Oseledets Theorem

Seminar in Homogenous Dynamics

February 2024

Definition: Let $(X, \mathcal{B}, \mu, T)$ be a probability preserving system $(T_* \mu = \mu)$. A sequence of functions $g_n: X \to \mathbb{R}$ of measurable functions is called subadditive, relative to f if $g_{m+n} \leq g_m + g_n \circ T^m$ for all $m, n \geq 1$

Examples: 1) If $f: X \to \mathbb{R}$ is measurable, then the ergodic sum $g_N = \sum_{n=0}^{N-1} f \circ T^n$ is an additive sequence, and particularly subadditive.

2) Given a measurable $A: X \to GL_d\mathbb{R}$, we define $A^n = (A \circ f^{n-1}) \cdot (A \circ f^{n-2}) \cdot \ldots \cdot A$ and $\varphi_n(x) = \log ||A^n(x)||$ with the norm being the operator norm. This is subadditive since the operator norm satisfies $\left|\left|AB\right|\right| < \left|\left|A\right|\right| \cdot \left|\left|B\right|\right|$ for any two $A, B \in M_d(\mathbb{R})$.

3) In the previous example if X is a smooth manifold, then the differential map of f, $D_x(f): T_x X \to T_{f(x)} X$ between tangent spaces satisfies

$$D_{f^n(x)}(f) = D_{f^{n-1}(x)}(f) \cdot \ldots \cdot D_x(f)$$

Any f with differential of full rank gives us a map $D(f): X \to GL_d(\mathbb{R})$, where $d = \dim X$.

Remark: A map $F: X \times \mathbb{R}^d \to X \times \mathbb{R}^d$ with $F(x, v) = (f(x), A(x)v)$ for A as in example 2 is called a linear cocycle and $F^n(x, v) = (f^n(x), A^n(x)v)$

Theorem 1 (Kingman's subadditive ergodic theorem): Let $(X, \mathcal{B}, \mu, T)$ be a p.p.s. and $g_n: X \to \mathbb{R}$ a subadditive sequence with $g_1^+ \in L^1(X, \mathcal{B}, \mu)$. Then the limit $g := \lim_{n\to\infty} \frac{1}{n} g_n$ exists μ –a.e. and it is an invariant function. Moreover, we have that $\int g d\mu = \lim_{n\to\infty} \frac{1}{n} \int g_n d\mu = \inf_{n\to\infty} \frac{1}{n} \int g_n d\mu$.

The following theorem is a direct consequence of the previous one (but was originally proved without using the subadditive ergodic theorem):

Theorem 2 (Furstenberg, Kesten): Let A be as above and assume that $\log^+(A^{\pm 1}) \in L^1$. Then the following limits exist:

$$\lambda_+(x) = \lim_{n\to\infty} \frac{1}{n} \log\|A^n(x)\|, \quad \lambda_-(x) = \lim_{n\to\infty} \frac{1}{n} \log\|A^n(x)^{-1}\|^{-1}$$

and they are invariant and integrable with

$$\int \lambda_+ \, d\mu = \lim_{n\to\infty} \frac{1}{n} \int \log\|A^n(x)\| \, d\mu$$

$$\int \lambda_- \, d\mu = \lim_{n\to\infty} \frac{1}{n} \int \log\|A^n(x)^{-1}\|^{-1} \, d\mu$$

$\lambda_\pm(x)$ are called the extremal Lyapunov exponents

The theorem of Furstenberg and Kesten provides us with growth rate for $||A^n(x)||$. Oseledets theorem, also known as the multiplicative ergodic theorem, provides growth rates for $||A^n(x)v||$ for all $v \in \mathbb{R}^d$. Here we only give the statement of the general case and provide a proof of the 2-dimensional case:

