CHILE LECTURES: MULTIPLICATIVE ERGODIC THEOREM, FRIENDS AND APPLICATIONS

1. Motivation

Context: $(\Omega, \mathbb{P})$ a probability space; $\sigma\colon \Omega \to \Omega$ is a measure-preserving transformation $\mathbb{P}(\sigma^{-1}A) = \mathbb{P}(A)$ for all measurable sets A. Generally, $\mathbb{P}$ will be *ergodic* as well: $\sigma^{-1}A = A$ implies $\mathbb{P}(A)$ is 0 or 1. This is an irreducibility condition.

This is sufficient for the following form of Birkhoff's theorem (ergodic theory's strong law of large numbers):

Theorem 1 (Birkhoff)**.** *Let σ be an ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Let $f \in L^1(\Omega)$ (in fact $f^+ \in L^1$ suffices). Then for $\mathbb{P}$-a.e. ω,*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(\sigma^n\omega) \to \int f(x)\, d\mathbb{P}(x) \text{ as } N \to \infty.$$

The *Kingman sub-additive ergodic theorem* is an important generalization. A sequence of functions, (f_n) is sub-additive if $f_1^+ \in L^1(\mathbb{P})$; and $f_{n+m}(\omega) \le f_n(\omega) + f_m(\sigma^n\omega)$ for each ω.

Theorem 2 (Kingman sub-additive ergodic theorem (1968))**.** *Let σ be an ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$ and let (f_n) be a sub-additive sequence of functions satisfying $\int f_1^+\, d\mathbb{P} < \infty$*

Then for $\mathbb{P}$-a.e. ω,

$$\tfrac{1}{n} f_n(\omega) \to \lim_{k\to\infty} \frac{1}{k}\int f_k(x)\, d\mathbb{P}(x) \text{ as } n \to \infty.$$

Example 3 (First Passage Site Percolation)**.** Let Ω be a probability space, and consider a family of measure-preserving maps $(\sigma_{\mathbf{n}})_{\mathbf{n}\in\mathbb{Z}^d}$ from Ω to Ω such that $\sigma_{\mathbf{n}+\mathbf{m}} = \sigma_{\mathbf{n}} \circ \sigma_{\mathbf{m}}$ Let $f\colon \Omega \to (0,\infty)$ be a sequence of weights and define

$$T_{\mathbf{a},\mathbf{b}}(\omega) = \inf_{\mathbf{a}=\mathbf{v}_0\to\mathbf{v}_1\to\cdots\to\mathbf{v}_n=\mathbf{b}} \sum_{i=0}^{n-1} f(\sigma_{\mathbf{v}_i}\omega),$$

the time of the shortest path from $\mathbf{a}$ to $\mathbf{b}$.

Now certainly $T_{\mathbf{a},\mathbf{c}}(\omega) \le T_{\mathbf{a},\mathbf{b}}(\omega) + T_{\mathbf{b},\mathbf{c}}(\omega)$. We also have $T_{\mathbf{a},\mathbf{b}}(\omega) = T_{\mathbf{0},\mathbf{b}-\mathbf{a}}(\sigma_{\mathbf{a}}\omega)$. In particular, for a fixed $\mathbf{v} \in \mathbb{Z}^d$,

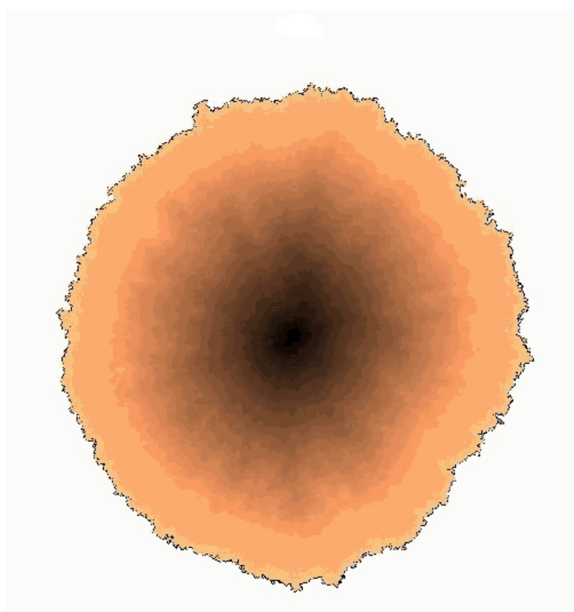
$$T_{\mathbf{0},(n+m)\mathbf{v}}(\omega) \le T_{\mathbf{0},n\mathbf{v}}(\omega) + T_{\mathbf{0},m\mathbf{v}}(\sigma_{n\mathbf{v}}\omega) = T_{\mathbf{0},n\mathbf{v}}(\omega) + T_{\mathbf{0},m\mathbf{v}}(\sigma_{\mathbf{v}}^n\omega).$$

Hence $t_n(\omega) = T_{\mathbf{0},n\mathbf{v}}(\omega)$ is a sub-additive sequence of functions (with respect to the transformation $\sigma_{\mathbf{v}}$. Hence $\frac{1}{n}T_{0,n\mathbf{v}}(\omega)$ converges to a constant, which we will call $1/s_{\mathbf{v}}$, for all $\mathbf{v} \in \mathbb{Z}^d$ and almost all $\omega \in \Omega$.

