MULTIPLICATIVE ERGODIC THEOREMS: A MINI-COURSE

ANTHONY QUAS

Abstract. Lecture notes for a mini-course at the University of Houston in May 2016.

1. Motivation

This mini-course is about:

- The sub-additive ergodic theorem;
- Lyapunov exponents;
- Multiplicative ergodic theorems;

This section aims to motivate these theorems.

1.1. **Sub-additive ergodic theorems.** Kingman (in 1968, following earlier work of Hammersley and Welsh) proved the Sub-additive ergodic theorem.

If $\sigma\colon \Sigma \to \Sigma$ is a measure-preserving transformation[1], a sequence of functions $(f_n)_{n\geq 1}$ is *sub-additive* (with respect to σ) if

$$f_{n+m}(\omega) \leq f_n(\omega) + f_m(\sigma^n\omega).$$

Examples of sequences of functions satisfying this condition?

(1) (Fekete's lemma) Let (a_n) be a sequence of real numbers such that $a_{n+m} \leq a_n + a_m$. (Here the functions (f_n) are constant functions(!)). Then $a_{mk+r} \leq ma_k + a_r$, so that $\limsup a_n/n \leq a_k/k$ for each k. Now $\limsup a_n/n \leq \inf_k a_k/k \leq \liminf a_n/n \leq \limsup a_n/n$. Hence a_n/n converges to a value $a = \inf a_k/k \in [-\infty, \infty)$.

(2) (Birkhoff averages) If f is an L^1 function on Ω, then $f_n(\omega) = f(\omega) + \ldots + f(\sigma^{n-1}\omega)$ is an *additive* sequence:

$$f_{n+m}(\omega) = f_n(\omega) + f_m(\sigma^n\omega).$$

(3) (First passage percolation) Let $\mathcal{E}$ denote the collection of edges in the $\mathbb{Z}^2$ lattice. Let ν be a probability measure on $(0, \infty)$ (with $\int x\, d\nu(x) < \infty$).

[1] all invariant measures will be probability measures

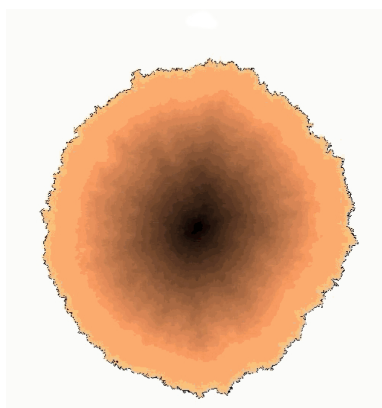
Let $\Omega = (0,\infty)^{\mathcal{E}}$ be the collection of all weightings of $\mathcal{E}$ and equip X with the probability measure $\nu^{\mathcal{E}}$ (so that in a realization $\omega \in \Omega$, each edge is assigned a weight from the distribution ν independently of all other edges). Define a $\mathbb{Z}^2$ action, $\tau_{\mathbf{v}}$ on Ω that translates the pattern of edge weightings through $-\mathbf{v}$.

Now for $\mathbf{v} \in \mathbb{Z}^2$, define $F_{\mathbf{v}}(\omega)$ to be the length of the shortest path from $\mathbf{0}$ to $\mathbf{v}$ (where the length of a path is the sum of the weights of the edges). Then

$$F_{\mathbf{u}+\mathbf{v}}(\omega) \le F_{\mathbf{u}}(\omega) + F_{\mathbf{v}}(\tau_{\mathbf{u}}(\omega)).$$

In particular, if $\mathbf{v}$ is any non-zero integer vector, then defining $\sigma = \tau_{\mathbf{v}}$ and $f_n(\omega) = F_{n\mathbf{v}}(\omega)$, (f_n) is a sub-additive sequence for the ergodic dynamical system $\sigma\colon (0,\infty)^{\mathcal{E}} \to (0,\infty)^{\mathcal{E}}$.

Hammersley and Welsh interpreted this as a 'wetting time': a rock is modelled by $\mathbb{Z}^2$. 'Water' is in contact with the rock at $\mathbf{0}$. The edge label determines the time it takes water to pass from one vertex to its neighbour. They were interested in the geometry of $\{\mathbf{v}\colon F_{\mathbf{v}}(\omega) < T\}$.



(image due to Jéremie Bettinelli)

(4) (Matrix products) If $\sigma\colon (\Omega,\mathbb{P}) \to (\Omega,\mathbb{P})$ is a measure-preserving transformation, and $A\colon (\Omega,\mathbb{P}) \to M_{d\times d}(\mathbb{R})$ is measurable, then one can form the *matrix cocycle*:

$$A^{(n)}(\omega) = A(\sigma^{n-1}\omega)\cdots A(\omega) \text{ for } n \in \mathbb{N} \text{ and } \omega \in \Omega.$$

Notice that

$$A^{(n+m)}(\omega) = A^{(m)}(\sigma^n\omega)A^{(n)}(\omega) \text{ (the cocycle relation).}$$

defining $f_n(\omega) = \log \|A^{(n)}(\omega)\|$, you obtain

$$f_{n+m}(\omega) \le f_n(\omega) + f_m(\sigma^n\omega)$$

(providing $\|\cdot\|$ satisfies $\|AB\| \le \|A\|\|B\|$ (e.g. operator norm))

Theorem 1 (Kingman Sub-additive ergodic theorem, 1968)**.** *Let* $\sigma\colon (\Omega,\mathbb{P}) \to (\Omega,\mathbb{P})$ *be an ergodic measure-preserving transformation. Let* $(f_n)_{n\geq 1}$ *be a sub-additive sequence of integrable functions. Then*

(1) $\lim_{n\to\infty} \frac{1}{n}\int f_n(\omega)\,\mathbb{P}(\omega)$ *converges to a constant* $c \in [-\infty,\infty)$*.*
(2) *For* $\mathbb{P}$*-a.e.* $\omega\in\Omega$, $\frac{1}{n}f_n(\omega) \to c$*.*