Theorem (Oseledets): Let $(X, \mathcal{B}, \mu, T)$ be a probability preserving system. And let $A: X \to GL_2$ be a measurable function with $\log^+ ||A^{\pm 1}|| \in L^1$. For almost every $x \in X$ there exist $k(x) \in \mathbb{N}$, numbers $\lambda_1(x) > \cdots > \lambda_{k(x)}$, and a sequence of subspaces $\mathbb{R}^d = V_x^1 \supset \cdots \supset V_x^{k(x)} \supset \{0\}$ such that for all $1 \le i \le k$:

1) $k(T(x)) = k(x), \lambda_i(T(x)) = \lambda_i(x), A(x)V_x^i = V_{T(x)}^i$
2) The maps k, λ_i, V_x^i are measurable.
3) For all $v \in V_x^i \setminus V_x^{i+1}$, $\lambda_i(x) = \lim_{n\to\infty} \frac{1}{n} \log\|A^n(x)v\|$

Remarks: 1) When μ is ergodic, $k(x), \lambda_i(x), \dim V_x^i$ are constant a.e.

2) The result does not depend on the choice of norm or basis on each of the subspaces because all norms on the Euclidean space are equivalent and constants disappear. Change of basis is equivalent to taking another norm.

Examples: 1) Take a matrix $A \in SL_d\mathbb{R}$ and take the cocycle $F: X \times \mathbb{R}^d \to X \times \mathbb{R}^d$ with $F(x, v) = (f(x), Av)$ and $F^n(x, v) = (f^n(x), A^n v)$. If A is diagonalizable with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ and eigenspaces $\mathbb{R}^d = V^{\lambda_1} \oplus \dots \oplus V^{\lambda_d}$, for every vector v with a component in V^{λ_1} the norm $\|A^n v\|$ grows like $\lambda^n \|v\|$. On the other hand, if $v \in V^{\lambda_2} \oplus \dots \oplus V^{\lambda_d}$ then the norm has a different rate of growth. We can iteratively use the largest eigenvalue to determine the rate of growth outside the sum of the eigenspaces with smaller eigenvalues. The theorem approves this intuitive argument in a much general setting.

2) Consider products of random matrices. Namely, take $A_0, A_1 \in GL_d\mathbb{R}$, $X = \{A_0, A_1\}^{\mathbb{Z}}$ and $\mu = \mu_p^{\mathbb{Z}}$ for some $p \in (0,1)$. As usual, we take $T: X \to X$ to be the shift map. Also define $A: X \to GL_d\mathbb{R}$ by $(a_n) \mapsto a_0$. Then the cocycle defined by A, T is $F: X \times \mathbb{R}^d \to X \times \mathbb{R}^d$ so that $F^m\big((a_n), v\big) = \big((a_{n+m}), a_{m-1} \cdot \dots \cdot a_0 v\big)$. The theorem provides us with the "growth rate" of this random walk.

3) Applying the theorem in the context given in example 3 from the previous discussion gives us a direction on the tangent space for which the cocycle has exponential growth rate of λ, and determines the rate for all the other directions (This is my intuition for the 2-dimensional case).

Before we get to the proof of the 2-dimensional case, we prove a lemma:

Lemma 1: Let $(X, \mathcal{B}, \mu, T)$ be a p.p.s. and $f \in L^1$ then for μ-a.e. $x \in X$, $\lim_{n\to\infty} \frac{1}{n} f\big(T^n(x)\big) = 0$

Proof: Let $\varepsilon > 0$ and denote $A_n = \{x \in X: |f(T^n(x))| \geq n\varepsilon\}$. Therefore,

$$\mu\big(A_n(\varepsilon)\big) = \mu(\{x \in X: |f(x)| > n\varepsilon\}) = \sum_{k=n}^{\infty} \mu\left(\left\{x \in X: k \leq \frac{|f(x)|}{\varepsilon} < k+1\right\}\right)$$

$$\sum_{n=1}^{\infty} \mu\big(A_n(\varepsilon)\big) = \sum_{k=1}^{\infty} k\, \mu\left(\left\{x \in X: k \leq \frac{|f(x)|}{\varepsilon} < k+1\right\}\right) \leq \int_X \frac{|f|}{\varepsilon} d\mu < \infty$$

