Birkhoff's Ergodic Theorem

In this lecture we prove Birkhoff's Ergodic Theorem (also called the Pointwise Ergodic Theorem):

Theorem: Let μ be a probability measure and $\psi \in L^1(\mu)$. Then the **ergodic average**

$$\psi^*(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \psi \circ T^i(x)$$

exists μ-a.e., and ψ^* is T-invariant, i.e., $\psi^* \circ T = \psi^*$ μ-a.e.
If in addition μ is ergodic then $\psi^* = \int_X \psi \, d\mu$ for μ-a.e. x

Or in short:

Space Average = Time Average (for typical points).

Proof of BET. Step 1

Maximal Ergodic Theorem Let $(X, T, \mathcal{B}, \mu)$ be a probability measure preserving dynamical system. Take $M_N = \max\{S_n : 0 \leq n \leq N\}$. Then

$$\int_{A_N} f \, d\mu \geq 0 \qquad \text{for } A_N = \{x \in X : M_N(x) > 0\}.$$

The Koopman operator $U_T : L^1(\mu) \to L^1(\mu)$ is defined as

$$U_T f = f \circ T.$$

Clearly U_T is linear and positive, i.e., $f \geq 0$ implies $U_T f \geq 0$.

We write the ergodic sum as

$$S_n = S_n f = \sum_{k=0}^{n-1} f \circ T^k \text{ and } S_0 \equiv 0.$$

Proof of BET. Step 1

Proof of the Maximal Ergodic Theorem: Clearly $M_N \geq S_n$ for all $0 \leq n \leq N$ and by positivity of the Koopman operator, also $U_T M_N \geq U_T S_n$. Add f: $U_T M_N + f \geq U_T S_n + f = S_{n+1}$. For $x \in A_N$, this means

$$\begin{aligned} U_T M_N(x) + f(x) &\geq \max_{1 \leq n \leq N} S_n(x) \\ &\geq_{x \in A_N} \max_{0 \leq n \leq N} S_n(x) = M_N(x). \end{aligned}$$

Therefore $f \geq M_N - U_T M_N$ on A_N, and (since $M_N \geq S_0 = 0$)

$$\begin{aligned} \int_{A_N} f \, d\mu &\geq \int_{A_N} M_N \, d\mu - \int_{A_N} U_T M_N \, d\mu \\ &= \int_X M_N \, d\mu - \int_{A_N} U_T M_N \, d\mu \\ &= \int_X M_N \, d\mu - \int_X U_T M_N \, d\mu = 0. \end{aligned}$$

This completes the proof.

Proof of BET. Step 2

Lemma: Let $(X, T, \mathcal{B}, \mu)$ be a probability measure preserving dynamical system, and $E \subset X$ a T-invariant subset. Let

$$B_\alpha := \{x \in X : \sup_n \frac{1}{n} S_n g(x) > \alpha\}.$$

Then

$$\int_{B_\alpha \cap E} g \; d\mu \geq \alpha \mu(B_\alpha \cap E).$$

For a measurable set $E \subset X$, we define a probability measure

$$\mu_E(B) = \frac{1}{\mu(E)} \mu(B \cap E)$$

If E is T-invariant then μ_E is T-invariant too.

Proof of BET. Step 2

Proof: If $\mu(E) = 0$ then there is nothing to prove. So assume that $\mu(E) > 0$. Take $f = g - \alpha$, so

$$B_\alpha = \cup_N A_N \quad \text{for} \quad A_N = \{x \in X : M_N(x) > 0\}.$$

Note also that $A_N \subset A_{N+1}$ for all N. Therefore for each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\int_{B_\alpha} f \, d\mu_E \geq \int_{A_N} f \, d\mu_E \geq -\varepsilon.$$

Since ε is arbitrary, $\int_{B_\alpha} f \, d\mu_E \geq 0$. Adding α again we have

$$\int_{B_\alpha} g \, d\mu_E = \int_{B_\alpha} f + \alpha \, d\mu_E \geq \alpha \mu_E(B_\alpha \cap E).$$

Multiply everything by $\mu(E)$ to get the lemma.

Proof of BET. Step 3

Proof of Birkhoff's Ergodic Theorem:

Recall $\psi \in L^1(\mu)$. Define

$$\overline{\psi} = \limsup_{n\to\infty} \frac{1}{n} S_n\psi \quad \text{and} \quad \underline{\psi} = \liminf_{n\to\infty} \frac{1}{n} S_n\psi.$$

Since

$$\left|\frac{n+1}{n}\frac{1}{n+1}S_{n+1}\psi - \frac{1}{n}S_n\psi \circ T\right| = \frac{1}{n}|\psi(x)| \to 0$$

as $n \to \infty$, we have $\overline{\psi} \circ T = \overline{\psi}$ and similarly $\underline{\psi} \circ T = \underline{\psi}$.