$s_{\mathbf{v}}$ is the rate at which the water spreads in direction $\mathbf{v}$. You can also make sense of $s_{\mathbf{v}}$ for non-integer $\mathbf{v}$'s by looking at the time to get to the nearest lattice point to $n\mathbf{v}$. A calculation shows $s_{t\mathbf{v}} = (1/t)s_{\mathbf{v}}$, so it is natural to look at $s_{\mathbf{v}}$ for $\mathbf{v}$ with $\|\mathbf{v}\| = 1$.

The set $S = \{x\mathbf{v} \colon x \le s_{\mathbf{v}}\}$ is called the *shape* of the percolation process.

Corollary 4 (Shape Theorem)**.** *For a.e.* ω, $\mathsf{Occ}_t(\omega)/t \to S$, *where* $\mathsf{Occ}_t(\omega)$ *is the set of sites that are occupied by time* t.



Example 5 (Matrix cocycles)**.** Let σ, $\mathbb{P}$ and Ω be as above. Let $A \colon \Omega \to M_d(\mathbb{R})$ and write A_ω for $A(\omega)$. Then a matrix cocycle is given by $A \colon \mathbb{N} \times \Omega \to M_d(\mathbb{R})$ by

$$A_\omega^{(n)} = A_{\sigma^{n-1}\omega} \cdots A_{\sigma\omega} A_\omega.$$

These matrices satisfy the cocycle relation: $A_\omega^{(n+m)} = A_{\sigma^n\omega}^{(m)} A_\omega^{(n)}$.

Equipping the matrices with the norm $\|A\| = \max_{\|x\|=1} \|Ax\|$, it's immediate that $\|AB\| \le \|A\|\,\|B\|$, so the functions $\omega \mapsto \|A_\omega^{(n)}\|$ as n runs over $\mathbb{N}$ form a sub-multiplicative sequence.

In particular, $f_n(\omega) = \log \|A_\omega^{(n)}\|$ is a sub-additive sequence of functions

Corollary 6 (Furstenberg-Kesten (1960))**.** *Let σ be an ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Let $A\colon \Omega \to M_d(\mathbb{R})$ satisfy $\int \log \|A_\omega\| \, d\mathbb{P}(\omega) < \infty$. Then there is an α such that for $\mathbb{P}$-a.e. ω, $\frac{1}{n} \log \|A_\omega^{(n)}\| \to \alpha$.*

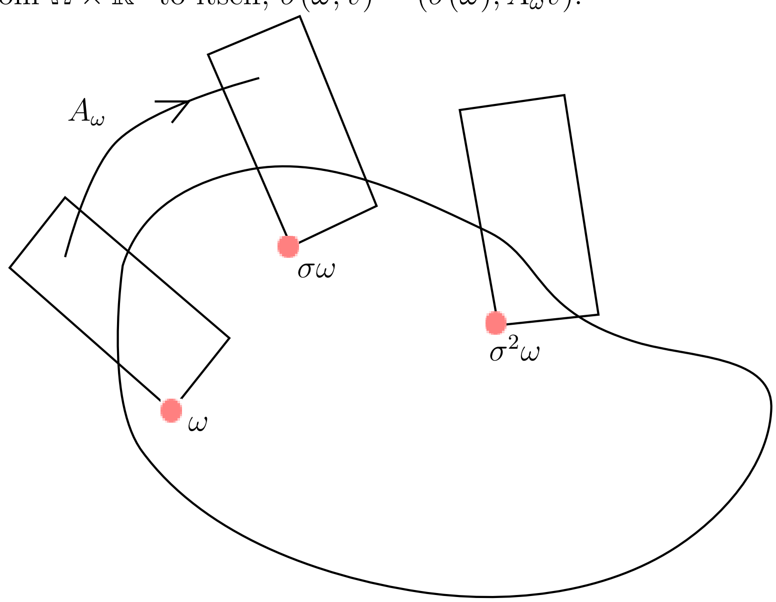
For a constant matrix cocycle: $A_\omega = M$ for each ω, $\frac{1}{n} \log \|M^n\| \to \log \rho(M)$, the spectral radius (equal to the logarithm of absolute value of the largest eigenvalue).

$\frac{1}{n} \log \|A_\omega^{(n)}\| \to \alpha$ is equivalent to the statement: for each $\epsilon > 0$, for all sufficiently large n, $e^{n(\alpha-\epsilon)} \le \|A^{(n)}\| \le e^{n(\alpha+\epsilon)}$. We will informally write $\|A^{(n)}\| \approx e^{n\alpha}$, so that α is the exponential growth rate of the matrix cocycle.

For a single matrix, and for any $\mathbf{v}$, $\|M^n \mathbf{v}\|$ grows roughly like a^npoly(n), where a is the absolute value of some eigenvalue. One might hope for a refinement of Furstenberg-Kesten giving analogues of the other eigenvalues.

