Math 6082 2025-01-24.

§3 Statistical behavior of orbits and ergodic theorems

§3.1 Asymptotical distribution and invariant measures.

Let (X, T) be a TDS, and let $B \subset X$ be a Borel set.

For $x \in X$, we would ask with what frequency do the elements of the set $\{x, Tx, T^2x, \cdots\}$ lie in the set B?

More precisely, this frequency is defined by

$$F_B(T, x) = \lim_{n\to\infty} \frac{1}{n} \#\{0 \le i \le n-1 : T^i x \in B\}$$

if the limit exists.

Let χ_B denote the characteristic function of B, i.e.

$$\chi_B(x) = \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{if } x \notin B. \end{cases}$$

Notice that $\#\{0 \le i \le n-1 : T^i x \in B\} = \sum_{i=0}^{n-1} \chi_B(T^i x)$.

So

$$F_B(T,x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_B(T^i x).$$

The above expression is called the time average or Birkhoff average of the function χ_B.

It is natural and more convenient to consider the Birkhoff average of continuous functions rather that characteristic functions.

Let $x \in X$. Assume that for any cts function φ on X,

$$I_x(\varphi) := \lim_{n\to\infty} \frac{1}{n} \left(\varphi(x) + \varphi(Tx) + \cdots + \varphi(T^{n-1}x) \right)$$

exists.

Let $C(X)$ be the space of all cts functions on X with the uniform topology. Then the time averge

$$I_x : C(X) \to \mathbb{R}$$

satisfies the following properties:

(1) Linearity $\left(I_x(\alpha\varphi+\beta\psi) = \alpha I_x(\varphi) + \beta I_x(\psi) \right)$

(2) Boundness $|I_x(\varphi)| \leq \sup_{y\in X} |\varphi(y)|$.

(3) positivity $I_x(\varphi) \geq 0$ if $\varphi \geq 0$.

(4) Invariance $I_x(\varphi) = I_x(\varphi\circ T)$.

According to (1), (2), (3), I_x is a bdd linear functional on $C(X)$. Hence by Riesz representation thm, $\exists$ a unique Borel prob. measure μ on X such that

$$I_x(\varphi) = \int \varphi \, d\mu, \qquad \varphi \in C(X).$$

Furthermore (4) implies that

$$\int \varphi\circ T \, d\mu = \int \varphi \, d\mu,$$

which implies that μ is T-invariant, i.e. $\mu = \mu\circ T^{-1}$.

Question: ① Are there points $x \in X$ such that I_x exists?

② For any T-invariant prob. measure μ, is there a point x such that

$$\int \phi \, d\mu = I_x(\varphi) \quad \text{for all } \varphi \in C(X) \, ?$$

The questions will be answered by a combination of two fundamental thms, one in TDS, one in ergodic theory.

§ 3.2 Existence of invariant measures.

Thm 3.1 (Krylov-Bogolubov Thm) Let (X, T) be a TDS. Then $\exists$ at least one T-invariant measure on X.

Pf. Let $y \in X$. Take a dense countable set

$$\{\varphi_1, \varphi_2, \cdots\} \subset C(X)$$

For each $m \in \mathbb{N}$, the sequence

$$\frac{1}{n} \sum_{i=0}^{n-1} \varphi(T^i y)$$

is bdd. Hence it contains a convergent subsequence.

By the diagonal process, one can find a subsequence (n_j) such that

$$\lim_{j\to\infty} \frac{1}{n_j} \sum_{i=0}^{n_j-1} \varphi_m(T^i y) := J(\varphi_m)$$

Then a standard approximation argument shows that

$$\lim_{j\to\infty} \frac{1}{n_j} \sum_{i=0}^{n_j-1} \varphi(T^i x) := J(\varphi)$$

for all $\varphi \in C(X)$. Then

$$J : C(X) \to \mathbb{R}$$

is a bdd positive linear functional. It is also invariant,

i.e. $J(\varphi) = J(\varphi \circ T)$.

By Riesz representation Thm, $\exists$ a Borel probability measure μ such that

$$J(\varphi) = \int \phi \, d\mu, \quad \forall \ \varphi \in C(X).$$

Since $J(\varphi) = J(\varphi \circ T)$, $\int \phi \circ T \, d\mu = \int \phi \, d\mu$

implying $\mu = \mu \circ T^{-1}$. ☒

§ 3.3. Birkhoff ergodic Thm.