Theorem 2 (Birkhoff ergodic theorem, 1931)**.** *Let* $\sigma\colon (\Omega,\mathbb{P}) \to (\Omega,\mathbb{P})$ *be an ergodic measure-preserving transformation. Let* $f\colon (\Omega,\mathbb{P}) \to \mathbb{R}$ *be an integrable function. Then*

$$\frac{1}{n}\sum_{i=0}^{n-1} f(\sigma^i x) \to \int f\,d\mathbb{P} \text{ for } \mathbb{P}\text{-a.e. } x.$$

Theorem 3 (Furstenberg, Kesten, 1960)**.** *Let* $\sigma\colon (\Omega,\mathbb{P}) \to (\Omega,\mathbb{P})$ *be an ergodic measure-preserving transformation. Let* $A\colon \Omega \to M_{d\times d}(\mathbb{R})$ *be a matrix-valued function with* $\int \log\|A(\omega)\|\,d\mathbb{P}(\omega) < \infty$*. Then*

$$\tfrac{1}{n}\log\|A(\sigma^{n-1}\omega)\cdots A(\omega)\| \to E \text{ for a.e. } \omega,$$

where $E = \lim_{n\to\infty}\frac{1}{n}\int \log\|A(\sigma^{n-1}\omega)\cdots A(\omega)\|\,d\mathbb{P}(\omega)$*.*

The Birkhoff and Furstenberg–Kesten theorems are immediate corollaries of Kingman's theorem.

1.2. Lyapunov Exponents and Multiplicative ergodic theorem.

If $T\colon I\to I$ is a differentiable self-map of the interval, then the chain rule gives $(T^n)'(x) = T'(T^{n-1}x)\cdot T'(T^{n-2}x)\cdots T'(x)$. The nth root of $(T^n)'(x)$ is a 'geometric average stretching rate' over n steps.

The *Lyapunov exponent for the one-dimensional map,* T at x is the logarithm of the limit of these rates: $\lambda(T,x) = \lim_{n\to\infty}\frac{1}{n}\log|(T^n)'(x)|$ (if it exists). That is: the derivative of T^n should be (logarithmically) close to $e^{n\lambda}$ (where 'logarithmically close' means between $e^{n(\lambda-\epsilon)}$ and $e^{n(\lambda+\epsilon)}$ for large n.)

We have

$$\lambda(T,x) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\log|T'(T^i x)|,$$

a Birkhoff sum.

In the particular case, $I=[0,1]$ and $T(x) = 4x(1-x)$, $|T'|(x) = 4|1-2x|$. T has an ergodic absolutely continuous invariant measure, μ with density $1/(\pi\sqrt{x(1-x)})$.

Now we can apply Birkhoff's theorem (writing $\phi(x) = \log|T'(x)| = \log 4 + \log|1-2x|$) to get

$$\begin{aligned}\lambda(T,x) &= \lim_{n\to\infty}\frac{1}{n}\sum_{0}^{n-1}\phi(T^i x)\\ &= \int_0^1 \phi(x)\,d\mu(x)\\ &= \log 4 + \int_0^1 \frac{\log|1-2x|}{\pi\sqrt{x(1-x)}}\,dx\\ &= \log 2\end{aligned}$$

for μ-a.e. $x \in [0,1]$.

We therefore expect $|T^{50}(0.3+10^{-30}) - T^{50}(0.3)|$ to be of the order of $e^{50\lambda}\cdot 10^{-30} = 2^{50}10^{-30} \approx 1.13\times 10^{-15}$. In fact, $|T^{50}(0.3+10^{-30}) - T^{50}(0.3)| \approx 3.44\times 10^{-16}$ (so the prediction was correct to one order of magnitude).

A positive Lyapunov exponent is one of the (many and inequivalent) definitions of 'chaos'.

Now we'll consider the case of a differentiable map, T, from a subset of $\mathbb{R}^d$ to itself (or a differentiable map from a manifold to itself). Writing $DT(x)$ for the Jacobian matrix of T at x, the Chain rule gives

$$DT^n(x) = DT(T^{n-1}x)\cdots DT(T(x))\cdot DT(x).$$

We'd like to make sense of how fast these matrices grow. We can apply the Furstenberg-Kesten theorem as soon as we have an invariant measure for T.

On the other hand, if A is a single matrix $\|A^n v\|$ grows at different rates depending on the eigenvectors that make up v. This suggests we might expect $DT^n(x)v$ to grow at different rates for different subspaces of $\mathbb{R}^d$. Also: the matrix $DT^n(x)$ depends on x, so we might expect the subspaces to depend on the point x.

Theorem 4 (Oseledets – non-invertible, 1965)**.** *Let σ be an ergodic measure-preserving transformation of $(\Omega,\mathbb{P})$. Let $A\colon \Omega\to M_{d\times d}(\mathbb{R})$ be a matrix-valued function with $\int \log\|A(\omega)\|\,d\mathbb{P}(\omega) < \infty$. Then there exist $\infty > \lambda_1 > \ldots > \lambda_k \geq -\infty$; $m_1,\ldots,m_k \in \mathbb{N}$ satisfying $m_1+\ldots+m_k = d$ and a measurable family of subspaces $F_1(\omega)$, $F_2(\omega)$, $\ldots F_k(\omega)$ such that*

(1) filtration*:* $\mathbb{R}^d = F_1(\omega) \supset F_2(\omega) \supset \ldots \supset F_k(\omega) \supset F_{k+1}(\omega) = \{0\}$*;*
(2) dimension*:* $\dim F_i(\omega) = m_i + \ldots + m_k$*; for a.e.* ω
(3) equivariance*:* $A(\omega)F_i(\omega) \subset F_i(\sigma(\omega))$ *for a.e.* ω

(4) growth*: If $v \in F_i(\omega) \setminus F_{i+1}(\omega)$ then $\frac{1}{n}\log\|A_\omega^{(n)}v\| \to \lambda_i$ for a.e. ω, where $A_\omega^{(n)} = A(\sigma^{n-1}\omega)\cdots A(\omega)$.*

The quantities λ_i are called *Lyapunov exponents* and the subspaces $F_i(\omega)$ are the collection of vectors expanding **at rate** λ_i **or less**.