By Borel-Cantelli $\mu\big(\text{limsup} A_n(\varepsilon)\big) = 0$. For every $x \notin \text{limsup} A(\varepsilon)$ there is some $p > 1$ such that for every $n \geq p$ such that $|f(T^n(x))| < n\varepsilon$. To conclude the proof, notice that $A = \bigcup_{i\geq 1} \text{limsup} A\left(\frac{1}{i}\right)$ has measure zero and $\lim_{n\to\infty} \frac{1}{n} f\big(T^n(x)\big) = 0$ on its complement.

Theorem: Let $(X, \mathcal{B}, \mu, T)$ be a probability preserving system. And let $A: X \to GL_2$ be a measurable function with $\log^+ ||A^{\pm 1}|| \in L^1$. Then for μ-a.e. $x \in X$ one of the following holds:

1) $\lambda_-(x) = \lambda_+(x)$ and for all $v \in \mathbb{R}^2 \setminus \{0\}$, $\lambda_\pm(x) = \lim_{n\to\infty} \frac{1}{n} \log\left||A_n(x)v|\right|$
2) $\lambda_+(x) > \lambda_-(x)$ and there exists a 1-dimensional subspace $E_x \subset \mathbb{R}^2$ such that:

$$\lim_{n\to\infty} \frac{1}{n} \log\left||A_n(x)v|\right| = \begin{cases} \lambda_-(x), & v \in E_x \setminus \{0\} \\ \lambda_+(x), & v \in \mathbb{R}^2 \setminus E_x \end{cases}$$

And $A(x)E_x = E_{T(x)}$

Note that the existence of $\lambda_\pm(x)$ follows of the theorem of Furstenberg-Kesten.

Proof: We begin with a reduction to the case where A takes values in SL_2. We justify this step by checking that for any $A(x) \in GL_d$ we can normalize $A(x)$ to obtain a matrix $B(x) \in SL_d$. If the integrability assumptions hold for $A(x)$, then they also hold for $B(x)$. Also, the corresponding one-dimensional lines specified by the theorem are the same for both cocycles, and the Lyapunov exponents of the one are just a shift of the exponents of the other by some additive constant dependent on x.

We show an equivalent definition for the operator norm taken with respect to the L^2 norm on $\mathbb{R}^n$

Proposition: $\|A\| = \sqrt{\lambda}$ where λ is the largest eigen value of A^tA

Proof: A^tA is symmetric and hence has an orthonormal basis of eigenvectors $\{e_1, \dots e_n\}$. Let $\lambda_1 \geq \cdots \geq \lambda_n$be the corresponding eigenvalues. Then for any $x \in \mathbb{R}^n \setminus \{0\}$ if $x = \sum_{i=1}^n \alpha_i e_i$ we have $\|A(x)\|^2 = \sum_{i=1}^n \lambda_i a_i^2 \leq \lambda_1 \cdot \|x\|$, so $\|A\| \leq \sqrt{\lambda_1}$. We also have $\|A\|^2 \geq \|Ae_1\|^2 = \lambda_1$ which ends the proof.

We prove two preliminary claims about SL_2:

Claim 1: For any $A \in SL_2$, $\|A\| = ||A^{-1}||$ where the operator norm is taken with respect to Euclidean L^2 norm on $\mathbb{R}^2$.

Proof: The characteristic polynomial of $A^t A$ is $x^2 - tr(A^t A)x + \det A^t A$ and it is easy to see that the same polynomial is obtained for $(A^{-1})^t(A^{-1})$ for $A \in SL_2$. From the previous proposition we the equality of the norms.

Claim 2: Let $A \in SL_2$, then there exist vectors $s, u \in \mathbb{R}^2$ such that $\left||As|\right| = \left||A|\right|^{-1}$ and $\left||Au|\right| = ||A||$. If $\|A\| \neq 1$ then u, v are orthogonal and unique up to sign.