Thm 3.2 (Birkhoff, 1931). Let $(X, \mathcal{B}, \mu)$ be a prob. space and $T: X \to X$ preserves μ.

Let $f \in L^1(\mu)$. Then

$$\frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$$

converges a.e. to a function $f^* \in L^1(\mu)$. Furthermore

$$\int f^* \, d\mu = \int f \, d\mu.$$

The proof of Birkhoff ergodic Thm is based on the following.

Thm 3.3 (Maximal ergodic Thm).

Let $f \in L^1(\mu)$. Set $f_0 = 0$ and $f_n(x) = f(x) + f(Tx) + \cdots + f(T^{n-1}x)$.

Let $F_N(x) = \max_{0 \le n \le N} f_n(x)$.

Then $$\int_{\{x : F_N(x) > 0\}} f \, d\mu \ge 0 \quad \text{for all } N \ge 1.$$

Pf.

$$F_N(Tx) + f(x)$$

$$= f(x) + \max_{0 \le n \le N} \{0, f(Tx), f(Tx)+f(T^2x), \cdots, f(Tx)+\cdots+f(T^N x)\}$$

$$= \max_{0 \le n \le N} \{f(x), f(x)+f(Tx), \cdots, f(x)+f(Tx)+\cdots+f(T^N x)\}$$

$$= \max_{1 \le n \le N+1} f_n(x)$$

$$\ge \max_{1 \le n \le N} f_n(x)$$

$$\ge F_N(x) \quad \text{on} \quad A := \{x : F_N(x) > 0\}.$$

Hence

$$f(x) \ge F_N(x) - F_N(Tx) \quad \text{on } A.$$

Thus

$$\int_A f \, d\mu \ge \int_A F_N(x) \, d\mu - \int_A F_N(Tx) \, d\mu$$

$$= \int_X F_N(x) \, d\mu - \int_A F_N(Tx) \, d\mu$$

(since $F_N(x) = 0$ on $X \setminus A$)

$$= \int_X F_N(Tx) \, d\mu - \int_A F_N(Tx) \, d\mu$$

$$= \int_{X \setminus A} F_N(Tx) \, d\mu \ge 0. \quad (\text{since } F_N \ge 0).$$

$\boxtimes$

Corollary 3.4 For any $g \in L^1(\mu)$ and $\alpha \in \mathbb{R}$, set

$$B_\alpha = \left\{ x : \sup_{n \geq 1} \frac{1}{n} \sum_{k=0}^{n-1} g(T^k x) > \alpha \right\}.$$

Then
$$\int_{B_\alpha} g \, d\mu \geq \alpha \cdot \mu(B_\alpha).$$

Moreover, if $A \in \mathcal{B}$ with $T^{-1}A = A$, then

$$\int_{B_\alpha \cap A} g \, d\mu \geq \alpha \cdot \mu(B_\alpha \cap A).$$

Pf. Set $f = g - \alpha$. Then $B_\alpha = \bigcup_{N=0}^{\infty} \{x : F_N(x) > 0\}$

where $F_N(x) = \sup\{0, f(x), f(x) + f(Tx), \dots, f(x) + \dots + f(T^{N-1}x)\}$

Notice that the sets

$\{x : F_N(x) > 0\}$ are increasing

Hence

$$\int_{B_\alpha} g \, d\mu = \int_{B_\alpha} f \, d\mu + \alpha \mu(B_\alpha)$$

$$= \lim_{N \to \infty} \int_{\{x : F_N(x) > 0\}} f \, d\mu + \alpha \mu(B_\alpha)$$

$$\geq \alpha \mu(B_\alpha).$$

When $A = T^{-1}A$, change X by A (i.e. Applying the result to $T|A$).

We obtain $\int_{B_\alpha \cap A} f \geq \alpha \mu(B_\alpha \cap A)$. ∎

Pf of the Birkhoff ergodic Thm :

WLOG, we may assume that f only takes real values.

Define $f^*(x) = \varlimsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$

$f_*(x) = \varliminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$.

Clearly, $f^*(x) = f^*(Tx)$, $f_*(x) = f_*(Tx)$.