We want to show that $\overline{\psi} = \underline{\psi}$ μ-a.e.

Proof of BET. Step 3

Let

$$E_{\alpha,\beta} = \{x \in X : \underline{\psi}(x) < \beta, \alpha < \overline{\psi}(x)\}$$

Then $E_{\alpha,\beta}$ is T-invariant, and

$$\{x \in X : \underline{\psi}(x) < \overline{\psi}(x)\} = \bigcup_{\alpha,\beta\in\mathbb{Q},\beta<\alpha} E_{\alpha,\beta}.$$

This is a countable union, and therefore it suffices to show that $\mu(E_{\alpha,\beta}) = 0$ for every pair of rationals $\beta < \alpha$.

Write $B_\alpha := \{x \in X : \sup_n \frac{1}{n} S_n\psi(x) > \alpha\}$ as in our Lemma. Since $E_{\alpha,\beta} = E_{\alpha,\beta} \cap B_\alpha$, this Lemma gives

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

Proof of BET. Step 3

From the previous slide:

$$\int_{E_{\alpha,\beta}} \psi \, d\mu \geq \alpha\mu(E_{\alpha,\beta}).$$

We repeat this argument replacing ψ, α, β by $-\psi, -\alpha, -\beta$. Note that $\overline{-\psi} = -\underline{\psi}$ and $\underline{-\psi} = -\overline{\psi}$. This gives

$$\int_{E_{\alpha,\beta}} \psi \, d\mu \leq \beta\mu(E_{\alpha,\beta}).$$

Since $\beta < \alpha$, this can only be true if $\mu(E_{\alpha,\beta}) = 0$.

Therefore $\overline{\psi} = \underline{\psi} = \psi^*$, i.e., the $\limsup$ and $\liminf$ are actually limits μ-a.e.

Proof of BET. Step 4

The next step is to show that $\psi^* \in L^1(\mu)$. From Measure Theory we have:

Fatou's Lemma If $(g_n)_{n\in\mathbb{N}}$ are non-negative $L^1(\mu)$-functions, then

$$\liminf_n g_n \in L^1(\mu) \text{ and } \int_X \liminf_n g_n \, d\mu \le \liminf_n \int_X g_n \, d\mu.$$

Here we apply this to $g_n = |\frac{1}{n} S_n\psi|$, which belong to $L^1(\mu)$ because (by T-invariance)

$$\int_X |\frac{1}{n} S_n\psi| \, d\mu \le \frac{1}{n} \sum_{k=0}^{n-1} \int_X |\psi| \circ T^k \, d\mu = \int |\psi| \, d\mu < \infty.$$

Hence in the limit: $\int_X |\overline{\psi}| \, d\mu \le \liminf_n \int_X |\psi| \, d\mu < \infty$.

Proof of BET. Step 5

Next, we need to show that $\int_X \psi^* \, d\mu = \int_X \psi \, d\mu$ (so without absolute value signs). Take

$$D_{k,n} = \{x \in X : \frac{k}{n} \leq \psi^*(x) < \frac{k+1}{n}\}.$$

Then $D_{k,n}$ is T-invariant, and $\cup_{k \in \mathbb{Z}} D_{k,n} = X \bmod \mu$. Also $B_{\frac{k}{n}-\varepsilon} \cap D_{k,n} = D_{k,n}$ for $\varepsilon > 0$. Therefore our Lemma gives

$$\begin{aligned}
\int_{D_{k,n}} \psi \, d\mu &= \int_{B_{\frac{k}{n}-\varepsilon} \cap D_{k,n}} \psi \, d\mu \\
&\geq (\frac{k}{n} - \varepsilon)\mu(B_{\frac{k}{n}-\varepsilon} \cap D_{k,n}) \\
&= (\frac{k}{n} - \varepsilon)\mu(D_{k,n}).
\end{aligned}$$

Proof of BET. Step 5

Since ε is arbitrary, we have $\frac{k}{n}\mu(D_{k,n}) \le \int_{D_{k,n}} \psi \, d\mu$. Therefore

$$\int_{D_{k,n}} \psi^* \, d\mu \le \frac{k+1}{n}\mu(D_{k,n}) \le \frac{1}{n}\mu(D_{k,n}) + \int_{D_{k,n}} \psi \, d\mu.$$

Summing over all $k \in \mathbb{Z}$, we find $\int_X \psi^* \, d\mu \le \frac{1}{n} + \int_X \psi \, d\mu$. Since $n \in \mathbb{N}$ is arbitrary, also

$$\int_X \psi^* \, d\mu \le \int_X \psi \, d\mu,$$

By the same argument for $-\psi$, we find $\int_X \psi^* \, d\mu \ge \int_X \psi \, d\mu$. Hence

$$\int_X \psi^* = \int_X \psi \, d\mu.$$

Finally, if μ is ergodic, the T-invariant function ψ^* has to be constant μ-a.e., so $\psi^* = \int \psi \, d\mu$. This completes the proof of Birkhoff's Ergodic Theorem.

