Multiplicative Ergodic Theorems and Applications

Cecilia González-Tokman and Anthony Quas
Department of Mathematics and Statistics, University of Victoria

Ergodic theory is the study of measure-preserving transformations. Its utility derives from two facts: (1) there are remarkably strong conclusions that apply to every measure-preserving transformation; and (2) measure-preserving transformations occur naturally in many branches of mathematics (for example number theory, combinatorics, probability, differential geometry, information theory). Formally, one considers maps σ from a space Ω (equipped with a measure μ and a σ-algebra $\mathcal{B}$) to itself, for which $\mu(\sigma^{-1}A) = \mu(A)$for all measurable sets A. We generally consider the case where μ is a probability measure. A simple example is rigid rotation of a circle, which preserves Lebesgue measure. The measure-preserving transformation $(\Omega, \mathcal{B}, \sigma, \mu)$, or more colloquially the measure μ, is said to be *ergodic* if $\sigma^{-1}(A) = A$ implies $\mu(A) = 0$ or $\mu(A^c) = 0$. This can be seen as an indecomposability condition, as if μ were not ergodic, one could restrict σ to A and A^c and study the systems separately.

A prototype theorem is the Birkhoff ergodic theorem.

Theorem 1 (Birkhoff, 1932 [1]). *Let σ be a measure-preserving transformation of a measure space Ω and let f be an integrable function f on Ω. Then $(1/n)(f(\omega) + \ldots + f(\sigma^{n-1}\omega))$ is convergent to some limit (a priori depending on ω) for almost every ω. If μ is ergodic, then the limit is $\int f\,d\mu$ for μ-almost every ω.*

The ergodic case is often reduced to the maxim "*the time average and space average agree.*" Notice that the theorem gives non-trivial conclusions even in the simplest example mentioned above. The strong law of large numbers in probability theory is an easy corollary of the Birkhoff ergodic theorem, as is the Borel normal numbers theorem.

Chaos has been defined in many ways, but all of the notions of chaotic behaviour of a map include infinitesimally close points becoming in some sense separated as the map is iterated. One natural approach to this, in the case where the phase space is a subset of $\mathbb{R}^d$ and the map is differentiable, is to look at the derivative of the nth power of the map. Using the chain rule, the derivative of the nth composition of T, $DT^n(\omega)$, is $DT(T^{n-1}\omega) \cdot DT(T^{n-1}\omega) \cdots DT(\omega)$. We focus temporarily on the case where T is a map of a one-dimensional space. In this case, $DT^n(\omega)$ is simply a product of reals. In this case, taking logarithms, one sees that $(1/n)log|DT^n(\omega)| = (1/n)(log|DT(\omega)| + \ldots + log|DT(T^{n-1}\omega)|)$. Providing T preserves an ergodic invariant measure μ, the limit of this is shown by the ergodic theorem to be $\int log|DT(\omega)|d\mu(\omega)$. This quantity is called the *Lyapunov exponent* of T. Large Lyapunov exponents indicate fast separation of infinitesimally close trajectories.

One definition of chaos is the existence of a positive Lyapunov exponent. In the higher-dimensional case, where T is a self-map of a manifold or space with differentiable structure, DT^n is a matrix product, so that the ergodic theorem does not directly apply as above. Oseledets' multiplicative ergodic theorem, another extremely general statement, is of interest in this setting.

Theorem 2 (Oseledets, 1968 [7]) *Let σ be a measure-preserving transformation of a space Ω and let $A\colon \Omega \to GL(d, \mathbb{R})$ be measurable. Assume that $\int \log A^{\pm 1}(\omega)\,d\mu(\omega) < \infty$. Then there exist $\lambda_1(\omega) > \lambda_2(\omega) > \ldots > \lambda_k(\omega)$ and a decomposition $\mathbb{R}^d = V_1(\omega) \oplus V_2(\omega) \oplus \ldots \oplus V_k(\omega)$ such that for μ-almost every ω, the following hold:*

- $\lambda_i(\sigma\omega) = \lambda_i(\omega)$;
- $A(\omega)V_i(\omega) = V_i(\sigma(\omega))$; and
- $\lim_{n\to\infty} (1/n) \log \|A(\sigma^{n-1}\omega) \cdots A(\omega)v\| = \lambda_i(\omega)$ for all $v \in V_i(\omega) \setminus \{0\}$.

In particular, if μ is ergodic, then the $\lambda_i(\omega)$ are constant–called Lyapunov exponents–and the $V_i(\omega)$ have constant dimension μ–almost everywhere.

We see this as a generalization to random matrices of a Jordan block decomposition. The vector space is expressed as a direct sum decomposition (with the decomposition depending equivariantly on ω) into pieces that have characteristic exponential rates of expansion given by the $\lambda_i(\omega)$. Applying this in the case where $A(\omega)$ is $DT(\omega)$ mentioned above, one obtains a splitting of the tangent space over each point into vector subspaces, each of which has a characteristic infinitesimal rate of expansion.