Question. *What can be said about the matrix $A_\omega^{(n)}$ for typical ω?*

It is helpful to think of the matrix cocycle as a vector bundle map from $\Omega \times \mathbb{R}^d$ to itself, $\bar{\sigma}(\omega, v) = (\sigma(\omega), A_\omega v)$.



Since different matrices are being applied at each ω, there's no reason to hope for the same vector space decomposition over each point.

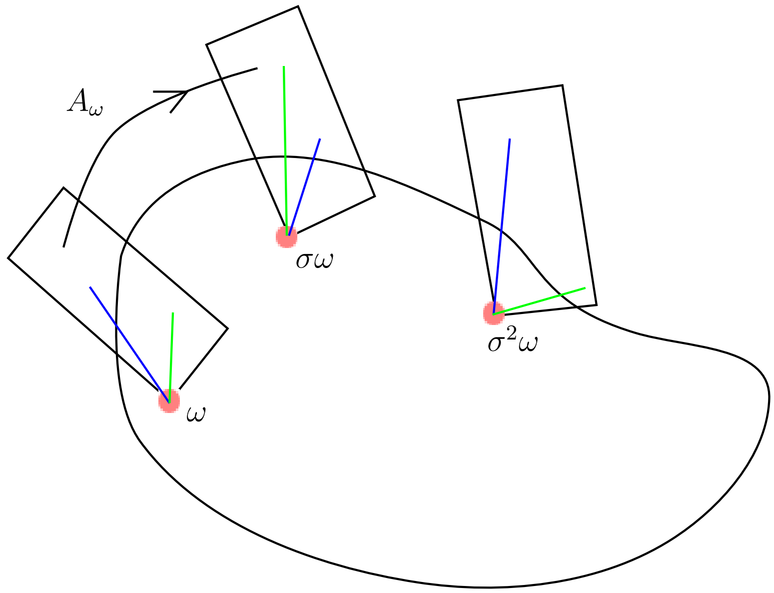


FIGURE 1. Oseledets subspaces

Theorem 7 (Oseledets Multiplicative Ergodic Theorem (invertible) (1965))**.** *Let σ be an* ***invertible*** *ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Let $A\colon \Omega \to GL_d(\mathbb{R})$ be a cocycle of* ***invertible*** *matrices with $\int \log \|A_\omega^{\pm 1}\| \, d\mathbb{P}(\omega) < \infty$.*

Then there exist $\infty > \lambda_1 > \lambda_2 > \ldots > \lambda_k > -\infty$ and subspaces $V_1(\omega), \ldots, V_k(\omega)$ satisfying

(1) ***Decomposition****: $\mathbb{R}^d = V_1(\omega) \oplus \ldots \oplus V_k(\omega)$;*
(2) ***Equivariance****: $A_\omega(V_i(\omega)) = V_i(\sigma(\omega))$;*
(3) ***Growth****: If $v \in V_i(\omega) \setminus \{0\}$, then $\frac{1}{n} \log \|A_\omega^{(n)} v\| \to \lambda_i$ as $n \to \pm\infty$.*

This may be thought of as something like a dynamical Jordan normal form.

The $V_i(\omega)$ are the *Oseledets subspaces* and the λ_i are the *Lyapunov exponents*.

There is also a non-invertible version.

Theorem 8 (Oseledets Multiplicative Ergodic Theorem (non-invertible))**.** *Let σ be an ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Let $A\colon \Omega \to M_d(\mathbb{R})$ be a cocycle of matrices with $\int \log \|A_\omega\| \, d\mathbb{P}(\omega) < \infty$.*

Then there exist $\infty > \lambda_1 > \lambda_2 > \ldots > \lambda_k > -\infty$ and subspaces $U_1(\omega), \ldots, U_k(\omega)$ satisfying:

(1) ***Filtration****: $\mathbb{R}^d = U_1(\omega) \supset \ldots \supset U_k(\omega)$;*

FIGURE 2. A decreasing collection of subspaces $U_1 \supset U_2 \supset \cdots U_k$ is called a *filtration* or a *flag*. A physical flag has a distinguished 1-dimensional subspace (the direction of the pole) that is a subspace of a distinguished 2-dimensional subspace (the plane of the flag) sitting in $\mathbb{R}^3$.

(2) ***Equivariance***: $A_\omega(U_i(\omega)) \subset U_i(\sigma(\omega))$;
(3) ***Growth***: *If* $v \in U_i(\omega) \setminus U_{i+1}(\omega)$, *then* $\frac{1}{n}\log\|A_\omega^{(n)}v\| \to \lambda_i$.

I sometimes call the $U_i(\omega)$ *slow spaces* – they are the vectors whose exponential growth rate is at most λ_i:

$$U_i(\omega) = \{v\colon \limsup \tfrac{1}{n}\log\|A_\omega^{(n)}v\| \le \lambda_i\}.$$

Notice: just like for a single matrix, the vectors growing at rate λ or less form a linear space; the vectors growing at rate λ or more do not form a linear space. I may also write $V_{\le\lambda_i}(\omega)$ for the slow spaces.

Fast spaces also play a role: $V_{\ge\lambda_i}(\omega)$.

Importantly, $\boxed{V_{\ge\lambda_i}(\omega) \ne \{v\colon \limsup \tfrac{1}{n}\log\|A_\omega^{(n)}v\| \ge \lambda_i\}}$.