Proof: Existence of both vectors follows from the fact that the continuous function $x \mapsto ||Ax||$ achieves maximal and minimal values on the compact set S^1. Let $s \in \mathbb{R}^2$ be the most contracted vector. Pick some unit vector $u \in \mathbb{R}^2$ orthogonal to s. Then, for every $x \in \mathbb{R}^2$, $x = \alpha_1 s + \alpha_2 u$ with $|\alpha_1|^2 + |\alpha_2|^2 = 1$.

We have $\left||Ax|\right| \leq |\alpha_1|\left||As|\right| + |\alpha_2|\left||Au|\right| \leq (|\alpha_1| + |\alpha_2|) \cdot \left||Au|\right| \leq ||Au||$. Therefore, $\left||A|\right| = ||Au||$ as required.

Now, fix some $x \in X$ for which theorem 2 holds.

Notice that $\lambda_+(x) + \lambda_-(x) = \lim_{n\to\infty} \frac{1}{n} \log(\|A^n(x)\| \cdot \|A^n(x)^{-1}\|^{-1}) = 0$.Therefore, we denote $\lambda(x) := \lambda_+(x) = -\lambda_-(x)$.

We prove that of $\lambda(x) = 0$ then (1) holds and if $\lambda(x) > 0$ then (2) holds.

Assume that $\lambda(x) = 0$, $(\lambda_+(x) = \lambda_-(x))$ then for any $v \in \mathbb{R}^2$,

$$\|A^n(x)\|^{-1}\|v\| = \|A^n(x)^{-1}\|^{-1}\|v\| \leq \|A^n(x)v\| \leq \|A^n(x)\|\|v\|$$

because $\|A^n(x)^{-1}\| = \sup\left\{\frac{\|u\|}{\|A^n(x)u\|} : u \in \mathbb{R}^2 \setminus \{0\}\right\} \geq \frac{\|v\|}{\|A^n(x)\|}$

And hence

$$\frac{1}{n}\log\|A^n(x)\|^{-1}\|v\| \leq \frac{1}{n}\log\|A^n(x)v\| \leq \frac{1}{n}\log\|A^n(x)\|\|v\|$$

Which shows the result for the first case in the theorem.

Now, assume that $\lambda(x) > 0$, then for n large enough $\|A_n(x)\| > 1$. By claim 2 we can choose $u_n(x), s_n(x) \in \mathbb{R}^2$ such that $\|A_n(x)s_n(x)\| = \|A^n(x)\|^{-1}$ and $\|A_n(x)u_n(x)\| = \|A^n(x)\|$. We proceed with the following lemmas:

Lemma 2: $\limsup \frac{1}{n} \log\left|\sin \angle\big(s_n(x), s_{n+1}(x)\big)\right| < -2\lambda(x)$

Proof: We denote $\alpha_n = \angle\big(s_n(x), s_{n+1}(x)\big)$ and decompose

$$s_n(x) = \sin\alpha_n\, u_{n+1}(x) + \cos\alpha_n\, s_{n+1}(x)$$

Then we use the properties of $u_n(x), s_n(x)$ to get:

$$|\sin\alpha_n| \cdot \|A^{n+1}(x)\| \leq \|A^{n+1}(x)s_n(x)\| \leq \big\|A\big(T^n(x)\big)\big\|\|A^n(x)\|^{-1}$$

Which implies

$$\frac{1}{n}\log|\sin(\alpha_n)| \leq \frac{1}{n}\log\big\|A\big(T^n(x)\big)\big\| - \frac{1}{n}\log\|A^n(x)\| - \frac{1}{n}\log\|A^{n+1}(x)\|$$

After taking limsup of both sides, the definition of $\lambda(x)$ and lemma 1 complete the proof.

Lemma 3: The sequence $\big(s_n(x)\big)_n$ is a Cauchy sequence.