The sequence of subspaces $F_1(\omega) \supset F_2(\omega) \supset \ldots \supset F_k(\omega)$ is called a *flag*.

Theorem 5 (Oseledets – invertible, 1965)**.** *Let σ be an* invertible *ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Let $A\colon \Omega \to GL(d, R)$ be a matrix-valued function with $\int \log\|A(\omega)\|\,d\mathbb{P}(\omega) < \infty$ and $\int \|(A(\omega))^{-1}\|\,d\mathbb{P}(\omega) < \infty$. Then there exist $\infty > \lambda_1 > \ldots > \lambda_k > -\infty$; $m_1, \ldots, m_k \in \mathbb{N}$ satisfying $m_1 + \ldots + m_k = d$ and measurable families of subspaces $V_1(\omega)$, $V_2(\omega)$, ..., $V_k(\omega)$ such that*

(1) decomposition*: $\mathbb{R}^d = V_1(\omega) \oplus V_2(\omega) \oplus \cdots \oplus V_k(\omega)$;*
(2) dimension*: $\dim V_i(\omega) = m_i$ for a.e. ω;*
(3) equivariance*: $A(\omega)V_i(\omega) = V_i(\sigma(\omega))$ for a.e. ω*
(4) growth*: If $v \in V_i(\omega) \setminus \{0\}$ then*

$\frac{1}{n}\log\|A_\omega^{(n)}v\| \to \lambda_i$ *and* $\frac{1}{n}\log\|A_\omega^{(-n)}v\| \to -\lambda_i$ *as $n \to \infty$ for a.e. ω,*

where

$$A_\omega^{(n)} = A(\sigma^{n-1}\omega)\cdots A(\omega) \text{ for } n \geq 0; \text{ and}$$
$$A_\omega^{(-n)} = A(\sigma^{-n}\omega)^{-1}\cdots A(\sigma^{-1}\omega)^{-1} \text{ for } n > 0.$$

The $V_i(\omega)$ are the vectors expanding **at rate** λ_i. These are the *Oseledets subspaces.*

2. Deducing Oseledets from Kingman: Preliminaries

In the next section, we'll sketch an argument of Raghunathan, giving a proof of the non-invertible form of Oseledets' theorem from the subadditive ergodic theorem.

As a warm-up, we need some reminders about the *singular value decomposition* of a matrix; and definition of the Grassmannian of a vector space; and the exterior algebra of a vector space.

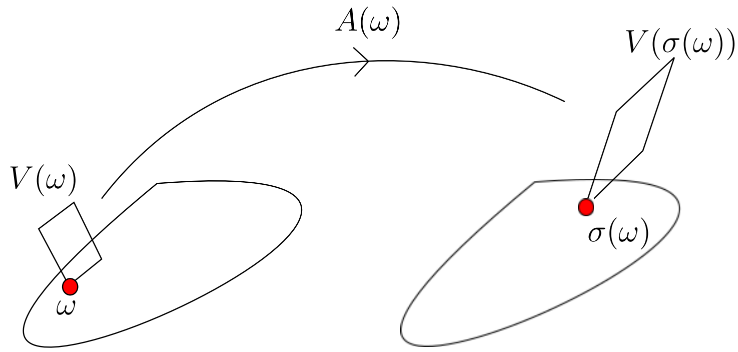


FIGURE 1. The Multiplicative Ergodic Theorem gives "A dynamical Jordan normal form decomposition."

2.1. **Singular Value Decomposition.**

Theorem 6 (Singular Value Decomposition)**.** *Let* $A \in M_{d\times d}(\mathbb{R})$*. Then there exist orthogonal matrices* O_1 *and* O_2 *and a diagonal matrix* D *with non-negative entries such that* $A = O_1 D O_2$*.*

Proof. The matrix A^*A is symmetric, and so there is an orthonormal basis of $\mathbb{R}^d$ consisting of eigenvectors. If $A^*Av = cv$, then $c = \langle A^*Av, v\rangle = \langle Av, Av\rangle \geq 0$, so that all eigenvalues are non-negative. Let the eigenvalues be $\lambda_1^2 \geq \lambda_2^2 \geq \ldots \geq \lambda_d^2$ with corresponding orthonormal eigenvectors $v_1, \ldots, v_d$. Let O_2 be the matrix with rows consisting of $v_1, \ldots, v_d$; D be the diagonal matrix with entries $\lambda_1, \ldots, \lambda_d$. Let k be the largest index such that $\lambda_k > 0$. For $i \leq k$, let $u_i = Av_i/\lambda_i$. If $k < d$, let $u_{k+1}, \ldots u_d$ be an orthonormal basis for $A(\mathbb{R}^d)^\perp$. Let O_1 be the matrix whose columns are $u_1, \ldots, u_d$.

Since the rows of O_2 are orthonormal, we see that $(O_2O_2^*)_{ij} = \langle v_i, v_j\rangle = \delta_{ij}$, so that O_2 is orthogonal. We have $O_1DO_2v_i = O_1De_i = O_1\lambda_ie_i = \lambda_iu_i = Av_i$, so that $O_1DO_2 = A$. Finally, notice that for $i < j \leq k$, $\lambda_i\lambda_j\langle u_i, u_j\rangle = \langle Av_i, Av_j\rangle = \langle v_i, A^Av_j\rangle = \lambda_j^2\langle v_i, v_j\rangle = 0$ for $i \neq j$, so that the first k columns of O_1 are orthonormal (and so are the rest by construction), so $O_1^*O_1 = I$ as required. □

The *singular values* of A, $\sigma_1(A) \geq \ldots \geq \sigma_d(A)$, are the entries of D. The *singular vectors* are the rows of O_2 and their images are multiples of the columns of O_1, so that $Av_i = \sigma_i(A)u_i$.

Remark. Singular value decomposition (SVD) also makes sense for non-square matrices.