Rather a fast space can be defined to be the largest equivariant measurable family subspace such that each vector expands at a rate at least λ_i.

1.1. **Oseledets theorem motivation.** Suppose that Ω is a manifold, σ is an invertible smooth map from Ω to itself and $\mathbb{P}$ is a (nice) ergodic σ-invariant measure.

Then the Multiplicative ergodic theorem gives for $\mathbb{P}$-a.e. ω, subspaces of the tangent space, $V_1(\omega), \ldots, V_k(\omega)$ each with a different exponent $\lambda_1, \ldots, \lambda_k$. If $\lambda_1, \ldots, \lambda_j$ are positive, while $\lambda_{j+1}, \ldots, \lambda_k$ are negative, then the directions $E^u(\omega) = V_1(\omega) \oplus \ldots \oplus V_j(\omega)$ are unstable directions at ω: perturbations in these directions tend to grow; while $E^s(\omega) =$

$V_{j+1}(\omega) \oplus \ldots \oplus V_k(\omega)$ are stable directions. They are also equivariant: $D\sigma(E^u(\omega)) = E^u(\sigma(\omega))$ and $D\sigma(E^s(\omega)) = E^s(\sigma(\omega))$.

If you can 'integrate' these subspaces (patch them together), you obtain stable and unstable manifolds at ω for $\mathbb{P}$-a.e. ω.

1.2. **Non-invertible to invertible I.** Notice that the hypotheses for the non-invertible theorem are satisfied in the invertible case. You can deduce the invertible theorem from the non-invertible theorem as follows. Let σ be an ergodic invertible measure-preserving transformation; let (A_ω) be a cocycle of invertible matrices (with $\int \log \|A_\omega\| \, d\mathbb{P} < \infty$ and $\int \log \|A_\omega^{-1}\| \, d\mathbb{P} < \infty$).

(1) Form the spaces $V_{\leq \lambda_i}(\omega)$ from the non-invertible version of the theorem.
(2) Form a new cocycle with base dynamical system σ^{-1} and generator $B_\omega = A^{-1}_{\sigma^{-1}\omega}$. Notice that $B^{(n)}_\omega = (A^{(n)}_{\sigma^{-n}\omega})^{-1}$. Hence the exponents of the B cocycle are $-\lambda_k > \ldots > -\lambda_1$.

    Form spaces $W_{\leq -\lambda_i}(\omega)$ consisting of vectors which when iterated backwards have expansion rate at most $-\lambda_i$. These same vectors (using equivariance here), when iterated forwards have expansion rate at least λ_i, so that $W_{\leq -\lambda_i}(\omega) = U_{\geq \lambda_i}(\omega)$.
(3) Finally $V_i(\omega) = U_{\leq \lambda_i}(\omega) \cap W_{\leq -\lambda_i}(\omega)$.

Corollary 9. *The slow spaces are determined by the future:* $V_{\leq \lambda_i}(\omega)$ *is determined by* $(A_{\sigma^n\omega})_{n \geq 0}$*; the fast spaces are determined by the past:* $V_{\geq \lambda_i}(\omega)$ *is determined by* $(A_{\sigma^n\omega})_{n<0}$.

I'll mention two other ways to obtain the invertible from the non-invertible case later.

2. Tricks of the trade

2.1. **Singular value decomposition (SVD).** Given a square (or not square) matrix, A, there are essentially unique orthonormal bases $\mathbf{v}_1, \ldots, \mathbf{v}_d$ and $\mathbf{w}_1, \ldots, \mathbf{w}_d$ and non-negative μ_i such that $A\mathbf{v}_i = \mu_i \mathbf{w}_i$. I like to write the μ_i in decreasing order. These are the *singular values* of A. The non-uniqueness is in choice of signs in the $\mathbf{v}_i$, or orthonormal basis choice if a singular value is repeated. The $\mathbf{v}_i$ are the *singular vectors* of A.

As an alternative formulation of SVD, A may be expressed as $O_1 D O_2$ where O_1 and O_2 are orthogonal matrices and D is diagonal. Here the rows of O_2 are the $\mathbf{v}_i$ and the columns of O_1 are the $\mathbf{w}_i$.

Geometrically, a linear map A maps a sphere to an ellipsoid. The $\mathbf{w}_i$ are the semi-principal axes and the $\mathbf{v}_i$ are their pre-images.

The proof of the singular value decomposition is straightforward, using diagonalization of $A^T A$. The eigenvectors are the $\mathbf{v}_i$; the eigenvalues are the μ_i^2.

One can see from the $O_1 D O_2$ formulation that $|\det A| = \prod_i \mu_i$. Also A^* has the same singular values as A, while the singular values of A^{-1} are the reciprocals of the singular values of A. A simple calculation shows that $\|A\| = \mu_1$; the vector achieving the maximum in the norm is just $\mathbf{v}_1$.