The L^p Ergodic Theorem

The L^p Ergodic Theorem is a generalisation of Von Neumann's L^2 version of the Ergodic Theorem, which predates[1] Birkhoff's Theorem, but nowadays, it is usually proved as a corollary of the pointwise ergodic theorem.

Theorem: Let $(X, T, \mathcal{B}, \mu)$ be a probability measure preserving dynamical system. If μ is ergodic, and $\psi \in L^p(\mu)$ for some $1 \leq p < \infty$ then there exists $\psi^* \in L^p(\mu)$ with $\psi^* \circ T = \psi^*$ μ-a.e. such that

$$\|\frac{1}{n} S_n \psi - \psi^*\|_p \to 0 \text{ as } n \to \infty.$$

[1] John von Neumann was earlier in proving his L^1-version, but Birkhoff delayed its publication until after the appearance of his own paper.

The L^p Ergodic Theorem

Proof of the L^p Ergodic Theorem.

First assume that ψ is bounded (and hence in $L^p(\mu)$). By Birkhoff's Ergodic Theorem there is ψ^* such that $\frac{1}{n}S_n\psi(x) \to \psi^*(x)$ μ-a.e., and ψ^* is bounded (and hence in $L^p(\mu)$ too). In particular,

$$|\frac{1}{n}S_n\psi(x) - \psi^*(x)|^p \to 0 \quad \mu\text{-a.e.}$$

By the Bounded Convergence Theorem, we can swap the limit and the integral:

$$\begin{aligned}
&\lim_{n\to\infty} \|\frac{1}{n}S_n\psi - \psi^*\|_p \\
&= \lim_{n\to\infty} \left(\int_X |\frac{1}{n}S_n\psi(x) - \psi^*(x)|^p \, d\mu \right)^{1/p} \\
&= \left(\int_X \lim_{n\to\infty} |\frac{1}{n}S_n\psi(x) - \psi^*(x)|^p \, d\mu \right)^{1/p} = 0.
\end{aligned}$$

The L^p Ergodic Theorem

Convergent sequences are Cauchy, so for every $\varepsilon > 0$ there is $N = N(\varepsilon, \psi)$ such that

$$\|\frac{1}{m}S_m\psi - \frac{1}{n}S_n\psi\|_p < \frac{\varepsilon}{2} \qquad (*)$$

for all $m, n \geq N$.

Now if $\phi \in L^p(\mu)$ is unbounded, we want to show that $\frac{1}{n}S_n\phi$ is a Cauchy sequence in $\| \ \|_p$. Let $\varepsilon > 0$ be arbitrary, and take ψ bounded such that $\|\phi - \psi\|_p < \varepsilon/4$. By T-invariance,

$$\|\frac{1}{n}S_n\phi - \frac{1}{n}S_n\psi\|_p \leq \|\phi - \psi\|_p \quad \text{for all } n \geq 1. \qquad (**)$$

The L^p Ergodic Theorem

Repeat (*) and (**) from the previous slide:

$$\|\frac{1}{m}S_m\psi - \frac{1}{n}S_n\psi\|_p < \frac{\varepsilon}{2} \quad \text{for all } m, n \geq N(\varepsilon, \psi) \qquad (*)$$

and

$$\|\frac{1}{n}S_n\phi - \frac{1}{n}S_n\psi\|_p \leq \|\phi - \psi\|_p < \frac{\varepsilon}{4} \quad \text{for all } n \geq 1. \qquad (**)$$

By the triangle inequality,

$$\begin{aligned}
\|\frac{1}{m}S_m\phi - \frac{1}{n}S_n\phi\|_p &\leq \|\frac{1}{m}S_m\phi - \frac{1}{m}S_m\psi\|_p \\
&\quad + \|\frac{1}{m}S_m\psi - \frac{1}{n}S_n\psi\|_p \\
&\quad + \|\frac{1}{n}S_n\phi - \frac{1}{n}S_n\psi\|_p \\
&< \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon
\end{aligned}$$

for all $m, n \geq N(\varepsilon, \psi)$.

The L^p Ergodic Theorem

Hence $\frac{1}{n}S_n\phi$ is Cauchy in $\| \, \|_p$ and thus converges to some $\phi^* \in L^p(\mu)$. We have

$$\left|\frac{n+1}{n}\frac{1}{n+1}S_{n+1}\phi(x) - \frac{1}{n}S_n\phi \circ T(x)\right| = |\frac{1}{n}\phi(x)|$$

for all x. Taking the limit $n \to \infty$ gives $\phi^* = \phi^* \circ T$ μ-a.e.

This concludes the proof of the L^p Ergodic Theorem.