Ergodic Theory & Entropy:

26/10/2017

Lecture 1:

webpage: http://www.math.tau.ac.il/~barakw/entropy

Books: General introduction to dynamical systems

Hassel Blatt - Katok

Einsiedle - Ward

Einsiedle - Lindenstrauss - Ward - "The entropy book"

Crash intro. to dynamical systems:

1) Abstract development

Top. dynamics (X, T)

compact topological space

continuous map $X \to X$

Let (X_1, T_1), (X_2, T_2) & $F: X_1 \to X_2$ continuous. F is a factor map if $F \circ T_2 = T_1 \circ F$

$$\begin{tikzcd} X_1 \arrow[r, "T_1"] \arrow[d, "F"'] & X_1 \arrow[d, "F"] \\ X_2 \arrow[r, "T_2"] & X_2 \end{tikzcd}$$

← commutative diagram

Questions:

1) מיון המרחבים עד כדי איזומורפיזם

2) which systems are factors of which systems?

3) given (X, T), what are all the isomorphisms $X \to X$

if T is a homomorphism, let us define

$$n\,x = T^n(x) = T(T \dots (T(x)) \dots)$$

this is an action of $\mathbb{Z}$ on X.

Ergodic Theory:

$(X, \mathcal{B}, \mu, T)$ p.p.s

set X — measure σ-alg on X; probability measure on $\mathcal{B}$; probability preserving transformation

$T^{-1}(\mathcal{B}) \subseteq \mathcal{B}$ ← T is measurable

$\forall A \in \mathcal{B} \quad \mu(T^{-1}(A)) = \mu(A)$

Let $(X_1, \mathcal{B}_1, \mu_1, T_1)$, $(X_2, \mathcal{B}_2, \mu_2, T_2)$ two p.p.s's.

& $F: X_1' \to X_2$, $\mu(X_1 \setminus X_1') = 0$, $X_1' \subseteq X_1$, $F^{-1}(\mathcal{B}_2) \subseteq \mathcal{B}_1$,

$F_* \mu_1 = \mu_2$, meaning: $F_* \mu_1(A) = \mu_1(F^{-1}(A)) = \mu_2(A)$ &

$T_2 \circ F = F \circ T_1$:

$$\begin{array}{ccc} X_1 & \xrightarrow{T_1} & X_1 \\ F \downarrow & & \downarrow F \\ X_2 & \xrightarrow[T_2]{} & X_2 \end{array}$$

← commutative.

then F is a 'Factor'. If F is also invertible then F is an isomorphism

Questions:

1) Classifying all p.p.s's up to Isomorphism
2) Which p.p.s's are Factors of which p.p.s's ?
3) Self-isomorphisms for a given p.p.s

Standard Lebeg Space:

Theorem:

If X localy compact space, $\mathcal{B}_0$ - Borel σ-algebra, μ non-atomic probability measure on $(X, \mathcal{B}_0)$. ($\mu(\{x_0\}) = 0 \quad \forall x_0 \in X$ - non atomic)

Let $\mathcal{B}$ be the completion of $\mathcal{B}_0$. then there exists $X' \subseteq X$ s.t. $\mu(X \setminus X') = 0$ & $f: X' \to [0,1]$ injective & measurable s.t. $F_* \mu$ is the Lebeg measure on $[0,1]$. $F_* \mu(A) = \mu(F^{-1}(A))$

Motivation:

n-body-Problem:

Let there be n-bodies - point masses in $\mathbb{R}^3$ with masses $m_1, \dots, m_n$ with initial velocities & initial positions. $\vec{v}_1, \dots, \vec{v}_n$. These bodies apply on one another gravitational forces by Newton's Gravity Law & they move by Newton's Second Law.

The question: How does this system work as a function of time? (There is a solution for $n=2$, but for $n=3$ it starts getting complicated).

3 problems in number-theory:

1) $\pi = 3.141\dots$ contains all finite words in $\{0, \dots, 9\}^*$

2) $x = 2^{\frac{1}{3}}$, $x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 \ddots}}}$. Do all a_i's appear?

Are the a_i bounded?

3) $\forall \alpha, \beta \in \mathbb{R}, \forall \varepsilon > 0 \;\; \exists n \in \mathbb{N}$ s.t. $\langle n\alpha \rangle \langle n\beta \rangle < \frac{\varepsilon}{n}$

where $\langle x \rangle = \text{dist}(x, \mathbb{Z}) = \min\limits_{p \in \mathbb{Z}} |p - x|$

Motivations for entropy which we hope to prove:

1) μ = Coin tossing measure on Cantor's middle thirds set. Then μ is base 2 generic

$$C = \left\{ \sum_{i=1}^{\infty} \frac{a_i}{3^i} \;\middle|\; a_i \in \{0,2\} \right\}$$

toss a coin - if heads then put 0, if tails put 2 (toss a coin to choose digits 0, 2 in base 3).