A generalization of this is the following:

Proposition 10. *Let A be a $d \times d$ matrix. Let the singular values be $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_d$ and the singular vectors be $\mathbf{v}_1, \ldots, \mathbf{v}_d$. Then*

(1) $\mu_k = \max_{\dim V = k} \min_{v \in V \cap S} \|Av\|$ *(where S denotes the unit sphere); the maximum is attained by $V = \mathrm{lin}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.*
(2) $\mu_k = \min_{\dim V = d-k+1} \max_{v \in V \cap S} \|Av\|$*; the minimum is attained by $V = \mathrm{lin}\{\mathbf{v}_k, \ldots, \mathbf{v}_d\}$.*
(3) $\mu_1 \cdots \mu_k = \max_{\dim V = k} |\det(A|_V)|$*, where* $\det(A|_V)$ *is the k-dimensional volume growth of A.*

The proof (of the first two parts) is fairly straightforward. Since orthogonal matrices are isometries, it suffices to consider the case where A is diagonal. The last part can be proved using the Cauchy-Binet theorem (more on this below).

2.2. **Exterior algebra (user's guide – apologies to Bourbaki).** If V is a vector space, define an abstract vector space $\bigwedge^k V$ spanned by elements $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k$ satisfying relations:

- (Multi-linearity) $\mathbf{v}_1 \wedge \cdots \wedge (a\mathbf{v}_i) \wedge \cdots \wedge \mathbf{v}_k = a\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_i \wedge \cdots \wedge \mathbf{v}_k$;
- (Antisymmetry) $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_i \wedge \cdots \wedge \mathbf{v}_j \wedge \cdots \wedge \mathbf{v}_k = -\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_j \wedge \cdots \wedge \mathbf{v}_i \wedge \cdots \wedge \mathbf{v}_k$.

If $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ forms a basis for V, then $\{\mathbf{v}_{i_1} \wedge \cdots \wedge \mathbf{v}_{i_k} : i_1 < \ldots < i_k\}$ forms a basis for $\bigwedge^k V$. In particular, $\bigwedge^k V$ is $\binom{d}{k}$- dimensional.

If V has an inner product, it turns out that $\bigwedge^k V$ inherits a natural inner product structure. A crude way to do this is to take an orthonormal basis for V, $\mathbf{e}_1, \ldots, \mathbf{e}_d$ and declare $\{\mathbf{e}_{i_1} \wedge \cdots \mathbf{e}_{i_k}\}$ to be an orthonormal basis for $\bigwedge^k V$. This turns out to define an inner product that does not depend on the choice of orthonormal basis (the proof is due to a clever determinant equality, the Cauchy-Binet theorem).

If A is a linear map on V, then there is a natural action on $\bigwedge^k V$, $A^{\wedge k}(\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k) = (A\mathbf{v}_1) \wedge \cdots \wedge (A\mathbf{v}_k)$ (you can check that the relations above are preserved). Singular value decomposition plays extremely nicely with exterior algebras. This is the basis of the so-called

Raghunathan trick, which we will use below in our proof of the (non-invertible) multiplicative ergodic theorem.

Suppose that $\mathbf{v}_1, \dots \mathbf{v}_d$ are the singular vectors for A with singular values $\mu_1, \dots, \mu_d$. Also, let $A\mathbf{v}_i = \mu_i \mathbf{w}_i$, so that the $\mathbf{w}_i$ are orthonormal. Then $A^{\wedge k}$ maps $\mathbf{v}_{i_1} \wedge \dots \wedge \mathbf{v}_{i_k}$ to $\mu_{i_1} \cdots \mu_{i_k} \mathbf{w}_{i_1} \wedge \dots \wedge \mathbf{w}_{i_k}$.

> The singular vectors of $\bigwedge^k A$ are the k-fold wedge products of the singular vectors of A; the singular values of $\bigwedge^k A$ are the k-fold products of the singular values of A.

The top singular value of $\bigwedge^k A$ is $\mu_1 \cdots \mu_k$ – we can interpret $\| \bigwedge^k A\|$ as the rate of growth of k-dimensional volume (see Proposition 10 above).

We sometimes think of $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_k$ as representing a normal (with magnitude) to a subspace spanned by $\mathbf{v}_1, \dots, \mathbf{v}_k$. (but what does $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_k + \mathbf{w}_1 \wedge \dots \wedge \mathbf{w}_k$ represent?)

2.3. **Grassmannians.** The Grassmannian, $\mathcal{G}_k(V)$, of a vector space V is the collection of all k-dimensional subspaces of V. It has a lot of structure! compact manifold, homogeneous space, scheme, ... Unlike exterior algebras, there's no additive structure though...

Mostly we'll take $V = \mathbb{R}^d$ and think of $\mathcal{G}_k(V)$ as a compact (and hence complete) metric space. The distance between two k-dimensional subspaces W and W' is the Hausdorff distance $d_H(W \cap S, W' \cap S)$, where S is the unit sphere.

Recall that

$$d_H(W \cap S, W' \cap S) = \max\Big(\max_{\mathbf{v} \in W \cap S} d(\mathbf{v}, W' \cap S), \max_{\mathbf{v} \in W' \cap S} d(\mathbf{v}, W \cap S)\Big).$$

It turns out that there is a 'symmetry of closeness' for subspaces of the same dimension. Each of the terms in the maximum is bounded by a constant multiple of the other, where the constant depends only on k (even in Banach spaces). In particular, to show that two k-dimensional subspaces are close it suffices to show that either of the two quantities in the maximum is small.

