measure-preserving transformation can be seen in the following way: Let $\mathscr{G} = \{B_1 \times B_2 : B_1 \in \mathscr{B}_1, B_2 \in \mathscr{B}_2\}$. Then $\mathscr{G}$ is a semi-algebra. One easily checks $T_1 \times T_2$ preserves measure of all rectangles in $\mathscr{G}$. Write $\mathscr{M} = \{A \in \mathscr{B}_1 \times \mathscr{B}_2 : (m_1 \times m_2)((T_1 \times T_2)^{-1}A) = (m_1 \times m_2)(A))\}$, then $\mathscr{M} \supseteq \mathscr{G}$ and $\mathscr{M}$ is a monotone class. Then by monotone class theorem, $\mathscr{M} = \mathscr{B}$.

The following theorem shows the connection between weak-mixing of T and the ergodicity of $T \times T$.

Theorem 3.14. *Let* $(X, \mathscr{B}, m, T)$ *be a MPS. The following are equivalent.*

(i) T *is weak-mixing.*

(ii) $T \times T$ *is ergodic.*

(iii) $T \times T$ *is weak-mixing.*

Proof. We show (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i). First consider (i) $\Rightarrow$ (iii). Let $A, B, C, D \in \mathscr{B}$. Since T is weak-mixing, there exist $J_1, J_2 \subseteq \mathbb{N}$ of density 0, such that

$$\lim_{J_1 \not\ni n \to \infty} m(T^{-n}A \cap C) = m(A)m(C) \text{ and } \lim_{J_2 \not\ni n \to \infty} m(T^{-n}B \cap D) = m(B)m(D).$$

Notice that

$$\begin{aligned}
&\lim_{\substack{n\to\infty \\ n \notin J_1 \cup J_2}} m \times m\big((T \times T)^{-n}(A \times B) \cap (C \times D)\big) \\
&= \lim_{\substack{n\to\infty \\ n \notin J_1 \cup J_2}} m \times m\big((T^{-n}A \cap C) \times (T^{-n}B \cap D)\big) \\
&= \lim_{\substack{n\to\infty \\ n \notin J_1 \cup J_2}} m(T^{-n}A \cap C)m(T^{-n}B \cap D) \\
&= m(A)m(B)m(C)m(D) \\
&= (m \times m)(A \times B)(m \times m)(C \times D).
\end{aligned}$$

Since $J_1 \cup J_2$ is of density 0, by Theorem 3.12 $T \times T$ is weak-mixing. (iii) $\Rightarrow$ (ii) is trivial. Now consider (ii) $\Rightarrow$ (i). Let $A, B \in \mathscr{B}$. Since $T \times T$ is ergodic, we have

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n-1} m(T^{-k}A \cap B) &= \frac{1}{n}\sum_{k=0}^{n-1} m \times m\big((T \times T)^{-k}(A \times X) \cap (B \times X)\big) \\
&\to m(A)m(B), \text{ as } n \to \infty,
\end{aligned}$$

and

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1} m(T^{-k}A\cap B)^2 &= \frac{1}{n}\sum_{k=0}^{n-1} m\times m\big((T^{-k}A\cap B)\times(T^{-k}A\cap B)\big)\\ &= \frac{1}{n}\sum_{k=0}^{n-1} m\times m\big(T^{-k}(A\times A)\cap(B\times B)\big)\\ &\to m(A)^2 m(B)^2,\ \text{as } n\to\infty.\end{aligned}$$

Hence we have

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n-1}\big(m(T^{-k}A\cap B)-m(A)m(B)\big)^2 &= \frac{1}{n}\sum_{k=0}^{n-1} m(T^{-k}A\cap B)^2\\ &\quad - 2m(A)m(B)\big[\frac{1}{n}\sum_{k=0}^{n-1} m(T^{-k}A\cap B)\big] + m(A)^2 m(B)^2\\ &\to 0,\ \text{as } n\to\infty.\end{aligned}$$

By Theorem 3.12, T is weak-mixing.

□