Lecture 14 : Ergodic Theorem

MATH275B - Winter 2012 *Lecturer: Sebastien Roch*

References: [Var01, Chapter 6], [Dur10, Section 6.2], [SS05, Section 6.5].

Previous class

In view of the canonical example in the previous lecture, we assume that we have $(\Omega, \mathcal{F}, \mathbb{P})$, $f \in \mathcal{F}$, T a measure-preserving transformation, and we let $X_n(\omega) = f(T^n\omega)$ for all $n \geq 0$.

We are interested in the convergence of empirical averages

$$n^{-1}S_n(\omega) = n^{-1}\sum_{m=0}^{n-1} X_m(\omega) = n^{-1}\sum_{m=0}^{n-1} f(T^m\omega).$$

1 Invariant sets

EX 14.1 *Let $\Omega = \{a, b, c, d, e\}$ and $\mathcal{F} = 2^\Omega$. Take $f = \mathbb{1}_A$ for some set $A \in \mathcal{F}$.*

1. *Suppose $T = (a, b, c, d, e)$. For T to be measure-preserving we require $\mathbb{P}[a] = \mathbb{P}[b] = \cdots$ so that $\mathbb{P}[a] = 1/5$ is the only possibility. (It is easy to see that T is indeed measure-preserving because the number of elements of Ω is invariant under T.) In that case, it is immediate that*

$$n^{-1}S_n \to \mathbb{P}[A] = \mathbb{E}[f].$$

2. *Suppose $T = (a, b, c)(d, e)$. Let $\Omega_1 = \{a, b, c\}$, $\mathcal{F}_1 = 2^{\Omega_1}$, $\Omega_2 = \{d, e\}$ and $\mathcal{F}_2 = 2^{\Omega_2}$. For T to be measure-preserving we need $\mathbb{P}[a] = \mathbb{P}[b] = \mathbb{P}[c] = \alpha/3$ and $\mathbb{P}[d] = \mathbb{P}[e] = \beta/2$. (Any choice of α, β with $\alpha + \beta = 1$ works because the number of elements of Ω_1 and Ω_2 is invariant under T.) Take $A = \{a, d\}$. Then $n^{-1}S_n \to 1/3$ with probability α (i.e. if $\omega \in \Omega_1$) and $n^{-1}S_n \to 1/2$ with probability β. Denoting $\hat{f}$ this limit, we note*

$$\mathbb{E}[\hat{f}] = \mathbb{P}[A] = \mathbb{E}[f],$$

but $\hat{f}$ is not constant if $\alpha, \beta \neq 0$. However, it is completely determined by whether $\omega \in \Omega_1$ or $\omega \in \Omega_2$.

DEF 14.2 *A set $A \in \mathcal{F}$ is* invariant *if*

$$(\{\omega : T\omega \in A\} =)T^{-1}A = A,$$

up to a null set. It is nontrivial *if $0 < \mathbb{P}[A] < 1$. The set of all invariant sets forms a σ-field $\mathcal{I}$. The transformation T is said* ergodic *if $\mathcal{I}$ is trivial, that is, all invariant sets are trivial. A function g is* invariant *if $g(T\omega) = g(\omega)$ a.s. Note that g is invariant iff $g \in \mathcal{I}$. (Exercise 6.1.1 in [Dur10].)*

2 Ergodic Theorem

It will be convenient to think of T as an operator of functions

$$Uf(\omega) = f(T\omega),$$

in which case $U^m f(\omega) = f(T^m\omega)$ and we define

$$A_n f = n^{-1}(I + \cdots + U^{n-1})f.$$

LEM 14.3 *If $g \in L^1$ then*

$$\mathbb{E}[Ug] = \mathbb{E}[g].$$

Moreover if $g, g' \in L^2$ then

$$\|Ug\| = \|g\|,$$

and

$$\langle Ug', Ug\rangle = \langle g', g\rangle.$$

Proof: Start from indicators. ■

THM 14.4 *Let $f \in L^1$. Then there is $\hat{f} \in \mathcal{I}$ s.t.*

$$A_n f \to \hat{f} \equiv \mathbb{E}[f \,|\, \mathcal{I}], \text{ a.s and in } L^1.$$

EX 14.5 (IID RVs) *Let $X_n(\omega) = \omega_n$ are iid rvs. If A is invariant then $\{\omega : \omega \in A\} = \{\omega : T\omega \in A\} \in \sigma(X_1, \ldots)$ and by induction*

$$A \in \cap_{n\geq 0}\sigma(X_n, \ldots) = \mathcal{T},$$

where $\mathcal{T}$ is the tail σ-field. Thus $\mathcal{I} \subseteq \mathcal{T}$. Since $\mathcal{T}$ is trivial by Kolmogorov's $0-1$ law, so is $\mathcal{I}$. Therefore T is ergodic and $\mathbb{E}[f \,|\, \mathcal{I}] = \mathbb{E}[f]$. Applying the ergodic thm to $f = X_0 \in L^1$ we get

$$n^{-1}\sum_{m=0}^{n-1} X_m(\omega) \to \mathbb{E}[X_0],$$

that is, we recover the SLLN.

3 L^2 Ergodic Theorem

THM 14.6 *Let $f \in L^2$. Then there is $\hat{f} \in \mathcal{I}$ s.t.*

$$A_n f \to \hat{f} \equiv \mathbb{E}[f \,|\, \mathcal{I}], \text{ in } L^2.$$

Proof: Let

$$H_0 = \{f \in L^2 \, : \, Uf = f \text{ a.s.}\},$$

and note that $A_n f = f$ for $f \in H_0$. We need the following lemma from basic Hilbert space theory (see [SS05, Lemma 6.5.2]).

LEM 14.7 *The following hold:*

1. $H_0 = \{f \in L^2 \, : \, U^* f = f \text{ a.s.}\}$.
2. $H_0^\perp = \overline{Range(I - U)}$.

Proof: See e.g. [SS05]. ■

For $\varepsilon > 0$, write $f = f_0 + f_1$ where $f_0 \in H_0$ and $\|f_1 - f_1'\|_2 < \varepsilon$ s.t. $f_1' = (I - U)g_1'$. Then

$$A_n f_0 = f_0, \text{ and } A_n f_1' = \frac{1}{n}(I - U^n)g_1',$$

so that

$$\begin{aligned}
\|A_n f - f_0\|_2 &= \|n^{-1}(I - U^n)g_1' + A_n(f_1 - f_1')\|_2 \\
&\leq (\|g_1'\|_2 + \|U^n g_1'\|_2)n^{-1} + n^{-1} \sum_{m=0}^{n-1} \|U^m(f_1 - f_1')\|_2 \\
&= 2\|g_1'\|_2 n^{-1} + n^{-1} \sum_{m=0}^{n-1} \|f_1 - f_1'\|_2 \\
&\to \varepsilon.
\end{aligned}$$

■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[SS05] Elias M. Stein and Rami Shakarchi. *Real analysis*. Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005. Measure theory, integration, and Hilbert spaces.

[Var01] S. R. S. Varadhan. *Probability theory*, volume 7 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2001.