The manifold $\mathcal{G}_k(\mathbb{R}^d)$ turns out to be $k(d-k)$-dimensional. I made use of a system of coordinates in a recent paper.

2.4. **Inverse cocycle.** We mentioned this one above. Here we require invertibility both of σ, and of the matrices defining the cocycle: $A_\omega \in GL(d, \mathbb{R})$. The inverse cocycle is a cocycle over σ^{-1} generated by $B_\omega = A^{-1}_{\sigma^{-1}\omega}$.

As mentioned above, $B^{(n)}_\omega = (A^{(n)}_{\sigma^{-n}\omega})^{-1}$. A little work with the sub-additive ergodic theorem shows that the exponents of the inverse cocycle are $-\mu_d \geq \ldots \geq -\mu_1$.

2.5. **Dual cocycle.** This cocycle requires invertibility of σ, but **not** invertibility of the matrices A_ω. Again it's a cocycle over σ^{-1}. The generator is $C_\omega = A^*_{\sigma^{-1}\omega}$.

We then calculate

$$\begin{aligned} C^{(n)}_\omega &= C_{\sigma^{-(n-1)}\omega} \cdots C_\omega \\ &= A^*_{\sigma^{-n}\omega} \cdots A^*_{\sigma^{-1}\omega} \\ &= (A_{\sigma^{-1}\omega} \cdots A_{\sigma^{-n}\omega})^* \\ &= (A^{(n)}_{\sigma^{-n}\omega})^*. \end{aligned}$$

Since the singular values of $A^{(n)}_{\sigma^{-n}\omega}$ are approximately $e^{n\mu_1}, \ldots, e^{n\mu_d}$, we see that the singular values of the dual are the same, and hence the dual cocycle has the same Lyapunov exponents as the primal cocycle.

2.6. **Temperedness.** If σ is a measure-preserving transformation and $f \in L^1$, then for $\mathbb{P}$-a.e. ω, $f(\sigma^n\omega)/n \to 0$.

The cheapest proof uses the Birkhoff ergodic theorem, but there is a more elementary proof using the first Borel-Cantelli lemma.

3. Multiplicative Ergodic Theorem: proof sketch

We'll give a rough sketch of the proof of the multiplicative ergodic theorem (the non-invertible version). The framework for the proof is from Raghunathan's proof of the MET (based on singular values). We'll deviate from it slightly, giving an inductive proof that we have modified to work in Banach spaces.

We've already indicated how to deduce the invertible version from the non-invertible version.

Let σ be an ergodic measure-preserving transformation of a probability space $(\Omega, \mathbb{P})$. Let $A\colon \Omega \to M_{d\times d}(\mathbb{R})$ generate a matrix cocycle and assume that A satisfies $\int \log \|A_\omega\| \, d\mathbb{P}(\omega) < \infty$.

3.1. **Finding the exponents.** For each k, define a new cocycle $A^{\wedge k}_\omega$. This satisfies $\int \log^+ \|A^{\wedge k}_\omega\| \, d\mathbb{P}(\omega) \leq k \int \log^+ \|A_\omega\| \, d\mathbb{P}(\omega) < \infty$. Now, $f^{\wedge k}_n(\omega) = \log \|{A^{\wedge k}}^{(n)}_\omega\|$ is a sub-additive sequence by sub-multiplicativity of norms, so that $\frac{1}{n} f^{\wedge k}_n(\omega)$ is convergent $\mathbb{P}$-a.e. to $\lim_{n\to\infty} \frac{1}{n} \int f^{\wedge k}_n \, d\mathbb{P}$ by the sub-additive ergodic theorem (or the Furstenberg-Kesten theorem).

We let this limit be $\mu_1 + \ldots + \mu_k$; that is, we define

$$\mu_k = \lim_{n\to\infty} \frac{1}{n} \int (f_n^{\wedge k}(\omega) - f_n^{\wedge (k-1)}(\omega))\, d\mathbb{P}(\omega),$$

where $f_n^{\wedge 0} \equiv 0$. From our comments before, $\mu_1 + \ldots + \mu_k$ measures the long-term k-dimensional volume growth rate of $A_\omega^{(n)}$. These μ_k's will be the Lyapunov exponents.

Recalling that $\|A^{\wedge k}{}_\omega^{(n)}\| = \prod_{j=1}^k s_j(A_\omega^{(n)})$, we see

$$e^{\mu_k} = \lim_{n\to\infty} s_k(A_\omega^{(n)})^{1/n}.$$

In particular, the μ_k are non-increasing as required.

Notice also, that

$$\begin{aligned}\mu_1 + \ldots + \mu_d &= \lim_{n\to\infty} \tfrac{1}{n} \log\left(s_1(A_\omega^{(n)}) \cdots s_d(A_\omega^{(n)})\right)\\ &= \lim_{n\to\infty} \tfrac{1}{n} \log|\det A_\omega^{(n)}|\\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log|\det A_{\sigma^i\omega}| = \int \log|\det A_\omega|\, d\mathbb{P}(\omega).\end{aligned}$$

In particular, if $A_\omega \in SL(d,\mathbb{R})$ for each ω, then $\mu_1 + \ldots + \mu_d = 0$.