X is base d dense if every word appears in its base d expansion.

Entropies of partitions:

Given $(X, \mathcal{B}, \mu)$

set — sigma algebra in 2^X — $\mu : \mathcal{B} \to [0,1]$

$\xi = \{A_1, A_2, \dots\} = \{A_i \mid i \in l\}$ is called a partition

if l is either countable or finite &

$$\begin{cases} i \neq j \implies A_i \cap A_j = \phi \\ \bigcup_{i \in l} A_i = X \end{cases}$$

$\xi = \{A_i \mid i \in l\}$ & $\eta = \{B_j \mid j \in l\}$ are equivalent

(sometimes equivalent mod 0, or equivalent mod μ)

if $\forall i \; \exists j$. s.t $\mu(A_i \triangle B_j) = 0$ $\qquad A_1 \triangle A_2 = (A_1 \setminus A_2) \cup (A_2 \setminus A_1)$

The atom of x w.r.t ξ is A_i, where $x \in A_i$.

Notation: $[x]_\xi$ $\leftarrow$ here the number of atoms is countable (it is l)

$\sigma(\xi) = \sigma$ algebra generated by ξ $=$ maybe

$=$ intersection of all σ-algebras containing ξ $=$

~~$= \left\{ \bigcap_i C_i^{\tau} \;\middle|\; \tau : l \to \{0,1\} \text{ and if } \tau(i) = 0, \; C_i^{(\tau)} = A_i \text{ if } \tau(i) = 1, \; C_i^{(\tau)} = X \setminus A_i \right\}$~~ ~~$X \setminus \bigcap_{i \in l} C_i^{\tau}$~~

more generally, for $\mathcal{B}_1 \subset 2^X$, $\sigma(\mathcal{B}_1)$ = intersection of all σ-algebras containing $\mathcal{B}_1$

Defenition:

$$I_\mu(x) = -\log(\mu([x]_\xi))$$

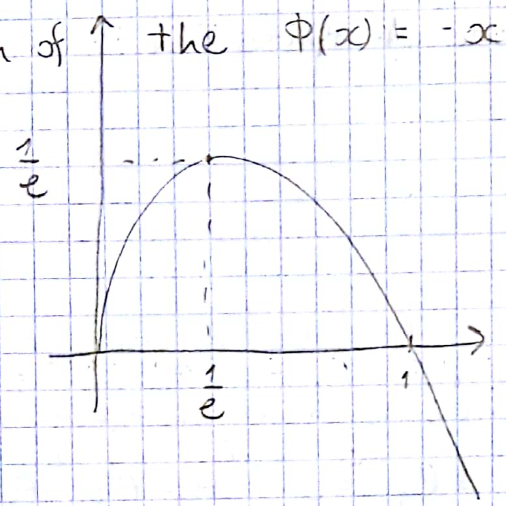
note: $\mu([x]_\xi)$ could be zero, this will have no effect, we can define $I_\mu(x) = 0$ when $\mu([x]_\xi) = 0$

Definition: (the entropy of ξ w.r.t μ)

$$H_\mu(\xi) = \int_X I_\mu(x)\, d\mu(x) = -\sum_i \mu(A_i)\log(\mu(A_i))$$

$$H(p_1, p_2, \dots) = -\sum_i p_i \log(p_i)$$

the graph of the $\phi(x) = -x\log x$ function:



If $\xi = \{A_i \mid i \in \ell\}$, $\eta = \{B_j \mid j \in \ell\}$ are two partitions, $\xi \vee \eta = \{A_i \cap B_j \mid (i,j) \in \ell \times \ell\}$

Note: $\sigma(\xi \vee \eta) = \sigma(\sigma(\xi) \cup \sigma(\eta))$ ← this is an exercise

If $\xi_1, \xi_2, \dots$ are partitions, then:

$$\bigvee_{i=1}^{\infty} \xi_i = \sigma\left(\bigcup_{i=1}^{\infty} \sigma(\xi_i)\right)$$

(the join of a finite amount of partition is a partition but the join of an infinite number of partition is defined as a σ-algebra)

$\xi_n \nearrow_{n\to\infty} \mathcal{B}$ means $\mathcal{B} = \bigvee_n \xi_n$ and $\sigma(\xi_n) \subset \sigma(\xi_{n+1})$ for all n,

$$T^{-1}\xi = \{T^{-1}A_i \mid i \in \ell\}$$

Definition: (The conditional entropy of ξ w.r.t μ conditioned on η)

$$H_\mu(\xi|\eta) = \sum_j \mu(B_j) H\underbrace{\left(\frac{\mu(A_1\cap B_j)}{\mu(B_j)}, \frac{\mu(A_2\cap B_j)}{\mu(B_j)}, \dots\right)}_{(*)} =$$

$$= \sum_{j\in J} \mu(B_j) H_{\mu|_{B_j}}(\xi)$$

where $\mu|_{B_j}(A) = \frac{1}{\mu(B_j)} \mu(B_j \cap A) \leftarrow$ normalized restriction of μ to B_j

if $\mu(B_j) = 0$, the respective element in $(*)$ is disregarded.