For $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$, define

$$E_{\alpha,\beta} = \{x : f_*(x) < \alpha < \beta < f^*(x)\}.$$

Then $E_{\alpha,\beta} = T^{-1} E_{\alpha,\beta}$.

Write $B_\beta = \{x : \sup_{n\geq 1} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) > \beta\}$.

Then $E_{\alpha,\beta} \cap B_\beta = E_{\alpha,\beta}$

Hence from the corollary

$$\int_{E_{\alpha,\beta}} f \, d\mu = \int_{E_{\alpha,\beta} \cap B_\beta} f \, d\mu \geq \beta \mu(E_{\alpha,\beta} \cap B_\beta)$$
$$= \beta \mu(E_{\alpha,\beta}).$$

Replace f by $-f$, then

$$E_{\alpha,\beta} = \{x : (-f)_*(x) < -\beta < -\alpha < (-f)^*(x)\}.$$

Similarly we have

$$\int_{E_{\alpha,\beta}} (-f)\,d\mu \geq (-\alpha)\,\mu(E_{\alpha,\beta})$$

i.e. $$\int_{E_{\alpha,\beta}} f\,d\mu \leq \alpha\,\mu(E_{\alpha,\beta}).$$

Hence we obtain $\quad \beta\,\mu(E_{\alpha,\beta}) \leq \alpha\,\mu(E_{\alpha,\beta}).$

As $\alpha<\beta$, this implies that $\mu(E_{\alpha,\beta})=0$.

Notice that

$$\{x : f_*(x) < f^*(x)\} = \bigcup_{\substack{\alpha,\beta\in\mathbb{Q} \\ \alpha<\beta}} E_{\alpha,\beta}$$

So $\quad \mu\{x : f_*(x) < f^*(x)\} = 0.$

Next we prove $\quad f^* \in L^1(\mu)$.

Set $\quad g=|f|$. clearly $\quad |f^*| \leq g^*$.

It suffices to show $g^* \in L^1(\mu)$.

As $\frac{1}{N}\sum_{k=0}^{N-1} g(T^k x) \to g^*$ μ-a.e,

by Fatou's lem,

$$\int g^* d\mu = \int \varliminf_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} g(T^k x)\, d\mu$$
$$\le \varliminf_{N\to\infty} \frac{1}{N}\int \sum_{k=0}^{N-1} g(T^k x)\, d\mu$$
$$= \int g\, d\mu < \infty.$$

Finally we show

$$\int_A f^* d\mu = \int_A f\, d\mu \quad \text{for all } A\in\mathcal{B} \text{ with } T^{-1}A = A.$$

Fix A with $T^{-1}A = A$. Define

$$D_n^k = \left\{x : \frac{k}{n} < f^*(x) \le \frac{k+1}{n}\right\}, \quad n\ge 1, \quad k\in\mathbb{Z}.$$

Recall $B_\alpha := \left\{x : \sup_{n\ge1} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) > \alpha\right\}$.

Then for any $\varepsilon > 0$,

$$\int_{D_n^k\cap A} f\, d\mu = \int_{D_n^k\cap A\cap B_{\frac{k}{n}-\varepsilon}} f\, d\mu \ge \left(\frac{k}{n}-\varepsilon\right)\mu\left(D_n^k\cap A\right).$$

(by Corollary 3.4).

Letting $\varepsilon \to 0$ gives

$$\int_{D_n^k \cap A} f \, d\mu \geq \frac{k}{n} \mu(D_n^k \cap A).$$

However $\int_{D_n^k \cap A} f^* \, d\mu \leq \frac{k+1}{n} \mu(D_n^k \cap A)$.

Summing over k yields

$$\int_A f^* d\mu \leq \int_A f \, d\mu + \frac{1}{n} \mu(A).$$

Letting $n \to \infty$ gives $\int_A f^* d\mu \leq \int_A f \, d\mu$

Similarly $\int_A (-f)^* d\mu \leq \int_A -f \, d\mu$

$$\Rightarrow \int_A f_* \, d\mu \geq \int_A f \, d\mu.$$

Since $f^* = f_*$ a.e., $\int_A f^* d\mu = \int_A f_* \, d\mu \geq \int_A f \, d\mu$.

Thus, $\int_A f_* \, d\mu = \int_A f \, d\mu$. $\blacksquare$