We let $\lambda_1 > \ldots > \lambda_l$ be the Lyapunov exponents without repetition, where λ_j occurs m_j times among the μ_i's. m_j is called the *multiplicity* of λ_j.

3.2. **Finding the slow spaces.** Let $\eta < \frac{1}{6}\min_i(\lambda_i - \lambda_{i+1})$. By the sub-additive ergodic theorem, there exists for $\mathbb{P}$-a.e. ω an n_0 such that for each $n > n_0$ and each i, $e^{(\mu_i-\eta)n} < s_i(A_\omega^{(n)}) < e^{(\mu_i+\eta)n}$. Let $1 \le j \le l$ and let $U_j^{(n)}(\omega)$ be the $(m_j + m_{j+1} + \ldots + m_l)$-dimensional space spanned by the singular vectors with the smallest singular values (these are all $e^{(\mu_j+\eta)n}$ or smaller).

Claim 1. *For $\mathbb{P}$-a.e. ω, the $U_j^{(n)}(\omega)$ form a Cauchy sequence in $\mathcal{G}_k(\mathbb{R}^d)$.*

We define the limit of the $U_j^{(n)}(\omega)$ as $n \to \infty$ to be $U_j(\omega)$.

3.3. **Finishing the proof.**

Claim 2. *$U_j(\omega)$ is equivariant.*

Claim 3. *For $\mathbb{P}$-a.e. ω, if $\mathbf{v} \notin U_j(\omega)$, then $\frac{1}{n}\log\|A^{(n)}\mathbf{v}\| \ge \lambda_{j-1} - \eta$ for all large n.*

Claim 4. *For $\mathbb{P}$-a.e. ω, if $\mathbf{v} \in U_j(\omega)$, then $\frac{1}{n}\log\|A^{(n)}\mathbf{v}\| \le \lambda_j + \eta$ for all large n.*

3.4. **Proof sketches.**

Claim 1 proof sketch. We want to show that for each j and for $\mathbb{P}$-a.e. ω, $U_j^{(n)}(\omega)$ is a Cauchy sequence of subspaces of $\mathbb{R}^d$.

The idea is quite simple:

Take $\mathbf{v} \in U_j^{(n)}(\omega) \cap S$ and write $\mathbf{v} = \mathbf{u} + \mathbf{w}$ where $\mathbf{u} \in U_j^{(n+1)}(\omega)$ and $\mathbf{w} \in U_j^{(n+1)}(\omega)^{\perp}$. Notice that $U_j^{(n+1)}(\omega)^{\perp}$ is the fast space for $A_\omega^{(n+1)}$ spanned by the top singular vectors, which all expand at least at exponential rate $\lambda_{j-1} - \eta$.

We have

$$\|A_\omega^{(n+1)}\mathbf{v}\| \le e^{(\lambda_j+\eta)n}\|A_{\sigma^n\omega}\|; \text{ but also}$$
$$\|A_\omega^{(n+1)}\mathbf{v}\| \ge e^{(\lambda_{j-1}-\eta)(n+1)}\|\mathbf{w}\|.$$

By temperedness, $\|A_{\sigma^n\omega}\| \le e^{\eta n}$, so that we end up with an estimate like $\|w\| \lesssim e^{(\lambda_j-\lambda_{j-1}-3\eta)n}$. Any vector $\mathbf{v}$ in $U_j^{(n)}(\omega) \cap S$ is exponentially close to a vector, $\mathbf{u}$, in $U_j^{(n+1)}(\omega)$. The symmetry of closeness described above implies $d(U_j^{(n)}(\omega), U_j^{(n+1)}(\omega)) \lesssim e^{(\lambda_j-\lambda_{j-1}-3\eta)n}$. Since these distances are summable, the sequence of subspaces is Cauchy and

$$d(U_j^{(n)}(\omega), U_j(\omega)) \lesssim e^{(\lambda_j-\lambda_{j-1}-3\eta)n}. \tag{1}$$

□

Claim 2 is proved similarly: we take an element $\mathbf{v}$ of $U_j^{(n+1)}(\omega) \cap S$ and write $A_\omega \mathbf{v}$ as $\mathbf{u}+\mathbf{w}$ with $\mathbf{u}$ in $U_j^{(n)}(\sigma\omega)$ and $\mathbf{w} \in U_j^{(n)}(\sigma\omega)^{\perp}$. Since $A_\omega^{(n+1)}\mathbf{v}$ isn't too large, one deduces that $\mathbf{w}$ is exponentially small

Claim 3 is straightforward. If $\mathbf{v} \notin U_j(\omega)$, then it has some distance δ from $U_j(\omega)$. It then has distance at least $\frac{\delta}{2}$ from $U_j^{(n)}(\omega)$ for large n. Writing $\mathbf{v}$ as $\mathbf{u}+\mathbf{w}$ with $\mathbf{u} \in U_j^{(n)}(\omega)$ and $\mathbf{w} \in U_j^{(n)}(\omega)^{\perp}$, we have $\|w\| \ge \frac{\delta}{2}$, so that $\|A_\omega^{(n)}\mathbf{v}\| \ge \|A_\omega^{(n)}\mathbf{w}\| \gtrsim e^{(\lambda_{j-1}-\eta)n}$.

