Birkhoff Ergodic Theorem

Jonathan L.F. King
University of Florida, Gainesville FL 32611-2082, USA
squash@ufl.edu
Webpage http://www.math.ufl.edu/~squash/
8 March, 2013 (at *20:22*)

Pointwise convergence

Fix an mpt $(T\colon X, \mathscr{X}, \mu)$ on a prob.space. Let $\mathbb{A}_N(f)$ denote the fnc whose value at x is $\mathbb{A}_N^k\big(f(T^k x)\big)$. Use $\mathsf{E}(\cdot \mid \cdot)$for the conditional expectation operator.

The proof below is a symmetrized version of the Katznelson-Weiss proof.

1: Birkhoff Ergodic Theorem. *Fix $f \in \mathbb{L}^1(\mu)$. Then this almost-everywhere limit exists,*

$$\operatorname*{a.e-lim}_{N\to\infty} \mathbb{A}_N(f) \quad = \quad \mathsf{E}(f \mid \mathscr{I}),$$

where $\mathscr{I}$ denotes the field of T-invariant sets. ◊

Reduction. (The crux is proving a.e-convergence. For then, identifying the limit function as $\mathsf{E}(f \mid \mathscr{I})$ is not difficult.) Let $\overline{f}$ and $\underline{f}$ denote the pointwise $\lim[\sup, \inf]_{N\to\infty} \mathbb{A}_N(f)$. It suffices to show that

$$\text{2:} \qquad \int \overline{f} \quad \leqslant \quad \int f.$$

For applying this to $-f$ yields $\int f \leqslant \int \underline{f}$. Hence $\int \overline{f} \leqslant \int f \leqslant \int \underline{f}$. Since

$$\textstyle\int \overline{f} \leqslant \int \underline{f} \qquad \text{yet} \qquad \overline{f}() \geqslant \underline{f}(),$$

it follows that $\overline{f} \overset{\text{a.e}}{=} \underline{f}$.

Let f_M denote $\mathrm{Max}(f, M)$, for $M \in \mathbb{Z}_-$. So

$$f^+ \;\geqslant\; f_{-1} \;\geqslant\; f_{-2} \;\geqslant\; f_{-3} \;\geqslant\; \dots,$$

and the Monotone Convergence Thm, implies $\int f_M \searrow \int f$ as $M \searrow -\infty$. Now suppose we had (2) for bounded-below integrable fncs. Applying (2) to an f_M gives $\int f_M \geqslant \int \overline{f_M}$. Each $f_M \geqslant f$, so $\overline{f_M} \geqslant \overline{f}$; hence $\boxed{\int f_M \geqslant \int \overline{f}.}$ So the foregoing MonoCT yields that $\int f \geqslant \int \overline{f}$. And this is (2) for f.

The upshot is that we we may assume that f is bounded below. And, since $\mu(X) < \infty$ we may add a constant to assume that $\boxed{f() \geqslant 0}$.

Bounding above. We now do another reduction, this time sending $M \nearrow \infty$. Let

$$h_M() \quad := \quad \mathrm{Min}\big(\overline{f}(), M\big) - \tfrac{1}{M}.$$

Note that h_M is T-invariant. And since

$$0 \leqslant h_1 \leqslant h_2 \leqslant \dots \quad \text{and} \quad h_M \nearrow \overline{f}$$

pointwise, the MonoCT forces $\int h_M \nearrow \int \overline{f}$.

Thus it suffices to establish $\int f \geqslant \int h$ for each bounded, invariant function h which is exceeded (pointwise) by $\overline{f}$. By scaling both f and h we may assume the bound is 1. Our goal, having fixed a positive ε, becomes

2′: *If a T-invariant function $h() \leqslant 1$ satisfies $h < \limsup_{n\to\infty} \mathbb{A}_N(f)$ pointwise, then*

$$\int h \;\leqslant\; 2\varepsilon + \int f.$$

□

Pf of (1) via (??′). Let $W(x)$ be the smallest posint N such that

$$*\text{:} \qquad \sum_{i=0}^{N-1} f(T^i x) \;\geqslant\; \sum_{i=0}^{N-1} h(T^i x).$$

The strict inequality in (2′) implies that this N exists since (by T-invariance of h) $\mathrm{RhS}(*) = h(x)$. From this **stopping time** we will obtain a sequence of stopping times.

Fix a sufficiently large constant W_{Max} so that $\mu(G^c) \leqslant \varepsilon$, where

$$G \;:=\; \big\{x \;\big|\; W(x) < W_{\mathrm{Max}}\big\}.$$

Filename: `Classwork/Analysis/Erg2000t/macs_ergbook.tex`
As of: *Monday 27Nov2000.* Typeset: *8Mar2013 at 20:22.*

For an $x \in X$, let $s_1 = s_1(x)$ be the smallest $m \geqslant 0$ so that $T^m(x) \in G$. For $k = 1, 2, \ldots$ inductively define stopping times s_k and ℓ_k and intervals J_k as follows:

$$\ell_k := W\Big(T^{s_k}(x)\Big) \overset{\text{note}}{\leqslant} W_{\text{Max}} \qquad \text{and}$$
$$J_k := [s_k \,..\, s_k + \ell_k) \,.$$

Finally, let s_{k+1} is the smallest integer $m \geqslant s_k + \ell_k$ such that $T^m(x) \in G$.

Summing along a name. Fix an $N \gg W_{\text{Max}}$ and let $K = K(x)$ be the largest index k with $J_k \subset [0 \,..\, N)$. Let

$$\mathbb{J} := \bigsqcup_{k=1}^{K} J_k \qquad \text{and} \qquad \mathbb{J}^c := [0 \,..\, N) \smallsetminus \mathbb{J} \,.$$

By definition of $W()$ we have the inequality $\sum_{i \in \mathbb{J}} h(T^i x) \leqslant \sum_{i \in \mathbb{J}} f(T^i x)$. Since $f() \geqslant 0$, then,

$$\sum_{i \in \mathbb{J}} h(T^i x) \leqslant \sum_{i \in [0 \,..\, N)} f(T^i x) \,.$$

Note also that an i can be in $\mathbb{J}^c$ only if either $i \in J_{K+1}$ or if $T^i(x) \in G^c$. Since $h() \leqslant 1$ this gives

$$\sum_{i \in \mathbb{J}^c} h(T^i x) \leqslant W_{\text{Max}} + \sum_{i \in [0 \,..\, N)} \mathbf{1}_{G^c}(T^i x) \,.$$

Adding these two inequalities then dividing by an $N > \frac{1}{\varepsilon} \cdot W_{\text{Max}}$, yields

$$\frac{1}{N} \sum_{i \in [0 \,..\, N)} h(T^i x) \leqslant \varepsilon + \frac{1}{N} \sum_{i \in [0 \,..\, N)} \mathbf{1}_{G^c}(T^i x) + \frac{1}{N} \sum_{i \in [0 \,..\, N)} f(T^i x) \,.$$

Observe that *this inequality has no mention of a stopping-time.* Thus, there is no issue as to what it means to integrate it w.r.t x. And indeed,

2″: $$\int h \leqslant \varepsilon + \mu(G^c) + \int f \leqslant \varepsilon + \varepsilon + \int f$$

is the result, yielding (**??**′) as desired. ♦