Proof: Using the previous lemma, we obtain $\|s_n(x) - s_{n+1}(x)\| \leq 2|\sin\alpha_n| \leq 2e^{n(-2\lambda(x)+\varepsilon)}$ for all $\varepsilon > 0$ such that $-2\lambda(x) + \varepsilon < 0$. Therefore, using the triangle inequality we can bound the differences between elements of the sequence:

$$\|s_n(x) - s_{n+k}(x)\| \leq Ce^{n(-2\lambda(x)+\varepsilon)}$$

For a suitable constant $C > 0$ and for arbitrary large enough $n, k > 0$.

Denote $s(x) = \lim_{n\to\infty} s_n(x)$

Lemma 3: $\lim_{n\to\infty}\frac{1}{n}\log\|A^n(x)s(x)\| = -\lambda(x)$

Proof: Denote $\beta_n = \angle\big(s(x), s_n(x)\big)$. We have $s(x) = \cos\beta_n\, s_n(x) + \sin\beta_n\, u_n(x)$. And hence

$$\begin{aligned}
&\operatorname{limsup}\frac{1}{n}\log\|A^n(x)s(x)\| \\
&\leq \max\left\{\operatorname{limsup}\frac{1}{n}\log(|\cos\beta_n|\|A^n(x)s_n(x)\|), \operatorname{limsup}\frac{1}{n}\log(|\sin\beta_n|\|A^n(x)u_n(x)\|)\right\} \\
&\leq \max\left\{\begin{matrix}\operatorname{limsup}\frac{1}{n}\log(\|A^n(x)\|^{-1}), \\ \operatorname{limsup}\frac{1}{n}\log(|\sin\beta_n|) + \operatorname{limsup}\frac{1}{n}\log(\|A^n(x)u_n(x)\|)\end{matrix}\right\} \\
&\leq \max\{-\lambda(x), -2\lambda(x) + \lambda(x)\} = -\lambda(x)
\end{aligned}$$

as required.

It remains to deal with vectors that are not on the line generated by $s(x)$.

Lemma 4: For any $v \in \mathbb{R}^2 \setminus Span(s(x))$, $\lim_{n\to\infty} \frac{1}{n} \log\|A^n(x)v\| = \lambda(x)$

Proof: $v, s(x)$ are not collinear, therefore, if $\gamma_n = \angle(v, s_n(x))$ then $|\sin\gamma_n| > 0$ for n large enough. Again, we decompose to get $v = \cos\gamma_n\, s_n(x) + \sin\gamma_n\, u_n(x)$. Then, a lower bound is $\|A^n(x)v\| \geq |\sin\gamma_n|\|A^n(x)u_n(x)\| - |\cos\gamma_n|\|A^n s_n(x)\|$. Using the fact that for any two real sequences a_n, b_n,

$$\text{liminf}\frac{1}{n}\log(a_n + b_n) \geq \max\left\{\text{liminf}\frac{1}{n}\log a_n\ ,\text{liminf}\frac{1}{n}\log b_n\right\}$$

we obtain $\text{liminf}\frac{1}{n}\log\|A^n(x)v\| \geq \lambda(x)$. On the other hand, $\text{limsup}\frac{1}{n}\log\|A^n(x)v\| \leq \lambda(x)$, which end the proof.

We now show that the line spanned by $s(x)$ is invariant under the action:

Claim: $A(x)s(x)$ and $s(T(x))$ are collinear.

Proof: By the previous lemma, it suffices to check that

$$\lim_{n\to\infty}\frac{1}{n}\log\left\|A^n(T(x))\cdot A(x)s(x)\right\| \neq \lambda(f(x))$$

And indeed, by lemma 3,

$$\lim_{n\to\infty}\frac{1}{n}\log\left\|A^n(f(x))\cdot A(x)s(x)\right\| = \lim_{n\to\infty}\frac{1}{n+1}\log\|A^{n+1}(x)s(x)\| = -\lambda(x)$$

From the invariance and positiveness of λ we gat that $\lambda(T(x)) = \lambda(x) \neq -\lambda(x)$.

This finishes the proof for the 2-dimensional case.