Lemma 7. *Let the singular values of A be $\sigma_1 \geq \ldots \geq \sigma_d$. Then for $1 \leq k \leq d$,*

$$\sigma_k = \max_{\dim V = k} \left(\min_{x \in V : \|x\|=1} \|Ax\| \right); \text{ and}$$
$$\sigma_k = \min_{\operatorname{codim} V = k-1} \left(\max_{x \in V :\, \|x\|=1} \|Ax\| \right).$$

Proof. (Exercise) □

In particular, from this characterization, you can see that $\sigma_1(A)$ is $\max_{\|x\|=1} \|Ax\|$, the norm of A and the first singular vector is a vector that is expanded most by A. By continuity, any vector close to v_1 is also expanded a lot by A, but v_2 is a vector in $\operatorname{lin}(v_1)^\perp$ that is expanded the most by A. etc.:

Lemma 8. *v_k is a vector in $\operatorname{lin}(v_1, \ldots v_{k-1})^\perp$ that is expanded the most by A.*

Proof. (Exercise) □

2.2. **Grassmannian of a vector space.** The k-dimensional Grassmannian of $\mathbb{R}^d$, $\operatorname{Gr}(d,k)$ is the collection of all k-dimensional subspaces of $\mathbb{R}^d$. This is a very nice space: a compact metric space, a smooth manifold etc.

To define a metric, we'll go with the most intuitive one: $d_{\mathrm{Gr}}(V, V') = d_H(V \cap S, V' \cap S)$, where S is the unit ball and d_H is the Hausdorff distance: for two non-empty compact sets, their Hausdorff distance is defined by $d_H(K, K') = \max\big(\max_{x\in K} \min_{y \in K'} d(x,y), \max_{y \in K'} \min_{x \in K} d(x,y)\big)$.

Exercise. *Prove that $Gr(d, 2)$ is sequentially compact.*

Exercise (Harder - for those that like manifolds)**.** *For a fixed k-codimensional subspace, W, of $\mathbb{R}^d$, let $\mathcal{U}$ denote those elements of $Gr(d,k)$ that have a trivial intersection with W.*

Prove that $\mathcal{U}$ is an open subset of $Gr(d,k)$.

Fix an element $V_0 \in \mathcal{U}$, a basis $e_1, \ldots, e_k$ for V_0 and a basis $f_1, \ldots, f_{d-k}$ for W. For an element $V \in \mathcal{U}$, since $V \oplus W = \mathbb{R}^d$, each e_i may be expressed uniquely as $v_i + w_i$ with $v_i \in V$ and $w_i \in W$. Form a matrix, $\Phi(V)$ whose ith column is the coefficients of w_i in the (f_j) basis.

Prove that Φ is a bijection from $\mathcal{U}$ to $\mathbb{R}^{(d-k)\times k}$.

Prove that the collection of $(\mathcal{U}, \Phi)$ (as W varies over $Gr(d, d-k)$, V_0 varies over $\mathcal{U}$ and the bases vary over bases for W and V_0) forms a smooth manifold structure on $Gr(d,k)$.

2.3. **Exterior power of a vector space.** A very useful construction in multiplicative ergodic theory is that of an exterior power of a vector space. For the formal construction of the kth exterior power, if V is a vector space, you form the free vector space F with basis consisting of all elements of the form $e_{v_1,\ldots,v_k}$ for $(v_1,\ldots,v_k)\in V^k$ (so that a typical element is $17e_{v_1,\ldots,v_k}+2e_{24v_1,v_2,\ldots,v_k}+12e_{0,v_2,\ldots,v_k}$). We then let Z be a subspace of elements of F that we want to identify with 0: Z is the subspace of F spanned by elements of the form

$$e_{v_1,\ldots,cv+c'v',\ldots,v_k}-ce_{v_1,\ldots,v,\ldots,v_k}-c'e_{v_1,\ldots,v',\ldots,v_k}\ \text{(multilinearity)}$$
$$e_{v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k}+e_{v_1,\ldots,v_j,\ldots,v_i,\ldots,v_k}\ \text{(antisymmetry)}$$

The kth exterior power of V, $\bigwedge^k V$ is then F/Z. We write $v_1\wedge\cdots\wedge v_k$ for $e_{v_1,\ldots,v_k}+Z$.

Exercise. *Verify that:*

- $v_1\wedge\cdots\wedge v_i\wedge\cdots\wedge v_j\wedge\cdots\wedge v_k=-v_1\wedge\cdots\wedge v_j\wedge\cdots\wedge v_i\wedge\cdots\wedge v_k$; *and*
- $v_1\wedge\cdots\wedge(cv+c'v')\wedge\cdots\wedge v_k=v_1\wedge\cdots\wedge v\wedge\cdots\wedge v_k+c'v_1\wedge\cdots\wedge v'\wedge\cdots\wedge v_k$
- *If* $e_1,\ldots,e_d$ *is a basis for* V*, then* $\{e_{i_1}\wedge\cdots\wedge e_{i_k}\colon i_1<\ldots<i_k\}$ *spans* $\bigwedge^k V$.

In fact, if $e_1,\ldots,e_d$ is a basis for V then $\{e_{i_1}\wedge\cdots\wedge e_{i_k}\colon i_1<\ldots<i_k\}$ forms a basis for $\bigwedge^k V$, but proving this goes through a universal algebraic property of $\bigwedge^k V$.

Some elements of $\bigwedge^k V$ may be expressed in the form $v_1\wedge\cdots\wedge v_k$. Others can only be expressed as a sum of elements of this form (cf matrices expressed as sums of rank 1 matrices). A 'pure' vector $v_1\wedge\cdots\wedge v_k$ can be roughly thought of as defining an element of $\mathrm{Gr}(d,k)$ (i.e. $\mathrm{lin}(v_1,\ldots,v_k)$) and a magnitude.

Exercise. *Let* V *be a* k*-dimensional subspace of* $\mathbb{R}^d$*. Let two bases for* V *be* $e_1,\ldots,e_k$ *and* $f_1,\ldots,f_k$.

