ERGODIC DECOMPOSITIONS

LIAM HODGKINSON

To begin, recall the ergodic theorem of Hurewicz proved last week:

Theorem (Hurewicz's Ergodic Theorem). *Suppose that T is a conservative, measure-preserving transformation of the σ-finite measure space $(X, \mathcal{X}, \mu)$. Then for almost every $x \in X$ and every $f, g \in L^1(\mu)$ with $g > 0$,*

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \hat{T}^k f(x)}{\sum_{k=1}^{n} \hat{T}^k g(x)} = \mathbb{E}_{\mu_g}\left[\frac{f}{g}\middle|\mathcal{T}\right](x),$$

where $\hat{T}$ is the transfer operator for T, $\mu_g(E) = \int_E g d\mu$ and $\mathcal{T}$ is the σ-algebra of T-invariant sets in $\mathcal{X}$.

Remark. If T is invertible (and as it turns out, even when it is not invertible), $\hat{T}^k f(x)$ can be replaced with $f(T^k x)$ and likewise for g. As a corollary of this, when $g \equiv 1$, Birkhoff's famous ergodic theorem is obtained:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k x) = \mathbb{E}_\mu[f|\mathcal{T}](x).$$

This result essentially states that the long-running averages are projections onto the space of T-invariant sets. As nice as this result is mathematically, it is not particularly practical at this point. What we would like is for the expectation to be a constant (not dependent on x).

Definition. A sequence $\{x_n\}_{n=1}^{\infty} \subset X$ is *uniform* with respect to μ if for every $f \in C(X)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(x_k) = \int_X f d\mu.$$

The easiest way to generate a uniform sequence in practice would be by applying an operation to an initial point repeatedly. What we want to look at are operators T such that the forward orbit $\{T^n x\}_{n=0}^{\infty}$ is uniform with respect to μ.

Proposition. *A measure-preserving system $(X, \mathcal{X}, \mu, T)$ is ergodic if and only if for almost every $x \in X$, the forward orbit $\{T^n x\}_{n=0}^{\infty}$ is uniform with respect to m.*

Proof. $\mathbb{E}_\mu[\mathbb{1}_A|\mathcal{T}] = \mu(A)$ μ-almost everywhere if and only if $\mathcal{T} = \{\emptyset, X\} \bmod \mu$, that is, $\mathcal{T}$ is ergodic on $(X, \mathcal{X}, \mu)$. □

Of course, not all measure-preserving transformations are ergodic, but it turns out that ergodic transformations act as their building blocks, and this notion follows from the ergodic theorems. Functions in L^2 can be approximated arbitrarily close by linear combinations of indicator functions. We can do the same thing for measure-preserving transformations via the following theorem, which allows for proofs of general theorems on measure-preserving transformations by proofs on the subset of ergodic transformations.

Theorem (Ergodic Decomposition)**.** *Suppose that T is a conservative, invertible, measure-preserving transformation of a standard σ-finite measure space $(X, \mathcal{X}, \mu)$, then there is a probability space $(Y, \mathcal{Y}, \nu)$ and a unique family of measures $\{\mu_y\}_{y\in Y}$ on X such that*
1. for every $y \in Y$, T is an invertible, conservative ergodic measure-preserving transformation of $(X, \mathcal{X}, \mu_y)$,
2. for every $E \in \mathcal{X}$, $\mu(E) = \int_Y \mu_y(E)\, d\nu(y)$.

Remark. For any $f \in L^1(\mu)$, there is

$$\int_X f d\mu = \int_Y \left(\int_X f(x)\, d\mu_y(x)\right) d\nu(y).$$

To prove this result, we need a powerful theorem from probability theory. This theorem can be thought of as a way of conditioning on events with zero measure.

Theorem (Disintegration Theorem)**.** *Let $(X, \mathcal{X}, \mu)$ be a probability space and $(Y, \mathcal{Y})$ a measurable space. Suppose $\pi : X \to Y$ is measurable and the pushforward measure $\nu = \mu \circ \pi^{-1} \in \mathcal{P}(Y, \mathcal{Y})$.*
Then there is a $Y_0 \in \mathcal{Y}$ such that $\nu(Y_0) = 1$ and there is a unique family of measures $\{\mu_y\}_{y\in Y_0}$ on X such that $\mu_y(\pi^{-1}\{y\}) = 1$ for every $y \in Y_0$ and for every $A \in \mathcal{X}$, $B \in \mathcal{Y}$,

$$\mu\left(A \cap \pi^{-1}B\right) = \int_B \mu_y(A)\, d\nu(y).$$

For each $y \in Y_0$, μ_y is called the fibre measure over $\pi^{-1}\{y\}$.

Remark. Observe that since $\pi^{-1}(Y_0)$ has full measure on μ, $\mu(A) = \int_{Y_0} \mu_y(A)\, d\nu(y)$. More generally, there is

$$\int_X f(g \circ \pi)\, d\mu = \int_Y \left(\int_X f d\mu_y\right) g(y)\, d\nu(y).$$

Proof Outline.