Claim 4 is tricky, but is straightforward for $j = 2$. If $\mathbf{v} \in U_2(\omega) \cap S$, then $\mathbf{v}$ can be written as $\mathbf{u}+\mathbf{w}$ with $\mathbf{u} \in U_2^{(n)}(\omega)$ and $\mathbf{w} \in U_2^{(n)}(\omega)^{\perp}$ with $\|\mathbf{w}\| \lesssim e^{-(\lambda_1-\lambda_2-3\eta)n}$ by (1). Now applying $A_\omega^{(n)}$ to $\mathbf{u}$ and $\mathbf{w}$ separately, both are of size $\lesssim e^{\lambda_2 n}$.

For $j > 2$, Raghunathan splits $\mathbf{w}$ into pieces with different growth rates and proves stronger estimates on the faster growing pieces. A technically simpler (but less pleasing proof) is obtained by working inductively down from the top, giving a priori estimates on the growth of the 'fast part'.

All of the above can be made to work in separable Banach spaces. First, a sub-multiplicative notion of k-dimensional volume growth is needed. From this, we can define the exponents as above. The spaces are obtained in essentially the same way.

This gives

Theorem 11 (Multiplicative ergodic theorem for Banach spaces; non-invertible version)**.** *Let σ be an ergodic measure-preserving transformation of a probability space $(\Omega, \mathbb{P})$. Let X be a Banach space with separable dual and let (A_ω) be a cocycle of quasi-compact operators on X. Then there exist $\lambda_1 > \lambda_2 > \ldots$, multiplicities $m_1, m_2, \ldots$ and $(m_1 + \ldots + m_{i-1})$-codimensional spaces $U_i(\omega)$ that are equivariant and have the correct growth conditions.*

We come back to the deduction of invertible-type METs from non-invertible METs. Taking a cue from Corollary 9, we insist on invertibility of the base transformation. We see what we can get without making any invertibility assumptions on the matrices.

3.5. **Non-invertible to invertible II: Duality.** Assume that X is reflexive: $X^{**} = X$ (this is certainly true in the finite-dimensional case). Then X^* is separable and the dual cocycle C_ω described above satisfies the non-invertible MET with the same exponents.

For matrices, a left eigenvector and right eigenvector with different eigenvalues are orthogonal. Analogously, a dual Oseledets subspace annihilates a primal Oseledets subspace with a different exponent. One can then show $U^*_{j+1}(\omega)^\circ$ is the fast space, $V_{\geq\lambda_j}(\omega)$. From here, you intersect with $V_{\leq\lambda_j}(\omega)$ to get $V_j(\omega)$ as previously.

3.6. **Non-invertible to invertible III: Power method.** To find the largest eigenvector of a matrix (assuming that the largest eigenvalue has multiplicity 1), you can just take (almost) any vector and iterate the matrix.

There's an analogous method to build the $V_j(\omega)$. I'll describe it for the case of $\mathbb{R}^d$, but there is a Banach space version also.

$$V_j(\omega) = \lim_{n\to\infty} A^{(n)}_{\sigma^{-n}\omega}(U_j(\omega) \ominus U_{j+1}(\omega)).$$

Start from the orthogonal complement of the $V_{\leq\lambda_{j+1}}(\sigma^{-n}\omega)$ inside $V_{\leq\lambda_j}(\sigma^{-n}\omega)$. This has some 'component' of the fast space inside it (you can make sense of this using exterior algebra). When it is iterated, the fast part dominates and the sequence converges to the true fast space.

The following version of the MET (that we call semi-invertible) follows from either of these approaches.

Theorem 12 (Multiplicative Ergodic Theorem (semi-invertible) - Froyland, Lloyd, Q)**.** *Let σ be an **invertible** ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$. Let $A\colon \Omega \to GL_d(\mathbb{R})$ be a cocycle of **not necessarily invertible** matrices with $\int \log \|A_\omega\| \, d\mathbb{P}(\omega) < \infty$.*

Then there exist $\infty > \lambda_1 > \lambda_2 > \ldots > \lambda_k > -\infty$ and subspaces $V_1(\omega), \ldots, V_k(\omega)$ satisfying

1. ***Decomposition**: $\mathbb{R}^d = V_1(\omega) \oplus \ldots \oplus V_k(\omega)$;*
2. ***Equivariance**: $A_\omega(V_i(\omega)) = V_i(\sigma(\omega))$;*
3. ***Growth**: If $v \in V_i(\omega) \setminus \{0\}$, then $\frac{1}{n} \log \|A_\omega^{(n)} v\| \to \lambda_i$ as $n \to +\infty$.*

There is also a version for quasi-compact operators on separable Banach spaces

4. Last lecture

To be decided: either METs on Banach spaces and applications to oceans (lighter weight); or sensitivity of Lyapunov exponents.