Prove that $f_1\wedge\cdots\wedge f_k=c\,e_1\wedge\cdots\wedge e_k$*, where* c *is the determinant of the matrix of coefficients of the* f *vectors in terms of the* e *vectors.*

If A is a linear self-map of V, then $\bigwedge^k A$ is a self-map of $\bigwedge^k V$ satisfying $\left(\bigwedge^k A\right)(v_1\wedge\cdots\wedge v_k)=(Av_1)\wedge\cdots(Av_k)$ for each $v_1\wedge\cdots\wedge v_k$.

Exercise. *Prove that there is a linear map,* $\bigwedge^k A$ *satisfying the above.*

The space $\bigwedge^k \mathbb{R}^d$ can be turned into a Euclidean space by letting $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : i_1 < \ldots < i_k\}$ be an orthonormal basis, where $e_1, \ldots, e_d$ is the standard basis.

It's completely non-obvious that if $f_1, \ldots, f_d$ is any orthonormal basis, then $\{f_{i_1} \wedge \cdots \wedge f_{i_k} : i_1 < i_2 < \ldots < i_k\}$ is orthonormal with respect to this inner product. But it's true!

Exercise. *Let u, v and w be three orthonormal vectors in $\mathbb{R}^3$. Prove that $u \wedge v$, $u \wedge w$ and $v \wedge w$ are orthonormal.*

2.4. **SVD of Exterior powers.** Singular value decomposition and exterior powers play *extremely* nicely together. Let A be a $d \times d$ matrix with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_d$ and singular vectors $v_1, \ldots, v_d$. Recall that these are orthonormal. By what we just said, $\{v_{i_1} \wedge \cdots \wedge v_{i_k} : i_1 < \ldots < i_k\}$ forms an orthonormal basis of $\bigwedge^k \mathbb{R}^d$.

Recall also that $Av_i = \sigma_i u_i$, where the u_i's are also orthonormal. This means that

$$\bigwedge^k A(v_{i_1} \wedge \cdots \wedge v_{i_k}) = (\sigma_{i_1} \cdots \sigma_{i_k}) u_{i_1} \wedge \cdots \wedge u_{i_k}.$$

The $\{u_{i_1} \wedge \cdots \wedge u_{i_k}\}$ are orthonormal also, so that we obtain

Lemma 9. *The singular values of $\bigwedge^k A$ are $\{\sigma_{i_1} \cdots \sigma_{i_k} : i_1 < \ldots < i_k\}$ and the singular vectors are $\{v_{i_1} \wedge \cdots \wedge v_{i_k} : i_1 < \ldots < i_k\}$.*

In particular,

$$\left\| \bigwedge^k A \right\| = \sigma_1 \cdots \sigma_k. \tag{1}$$

3. Deducing non-invertible Oseledets from Kingman

3.1. **The Raghunathan trick.** Recall the notation $A_\omega^{(n)} = A(\sigma^{n-1}\omega) \cdots A(\omega)$. For each $1 \leq k \leq d$, define $f_n^{\wedge k}(\omega) = \log \| \bigwedge^k A_\omega^{(n)} \|$. Since $A_\omega^{(n+m)} = A_{\sigma^n \omega}^{(m)} A_\omega^{(n)}$ and $\bigwedge^k (AB) = \bigwedge^k A \bigwedge^k B$, we see that

$$f_{n+m}^{\wedge k}(\omega) \leq f_m^{\wedge k}(\sigma^n \omega) + f_n^{\wedge k}(\omega).$$

Hence the Kingman sub-additive ergodic theorem (or Furstenberg-Kesten theorem) applies. There exist $L_1, \ldots, L_d$ such that $f_n^{\wedge k}(\omega)/n \to L_k$ for each k and a.e. ω.

Notice also that by (1), $f_n^{\wedge k}(\omega) = \log \| \bigwedge^k A_\omega^{(n)} \| = \sum_{i=1}^k \log \sigma_i(A_\omega^{(n)})$. Hence $f_n^{\wedge k}(\omega) - f_n^{\wedge (k-1)}(\omega) = \log \sigma_k(A_\omega^{(n)})$. Dividing by n and taking the limit, we obtain

$$\lim_{n \to \infty} \frac{1}{n} \log \sigma_k(A_\omega^{(n)}) = L_k - L_{k-1}.$$

Define $\mu_k = L_k - L_{k-1}$. By the above, we have

$$\infty > \int \|\log A(\omega)\|\, d\mathbb{P}(\omega) > \mu_1 \geq \mu_2 \geq \ldots \geq \mu_d \geq -\infty.$$

These are the *Lyapunov exponents*. It is useful to group them by multiplicity:

$$\{\lambda_1, \ldots, \lambda_k\} = \{\mu_1, \ldots, \mu_d\}$$
$$\infty > \lambda_1 > \lambda_2 > \ldots > \lambda_k \geq -\infty$$
$$\mu_{m_1+\ldots+m_{i-1}+j} = \lambda_i \text{ for } 1 \leq j \leq m_i.$$

3.2. **Equivariant subspaces.** That was the easy part! Now we need the subspaces... We're trying to find equivariant spaces $F_j(\omega)$ of dimension $m_d + \ldots + m_j$ consisting of vectors expanding at rate λ_j or lower, *"the jth slow space"*. We'll get at these using the slow singular vectors of $A_\omega^{(n)}$.

Let $M_{j-1} = m_1 + \ldots + m_{j-1}$ for $1 \leq j \leq k$. This is the dimension of the "$(j-1)$*st fast space*", the number of exponents larger than λ_j. The jth slow space should be spanned by singular vectors with exponents λ_j and below: by the $(M_{j-1}+1)$st to dth singular vectors. Let $O_j = m_j + \ldots + m_k$.

The idea is to define $F_j^{(n)}(\omega)$ to be the space spanned by the $(M_{j-1}+1)$st to dth singular vectors of $A_\omega^{(n)}$, and prove:

(1) these subspaces converge to a limit, $F_j(\omega)$, as $n \to \infty$;
(2) $F_j(\omega)$ is equivariant: $A(\omega)F_j(\omega) \subset F_j(\sigma(\omega))$;
(3) if $v \notin F_j(\omega)$, then $\liminf_{n\to\infty} \frac{1}{n} \log \|A_\omega^{(n)} v\| \geq \lambda_{j-1}$;
(4) if $v \in F_j(\omega)$, then $\limsup_{n\to\infty} \frac{1}{n} \log \|A_\omega^{(n)} v\| \leq \lambda_j$.