One caveat with the above theorems is that the strength of the conclusion is only as good as the invariant measure that one is using. Suppose that T is a map from X to itself (where there is no measure specified yet). To apply ergodic theory, one needs to find a T-invariant measure. In general there may be (uncountably) many invariant measures for a given transformation. If T has a fixed point, x_0, then the Dirac δ-measure, giving mass 1 to x_0 and mass 0 to the remainder of the space is an invariant measure. Unsurprisingly, this measure is quite useless from our point of view, because the ergodic theorems give conclusions that hold for almost every point *with respect to the invariant measure*. In the case of a δ-measure, the conclusion only holds at the single fixed point! If one wants a conclusion that holds for Lebesgue almost every point (assuming X is a subset of $\mathbb{R}^d$ or a manifold), one needs an ergodic invariant measure equivalent to Lebesgue measure (that is, an *absolutely continuous invariant measure*).

A standard technique for searching for absolutely continuous invariant measures is to use the *Ruelle-Perron-Frobenius operator*. For a dynamical system T whose phase space is a subset of $\mathbb{R}^d$, its Perron-Frobenius operator is a linear map

$\mathcal{L}$ from $L^1(\mathbb{R}^d)$ to itself, which may be thought of as describing the evolution of densities: if a random variable Z has a distribution with density $\rho(x)$, then $T(Z)$ (the random variable obtained by sampling a point from Z and applying T) has density $\mathcal{L}(\rho)$. As an example, if T is the map of $\mathbb{R}$ sending x to $x/2$, $\mathcal{L}$ would send a density ρ to a density $\mathcal{L}(\rho)(x) = 2\rho(2x)$ (i.e. a density with half the width and twice the height). Absolutely continuous invariant measures are in a one-to-one correspondence with non-negative fixed points of this operator.

A good analogy for Ruelle-Perron-Frobenius operators is transition matrices for finite Markov chains. Invariant distributions for Markov chains are exactly fixed points (i.e. left eigenvectors with eigenvalue 1) for the transition matrix. In the nice situation where all other eigenvalues of the transition matrix have absolute value less than 1, there is said to be a *spectral gap*. The size of the spectral gap is related to the convergence rate of the Markov chain to its invariant distribution. The *mixing time* is a closely related statistic (see [6] for more information). There is important information also in the left eigenvectors corresponding to (real) eigenvalues close to 1. The level sets can be used to determine *almost invariant* sets for the Markov chain: if one has a Markov chain with a 'bottleneck' – that is, the state space is divided into pieces such that it takes a long time to transition between them – then eigenvectors are almost constant on the almost invariant sets.

We combine all of the above ingredients to understand the evolution of densities in forced dynamical systems. Here one imagines that there is a dynamical system (the *forcing system*) running autonomously in the background, but whose state affects the evolution of a second dynamical system. For a simple illustration, imagine the moon evolving autonomously and affecting, but not being affected by, tidal motions.

Formally, we are talking about dynamical systems of the form $S\colon \Omega \times X \to \Omega \times X$ given by $S(\omega, x) = (\sigma(\omega), T_\omega(x))$. We further assume that X is a subset of $\mathbb{R}^d$ and T_ω is a differentiable mapping. Writing $\mathcal{L}_\omega$ for the Ruelle-Perron-Frobenius operator of T_ω, we can study the evolution of densities under the forced dynamical system. A recent generalization of the multiplicative ergodic theorem due to Lian and Lu [5] (following earlier work by Mañé, Ruelle, Thieullen and others) allows us to decompose suitable Banach spaces of densities into a direct sum of finite-dimensional subspaces, the top one with exponent 0, and others with negative exponents, together with a residual faster-contracting subspace. Just as with the Markov chain analysis, this may be used to to identify principal bottlenecks responsible for inhibiting mixing. In recent work and work in progress with Gary Froyland [2,3,4] as part of a long-term program of Froyland and his collaborators, we aim to apply these techniques to studying slow-mixing regions of the oceans (called *gyres*) which have been implicated in phenomena such as the *Great Pacific Garbage Patch*.

References

[1] G.D. Birkhoff. Proof of the ergodic theorem. *Proc. Nat. Acad. Sci.*, 17:656–660, 1931.

[2] G. Froyland, C. González-Tokman, and A. Quas. Stability and approximation of random invariant densities for Lasota-Yorke map cocycles. submitted, 2012.

[3] G. Froyland, N. Santitissadeekorn, and A. Monahan. Transport in time-dependent dynamical systems: Finite-time coherent sets. *Chaos*, 20:043116, 2010

[4] C. González-Tokman and A. Quas. A semi-invertible operator Oseledets theorem. *Ergodic Theory Dynamical Systems* (to appear).

[5] Z. Lian and K. Lu. Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space. *Mem. Amer. Math. Soc.*, 206, 2010.

[6] D.A. Levin, Y. Peres, and E.L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence R.I., 2008.

[7] V. Oseledets. A multiplicative ergodic theorem. Characteristic Ljapunov exponents of dynamical systems. Trudy Moskov. Mat. Obšč, 19: 179–210, 1968.