- For each set $A \in \mathcal{X}$, define measures $\nu_A(B) := \mu\left(A \cap \pi^{-1}B\right)$.
- If $\nu(B) = 0$, then $\nu_A(B) \le \mu \circ \pi^{-1}(B) = 0$ and so we can define u_y as the mapping $A \mapsto \frac{\partial \nu_A}{\partial \nu}(y)$, which satisfies
$$\mu\left(A \cap \pi^{-1}B\right) = \int_B u_y(A)\, \nu(y).$$
- It only remains to make u_y into measures. It is easy to check that it is countably additive on disjoint sets.
- Performing Caratheodory construction on the outer measures generated by u_y, we obtain the desired μ_y.

□

From this, we can prove our main theorem.

Proof of Main Theorem.

Reduction to the Probability Space

- Let P be a probability measure on $(X, \mathcal{X})$ with the same null sets as μ and let $\mathcal{T}$ be the σ-algebra of invariant sets of T.
- T is a conservative, invertible, non-singular transformation of the probability space $(X, \mathcal{X}, P)$.
- By the factor proposition, there is a probability space $(Y, \mathcal{Y}, \nu)$ and a factor map $\pi : X \to Y$ such that $\pi^{-1}\mathcal{Y} = \mathcal{T}$ and $\nu = \mu \circ \pi^{-1}$.
- Apply the disintegration theorem to get $Y_0 \in \mathcal{Y}$ with full ν-measure and a family of measures $\{P_y\}_{y \in Y_0}$ on X.

Constructing the Family of Measures (Property 2)

- Define the family of measures $\{\mu_y\}_{y \in Y}$ by

$$\mu_y (E) = \int_E \frac{\partial \mu}{\partial P} dP_y.$$

- Observe that there is

$$\int_Y \mu_y (E) \, d\nu (y) = \int_Y \left(\int_X \mathbb{1}_E (x) \frac{\partial \mu}{\partial P} dP_y (x) \right) d\nu (y) = \int_E \frac{\partial \mu}{\partial P} dP = \mu (E).$$

T is Invertible and Conservative on μ_y

- T is still invertible on μ_y.
- Using standard results, it is straightforward to show T is conservative on $(X, \mathcal{X}, P_y)$ for every $y \in Y$.
- Also implies that T is conservative on μ_y.

T is Measure-Preserving on μ_y

- Let f be the Radon-Nikodym derivative of μ with respect to P. Then using the chain rule of Radon-Nikodym derivatives repeatedly, there is

$$\frac{\partial (P \circ T)}{\partial P} = \frac{f}{f \circ T}.$$

- Since T is non-singular, $P_y \circ T \sim P_y$. Indeed, for almost every $y \in Y_0$, by the uniqueness of the disintegration,

$$\frac{\partial (P_y \circ T)}{\partial P_y} = \frac{\partial (P \circ T)}{\partial P} = \frac{f}{f \circ T} \ P_y\text{-almost everywhere.}$$

because

$$(P \circ T)(B) = \int_{Y_0} (P_y \circ T)(B)\, d\nu(y) = \int_{Y_0} \int_B \frac{\partial (P_y \circ T)}{\partial P_y} dP_y d\nu(y).$$

- From this, we can deduce that

$$\begin{aligned}\frac{\partial(\mu_y \circ T)}{\partial \mu_y} &= \frac{\partial(\mu_y \circ T)}{\partial(P_y \circ T)} \cdot \frac{\partial(P_y \circ T)}{\partial P_y} \cdot \frac{\partial P_y}{\partial \mu_y} \\ &= \frac{f \circ T}{f} \cdot \frac{f}{f \circ T} = 1.\end{aligned}$$

- Thus T is measure-preserving on μ_y.

T is Ergodic on μ_y

- By the Birkhoff ergodic theorem for T on $(X, \mathcal{X}, P)$, for almost every $x \in X$ and $B \in \mathcal{X}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_B\left(T^k x\right) = \mathbb{E}_P\left[\mathbb{1}_B | \mathcal{T}\right](x) = P_{\pi x}(B).$$

- Since $P_y\left(\pi^{-1}\{y\}\right) = 1$, for P_y-almost every $x \in X$, $\pi x = y$. Thus, for almost every $y \in Y$, $P_{\pi x}(B) = P_y(B)$ P_y-almost everywhere.
- By the Birkhoff ergodic theorem for T on $(X, \mathcal{X}, P_y)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_B\left(T^k x\right) = \mathbb{E}_{P_y}\left[\mathbb{1}_B | \mathcal{T}\right](x).$$

- Equating the two gives for every $B \in \mathcal{X}$ and almost every $y \in Y$, and P_y-almost every $x \in X$,

$$\mathbb{E}_{P_y}\left[\mathbb{1}_B | \mathcal{T}\right](x) = P_y(B).$$

- Thus, T is ergodic on P_y, and T is still ergodic on μ_y.

□