Of these, (1), (2) and (3) are relatively straightforward, while (4) is the trickiest.

3.3. **A sketch of** (1)**.** Remember that $\mathrm{Gr}(d, O_j)$ is compact metric (hence complete). The idea is to show that the distance from $F_j^{(n)}(\omega)$ to $F_j^{(n+1)}(\omega)$ is $O(e^{-n(\lambda_{j-1}-\lambda_j-\epsilon)})$. Then the subspaces form a 'fast Cauchy sequence'.

How to do this? Take a unit vector, v, in $F_j^{(n)}(\omega)$ and write it as an orthogonal sum $u + w$ of a part u in $F_j^{(n+1)}(\omega)$ (the span of the slow singular vectors for $n+1$ step evolution) and w in $F_j^{(n+1)}(\omega)^\perp$ (the fast singular vectors). Since we know that $\|A^{(n)} v\| \lesssim e^{n\lambda_j}$, [2] it follows that

[2] I'll write $x_n \lesssim e^{an}$ to mean for any ϵ, $x_n \leq e^{(a+\epsilon)n}$ for large n. That is, the exponential growth rate is at most a.

$\|A^{(n+1)}v\| \lesssim e^{n\lambda_j}$. But $\|A_\omega^{(n+1)}v\|$ is the sum of the orthogonal vectors $A_\omega^{(n+1)}u$ and $A_\omega^{(n+1)}w$. Hence $\|A_\omega^{(n+1)}w\| \lesssim e^{n\lambda_j}$. Since w is in the fast space (so grows at rate λ_{j-1} or faster), this implies $\|w\| \lesssim e^{n(\lambda_j-\lambda_{j-1})}$. That is: every unit vector in $F_j^{(n)}(\omega)$ is $e^{-n(\lambda_{j-1}-\lambda_j-\epsilon)}$-close to something in $F_j^{(n+1)}(\omega)$.

3.4. **A sketch of** (2)**.** We show that elements of $A(\omega)(F_j^{(n+1)}(\omega))$ are exponentially close to $F_j^{(n)}(\sigma(\omega))$ and take the limit as $n \to \infty$ using claim (1).

Take $v = A(\omega)z$ in the unit ball of $A(\omega)(F_j^{(n+1)}(\omega))$; express it as $u + w$ with $u \in F_j^{(n)}(\omega)$ and $w \in F_j^{(n)}(\omega)^\perp$. The rest of the argument is like the previous step.

3.5. **A sketch of** (3)**.** If v is a unit vector not in $F_j(\omega)$, it is some positive distance, δ, from $F_j(\omega)$. By the triangle inequality, it is at least $\frac{\delta}{2}$ from $F_j^{(n)}(\omega)$ for all large n. That means that if v is decomposed into components in the slow space, $F_j^{(n)}(\omega)$ and the fast space, $F_j^{(n)}(\omega)^\perp$, there is a vector of length at least $\frac{\delta}{2}$ in the fast space. When you apply $A_\omega^{(n)}$, you get a vector of length $\frac{\delta}{2}e^{n(\lambda_{j-1}-\epsilon)}$, as required.

3.6. **Sketch of a sketch of** (4)**.** Write $V_i^{(n)}(\omega)$ for the space spanned by the $(M_{i-1}+1)$st to M_ith singular vectors of $A_\omega^{(n)}$. The idea is to

> Show that if v is a unit vector in $F_j(\omega)$, then the component of v in $V_i^{(n)}$ is of size at most $e^{(\lambda_j-\lambda_i+\epsilon)n}$ for each $i < j$.

Now when you apply $A_\omega^{(n)}$ to v, the vector obtained is of size at most $e^{(\lambda_j+\epsilon)n}$ (as seen working component by component and using the triangle inequality).

Raghunathan shows the above by clever estimates on the inverse of a matrix.

As an alternative, step (1) already gives the desired estimate in the case of $F_2(\omega)$. This is enough to show that elements of $F_2(\omega)$ grow at rate λ_2 or less. Now, one can look at the restriction of $A(\omega)$ to $F_2(\omega)$ and deduce that $F_3(\omega)$ grows at rate λ_3 or less and obtain the result inductively. (This argument is carried out in a Banach space setting in papers of Alex Blumenthal, and of Cecilia González-Tokman and myself).

4. Deducing invertible Oseledets from non-invertible Oseledets

For this section, we're assuming that the base dynamics, σ, is invertible, and also that the matrices $A(\omega)$ are invertible (and $\|(A(\omega))^{-1}\|$ is log-integrable). It turns out that the first condition is crucial, whereas the second condition is not.

Recall the definition of the stable and unstable manifolds of a fixed point of an invertible map T

$$W_s(p) = \{x\colon d(T^n x, p) \to 0 \text{ as } n \to \infty\};$$
$$W_u(p) = \{x\colon d(T^n x, p) \to 0 \text{ as } n \to -\infty\}.$$

At first sight, the definition of the unstable manifold may be surprising:

Exercise. *Why is the unstable manifold defined this way? (Think about the map* $T(x) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} x$ *for a concrete example).*

Like this, we will obtain fast spaces as the slow spaces of the inverse system.

Exercise. *Show that if the invertible matrix* A *has singular values* $\sigma_1 > \ldots > \sigma_d$*, then* A^{-1} *has singular values* $\sigma_d^{-1} > \ldots > \sigma_1^{-1}$.

4.1. Non-invertible implies invertible using The Inverse system.

The map σ^{-1} is another ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Define the matrix $B(\omega) = A(\sigma^{-1}\omega)$ and build the matrix cocycle $B_\omega^{(n)} = B(\sigma^{-(n-1)}\omega) \cdots B(\omega)$. Notice that $B_\omega^{(n)} = (A_{\sigma^{-n}\omega}^{(n)})^{-1}$.