Informally:

if $(X, \mathcal{B}, \mu)$ is a space on which we perform an experiment, then $\xi = (A_i)$ is a list of possible outcomes. $H_\mu(\xi)$ measures the amount of information gained on average by performing the experiment.

$H_\mu(\xi|\eta)$ is the average amount of information gained by performing ξ, given that we've already performed η.

$k \in \mathbb{N}$, $\Delta_k = \{(p_1, \dots, p_k) \mid \sum p_i = 1, \ p_i \geq 0 \ \forall i\}$

$\Delta_\infty = \{(p_1, \dots) \mid \sum p_i = 1, \ p_i \geq 0 \ \forall i\}$

$\Delta = \bigcup_{k\in\mathbb{N}} \Delta_k$

$H : \Delta \longrightarrow \mathbb{R}$

$H(p_1, p_2, \dots, p_k) = -\sum_{i=1}^{k} p_i \log(p_i)$

Properties:

1) $H(p_1, \dots, p_k) \geq 0$, $H(p_1, \dots, p_k) = 0 \iff$ one of the p_i is 1
2) $H(p_1, \dots, p_k, 0) = H(p_1, \dots, p_k)$
3) H is continuous, invariant under permuting the p_i & achieves a max on Δ_k of $\log k$ exactly on $p_1 = \dots = p_k = \frac{1}{k}$.
4) $H_\mu(\xi \vee \eta) = H_\mu(\eta) + H_\mu(\xi|\eta)$

Theorem:

H is the unique function satisfying ① - ④

Proof of the properties:

1) is clear from graph of ϕ

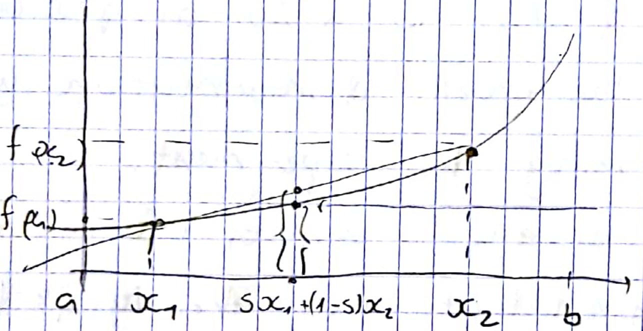
2) clear from definition.

for 3), we will need to discuss convex functions

Definition:

$f: (a,b) \to \mathbb{R}$ is convex if for all $x_1, x_2 \in (a,b)$, & all $s \in [0,1]$:

$$f(sx_1 + (1-s)x_2) \leq s f(x_1) + (1-s) f(x_2) \qquad (*)$$



if f is C^2 & $f''(x) \geq 0$ for all $x \in (a,b)$ then f is convex.

f is strictly convex if equality in $(*)$ implies either $x_1 = x_2$ or $s \in \{0,1\}$

Note: $-\log x$ & $-\phi$ are convex.

Jensens inequality:

if f is convex & μ is a probability measure on (a,b) then:

$$f\left(\int x \, dx\right) \leq \int f(x) \, d\mu \qquad (**)$$

Note: $(*)$ is a special case of $(**)$, with $\mu = s\delta_{x_1} + (1-s)\delta_{x_2}$

furthermore, if f is strictly convex then equality in $(**)$ implies $\mu = \delta_x$.

Proof of 3) :

$$\frac{1}{k} H(p_1, \dots, p_k) = -\frac{1}{k} \sum p_i \log(p_i) = \sum \frac{1}{k} \phi(p_i) \underset{\uparrow}{\leq} \phi\left(\sum_i \frac{1}{k} p_i\right) =$$

Jensen, $-\phi$ is convex

$$= \phi\left(\frac{1}{k}\right) = -\frac{1}{k} \log \frac{1}{k} = \frac{1}{k} \log k$$

multiply through by k, get $H(p_1, \dots, p_k) \leq \log k$, by strict convexity, only have equality when $p_1 = \dots = p_k$ which implies $p_i = \frac{1}{k}$ $\forall i$.