Exercise. *Check from the previous exercise and the Kingman subadditive ergodic theorem that the Lyapunov exponents for this cocycle are* $-\lambda_k, \ldots, -\lambda_1$*, with multiplicities* $m_k, \ldots, m_1$.

Applying the one-sided Oseledets theorem to the inverse system, we obtain a family of subspaces $\mathbb{R}^d = E_k(\omega) \supset E_{k-1}(\omega) \supset \cdots E_1(\omega)$ such that:

- (dimension): $\dim E_j(\omega) = m_1 + \ldots + m_j$;
- (equivariance): $B(\omega)E_j(\omega) \subset E_j(\sigma^{-1}\omega)$;
- (growth): $v \in E_j(\omega) \setminus E_{j+1}(\omega)$ implies $\frac{1}{n} \log \|B_\omega^{(n)} v\| \to -\lambda_j$;

Since $B(\omega) = A(\sigma^{-1}(\omega))^{-1}$, the equivariance condition can be rephrased as $E_j(\omega) \subset A(\sigma^{-1}(\omega))E_j(\sigma^{-1}\omega)$, or $E_j(\sigma(\omega)) \subset A(\omega)E_j(\omega)$. Since $A(\omega)$ is invertible, and $\dim E_j(\omega) = m_1 + \ldots + m_j$ for a.e. ω, we deduce $E_j(\omega)$ *is an equivariant family.*

If $v \in E_j(\omega)$, then $\|B^{(n)}_\omega v\| \lesssim e^{-(\lambda_j-\epsilon)n}\|v\|$. Since $B^{(n)}_\omega = (A^{(n)}_{\sigma^{-n}\omega})^{-1}$, it follows that for $w \in E_j(\sigma^{-n}\omega)$ (writing w as $(A^{(n)}_{\sigma^{-n}\omega})^{-1}v$), $\|w\| \lesssim e^{-(\lambda_j-\epsilon)n}\|A^{(n)}_{\sigma^{-n}\omega}w\|$ or $\|A^{(n)}_{\sigma^{-n}\omega}w\| \gtrsim e^{(\lambda_j-\epsilon)n}\|w\|$.

This (plus a little more work) shows that $E_j(\omega)$ is the jth fast space: the vectors expanding at rate λ_j or faster.

Now: $V_j(\omega) = E_j(\omega) \cap F_j(\omega)$ is an equivariant space consisting of vectors expanding at exactly rate λ_j. The last thing to check is that it has the correct dimension, m_j. Since $\dim E_j(\omega) = m_1 + \ldots + m_j$ and $\dim F_j(\omega) = m_j + \ldots + m_k = (d - \dim E_j(\omega)) + m_j$, we see from the formula $\dim(U \cap V) = \dim U + \dim V - \dim(U+V)$ that $\dim V_j(\omega) \geq m_j$. It is not hard to see that the $(V_j(\omega))$ are mutually linearly independent: Suppose that that $v_1 + \ldots + v_k = 0$, where $v_i \in V_i(\omega)$. Suppose for a contradiction that the v_i are not all 0. Then let ℓ be the smallest index such that $v_\ell \neq 0$. Now $A^{(n)}_\omega v_\ell$ grows at rate λ_ℓ, while $A^{(n)}_\omega(v_{\ell+1} + \ldots + v_k)$ grows at rate at most $\lambda_{\ell+1}$, so that they cannot cancel for large n, contradicting the assumption that $v_1 + \ldots + v_k = 0$ (hence $A^{(n)}_\omega(v_1 + \ldots + v_k) = 0$).

A key observation: *The $E_j(\omega)$ were the slow spaces for the inverse system – that is these are determined by $(A(\sigma^n\omega))_{n<0}$, while the $F_j(\omega)$ are governed by $(A(\sigma^n\omega))_{n\geq 0}$.*

4.2. **Non-invertible implies invertible using duality.** In this subsection, we'll prove the same result using duality. It is still important that σ is invertible, but we never take inverses of the matrices.

Define $C(\omega) = A(\sigma^{-1}\omega)^*$ and build a cocycle over σ^{-1}: $C^{(n)}_\omega = C(\sigma^{-(n-1)}\omega)\cdots C(\omega) = (A^{(n)}_\omega)^*$.

Theorem 10. *Let σ be an ergodic invertible measure-preserving transformation. Let $A\colon \Omega \to M_{d\times d}(\mathbb{R})$ be such that $\|A(\cdot)\|$ is log-integrable.*

Let $C^{(n)}_\omega$ be the dual cocycle over σ^{-1} as above. Then the Lyapunov exponents of the dual cocycle are the same as those of $A^{(n)}_\omega$.

Let the slow spaces for the dual cocycle be $G_1(\omega), \ldots G_k(\omega)$. Then:

1. *$A(\omega)G_j(\omega)^\perp = G_j(\sigma(\omega))^\perp$ for a.e. ω;*
2. *$G_j(\omega)^\perp \cap F_{j-1}(\omega) = V_{j-1}(\omega)$ for a.e. ω.*

Proof. To prove (1), let $v \in G_j(\omega)^\perp$ and $y \in G_j(\sigma(\omega))$. Then we have

$$\langle A(\omega)v, y\rangle = \langle v, A(\omega)^* y\rangle = \langle v, C(\sigma(\omega))y\rangle.$$

Since $C(\sigma(\omega))G_j(\sigma(\omega)) \subset G_j(\omega)$, we have $C(\sigma(\omega))y \in G_j(\omega)$, so that $\langle A(\omega)v, y\rangle = 0$ and $A(\omega)v \in G_j(\sigma(\omega))^\perp$, as required.

We'll just sketch the proof of (2). The main idea is to show that $G_j(\omega)^\perp$ has a trivial intersection with $F_j(\omega)$. Assuming this for now, since they have complementary dimensions ($F_j(\omega)$ and $G_j(\omega)$ have the same dimension as the A and C cocycles have the same Lyapunov exponents), it will then follow that $\mathbb{R}^d = F_j(\omega) \oplus G_j(\omega)^\perp$. From here (and (3) of section 3.2), it follows that everything in $G_j(\omega)^\perp$ expands at rate λ_{j-1} or faster. Now we have $G_j(\omega)^\perp \cap F_{j-1}$ is an equivariant subspace consisting of vectors expanding at rate λ_{j-1}. From the formula $\dim(U\cap V) = \dim U + \dim V - \dim(U+V)$, we see that $G_j(\omega)^\perp \cap F_{j-1}$ is of dimension m_{j-1}.

To prove the trivial intersection, let $Z = F_j(\omega)^\perp$. By section 3.2 (3), we have $\|A_\omega^{(n)} z\| \gtrsim e^{\lambda_{j-1}n}$ for all $z \in Z \cap S$. On the other hand, we have $d(A_\omega^{(n)} z, G_j(\sigma^n\omega)^\perp) = \max_{y\in G_j(\sigma^n\omega)\cap S}\langle A_\omega^{(n)} z, y\rangle$. For any $y \in G_j(\sigma^n\omega) \cap S$, we have $\langle A_\omega^{(n)} z, y\rangle = \langle z, B_{\sigma^n\omega}^{(n)} y\rangle \lesssim e^{\lambda_j n}$. Hence for any $z \in Z \cap S$, the component of $A_\omega^{(n)} z$ in the direction perpendicular to $G^\perp(\sigma^n\omega)$ is $\lesssim e^{\lambda_j n}$. We deduce $\angle(A_\omega^{(n)} Z, G^\perp(\sigma^n\omega)) \lesssim e^{-n(\lambda_{j-1}-\lambda_j)}$.

To finish, we show that elements of $A_\omega^{(n)} Z$ are forced to lie far from $F_j(\sigma^n\omega)$. One can show (with a little determinant magic) $\|\bigwedge^k A_\omega^{(n)}\| \approx \|\bigwedge^k A_\omega^{(n)}|_{\bigwedge^k Z}\|$ for all large n and so $\|\bigwedge^k A_\omega^{(n)}|_{\bigwedge^k Z}\| \approx e^{n(m_1\lambda_1+\ldots+m_{j-1}\lambda_{j-1})}$. If an element z of Z had the property that $A_\omega^{(n)} z$ was e^{-an} close (in angle) to $F_j(\sigma^n\omega)$, then the above growth condition is contradicted. In particular, we deduce $A_\omega^{(n)} Z$ is far (at an exponential scale) in every direction from $F_j(\sigma^n\omega)$; but $A_\omega^{(n)} Z$ is close to $G_j(\sigma^n\omega)$. Hence $F_j(\sigma^n\omega) \cap G_j(\sigma^n\omega) = \{0\}$ for large n. Hence $F_j(\omega) \cap G_j(\omega) = \{0\}$ a.e. by the Poincaré recurrence theorem.

□

Corollary 11 (Oseledets theorem: semi-invertible case)**.** *Let σ be an* invertible *ergodic measure-preserving transformation of* $(\Omega, \mathbb{P})$*. Let $A\colon \Omega \to M_{d\times d}$ be a matrix-valued function with $\int \log\|A(\omega)\|\,d\mathbb{P}(\omega) < \infty$. Then there exist $\infty > \lambda_1 > \ldots > \lambda_k \geq -\infty$; $m_1, \ldots, m_k \in \mathbb{N}$ satisfying $m_1 + \ldots + m_k = d$ and measurable families of subspaces $V_1(\omega)$, $V_2(\omega)$, $\ldots$, $V_k(\omega)$ such that*

(1) decomposition*: $\mathbb{R}^d = V_1(\omega) \oplus V_2(\omega) \oplus \cdots \oplus V_k(\omega)$;*
(2) dimension*: $\dim V_i(\omega) = m_i$ for a.e. ω;*
(3) equivariance*: $A(\omega)V_i(\omega) = V_i(\sigma(\omega))$ for a.e. ω*
(4) growth*: If $v \in V_i(\omega) \setminus \{0\}$ then $\frac{1}{n}\log\|A_\omega^{(n)} v\| \to \lambda_i$ as $n \to \infty$ for a.e. ω, where $A_\omega^{(n)} = A(\sigma^{n-1}\omega)\cdots A(\omega)$.*

The hypotheses are a hybrid of the two original Oseledets theorems: the underlying system must be invertible; there is no invertibility requirement for the matrices. The good news: we can still get a decomposition: $\mathbb{R}^d = V_1 \oplus \ldots \oplus V_k$ rather than a filtration ($\mathbb{R}^d = F_1 \supset \ldots \supset F_k$). We did lose something though: we have no backwards growth bounds on $\|A_\omega^{(-n)} v\|$ – the inverse matrices needn't even exist.

Theorem 12 (Oseledets theorem: Banach space version)**.** *Let σ be an invertible ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Let B be a separable Banach space. Let $\mathcal{L}\colon \Omega \to L(B,B)$ be an operator-valued function with $\int \log \|\mathcal{L}_\omega\| \, d\mathbb{P}(\omega) < \infty$.*

Suppose that $\frac{1}{n} \int \log \|\mathcal{L}_\omega^{(n)}\| \, d\mathbb{P}(\omega) \to \lambda$ and $\frac{1}{n} \int \log \kappa(\mathcal{L}_\omega^{(n)}) \, d\mathbb{P}(\omega) \to \alpha < \lambda$, where $\kappa(\mathcal{L}) = \inf\{r \colon \mathcal{L}(B)$ can be covered by balls of radius $r\}$ and B is the unit ball.

Then there exist $1 \le k \le \infty$, $\lambda_1 > \lambda_2 > \ldots > \lambda_k$ and equivariant subspaces $V_1(\omega), \ldots, V_k(\omega)$ and $R(\omega)$ such that $B = V_1(\omega) \oplus \ldots \oplus V_k(\omega) \oplus R(\omega)$ and the growth conditions of the matrix Oseledets theorem hold.

The proofs are based on defining suitable notions of singular values (or volume growth) for maps of linear maps on Banach spaces. There are *many* possibilities – all giving the same growth rates.