Lyapunov Exponents and Osoledec's Multiplicative Ergodic Theorem

Jim Kelliher
Spring 2003 Working Dynamical Systems Seminar
UT Austin

The last substantial changes to this document were made in **September 2003**. This document was last compiled on **January 12, 2005**, which is an upper bound for when it was last updated. The most recent version is at:

`http://www.ma.utexas.edu/users/kelliher/Geometry/Geometry.html`

Motivation

Setting: A smooth, compact, Riemannian manifold (M, g) of dimension d.

$\varphi : \mathbb{R} \times M \to M$ is a smooth measure-preserving flow:

- Write $\varphi(t, x)$ or $\varphi_t(x)$,
- $\varphi_0 =$ identity,
- $\varphi(s, \varphi(t, x)) = \varphi(s + t, x)$ or $\varphi_s \circ \varphi_t = \varphi_{s+t}$
- $\det d\varphi_t = 1$,
- $d\varphi_t : T_x M \to T_{\varphi_t(x)} M$

Extend $d\varphi_t$ to be a map from TM to itself, where TM is the tangent bundle of M as follows (think of the tangent bundle as the phase space for a particle in a Hamiltonian system):

$$d\varphi_t(x, v) = (\varphi_t(x), (d\varphi_t)_x(v)).$$

This can be a little confusing because of the double use of $d\varphi_t$.

Example: Define φ on $M = \mathbb{R}^d$ such that [note that in this example M is not compact]

$$\varphi_1(x) = Ax,$$

where A is a constant self-adjoint, positive-definite matrix with $\det A = 1$. Then for all $x \in M$,

$$(d\varphi_1)_x = A,$$

$$(d\varphi_n)x = A^n,$$

$$(d\varphi_1)(x, v) = (Ax, Av),$$

$$(d\varphi_n)(x, v) = (A^n x, A^n v).$$

Notice that

$$(d\varphi_{s+t})_x = (d\varphi_s)_{\varphi_t(x)}(d\varphi_t)_x,$$

$$\begin{aligned}
d\varphi_{s+t}(x,v) &= (\varphi_{s+t}(x), (d\varphi_{s+t})_x v) \\
&= (\varphi_s(\varphi_t(x)), (d\varphi_s)_{\varphi_t(x)}(d\varphi_t)_x v) \\
&= d\varphi_s(\varphi_t(x), (d\varphi_t)_x v) \\
&= d\varphi_s(d\varphi_t(x,v)),
\end{aligned}$$

so

$$d\varphi_{s+t} = d\varphi_s \circ d\varphi_t.$$

If we fix a time increment of 1 then, being a dynamical systems talk, it is natural to consider the forward orbit of a point $(x,v) \in TM$ under iteration of $d\varphi_1$. The function φ_1 moves the point around on the manifold, while its differential moves the vector in the corresponding tangent spaces.

Let

$$\begin{aligned}
&r : TM \to \mathbb{R}, \\
&r(x,v) = \frac{\|(d\varphi_1)_x v\|_{\varphi_1(x)}}{\|v\|_x},
\end{aligned}$$

where we have emphasized that the norms are evaluated at different points in the Riemannian manifold, and so the Riemannian metric is involved here. We will suppress the subscripts on the norms from now on, though.

$r(x,v)$ is the factor by which the length of v expands or shrinks under the differential map of φ. (The lengths are measured at different points on the manifold.) Evaluating r at the k-th point in the forward orbit of (x,v) gives,

$$\begin{aligned}
r((d\varphi_1)^k(x,v)) &= r(d\varphi_k(x,v)) = r(\varphi_k(x), (d\varphi_k)_x v) \\
&= \frac{\|(d\varphi_1)_{\varphi_k(x)}(d\varphi_k)_x v\|}{\|(d\varphi_k)_x v\|} = \frac{\|(d\varphi_{k+1})_x v\|}{\|(d\varphi_k)_x v\|}.
\end{aligned}$$

Thus, the *geometric* mean of r as we move through the first n points in the forward orbit is:

$$\begin{aligned}
r_n(x,v) :&= \left\{ r(x,v) r(d\varphi_1(x,v)) r((d\varphi_1)^2(x,v)) \cdots r((d\varphi_1)^{n-1}(x,v)) \right\}^{1/n} \\
&= \left\{ \frac{\|(d\varphi_1)_x v\|}{\|v\|} \frac{\|(d\varphi_2)_x v\|}{\|(d\varphi_1)_x v\|} \cdots \frac{\|(d\varphi_n)_x v\|}{\|(d\varphi_{n-1})_x v\|} \right\}^{1/n} \\
&= \left\{ \frac{\|(d\varphi_n)_x v\|}{\|v\|} \right\}^{1/n}.
\end{aligned}$$

Back to our example:

$$r(x,v)=\frac{\|Av\|}{\|v\|},\quad r_n(x,v)=\left\{\frac{\|A^n v\|}{\|v\|}\right\}^{1/n}.$$

We can do much better than this, however. Suppose (v,λ) is an eigenvector, eigenvalue of A. Then $Av=\lambda v$, so

$$r(x,v)=\frac{\|\lambda v\|}{\|v\|}=|\lambda|\,,\quad r_n(x,v)=\left\{\frac{\|\lambda^n v\|}{\|v\|}\right\}^{1/n}=|\lambda|\,.$$

This says nothing more than that an eigenvector is stretched by a factor equal to the modulus of its eigenvalue. But what about an arbitrary vector v?

Let $v_1,\ldots,v_d$ be a complete orthonormal set of eigenvectors of A with corresponding eigenvalues $\lambda_1,\ldots,\lambda_d$, and assume that $0\le\lambda_1\le\cdots\le\lambda_d$. Write $v\in M$ as

$$v=\alpha_1v_1+\cdots+\alpha_dv_d.$$

Then

$$\begin{aligned}r(x,v)&=\frac{\|A(\alpha_1v_1+\cdots+\alpha_dv_d)\|}{\|\alpha_1v_1+\cdots+\alpha_dv_d\|}=\frac{\|\lambda_1\alpha_1v_1+\cdots+\lambda_d\alpha_dv_d\|}{\|\alpha_1v_1+\cdots+\alpha_dv_d\|}\\&=\left\{\frac{\lambda_1^2\alpha_1^2+\cdots+\lambda_d^2\alpha_d^2}{\alpha_1^2+\cdots+\alpha_d^2}\right\}^{1/2}\end{aligned}$$

and there's not much more to say about that. But,

$$\begin{aligned}r_n(x,v)&=\left\{\frac{\|A^n(\alpha_1v_1+\cdots+\alpha_dv_d)\|}{\|\alpha_1v_1+\cdots+\alpha_dv_d\|}\right\}^{1/n}\\&=\left\{\frac{\|\lambda_1^n\alpha_1v_1+\cdots+\lambda_d^n\alpha_dv_d\|}{\|\alpha_1v_1+\cdots+\alpha_dv_d\|}\right\}^{1/n}\\&=\left\{\frac{\lambda_1^{2n}\alpha_1^2+\cdots+\lambda_d^{2n}\alpha_d^2}{\alpha_1^2+\cdots+\alpha_d^2}\right\}^{1/2n}.\end{aligned}$$

Suppose k is the largest index such that $\alpha_k\neq0$. Then

$$\begin{aligned}r_n(x,v)&=\left\{\frac{\lambda_1^{2n}\alpha_1^2+\cdots+\lambda_k^{2n}\alpha_k^2}{\alpha_1^2+\cdots+\alpha_k^2}\right\}^{1/2n}\\&=\left\{(\lambda_k)^{2n}\frac{\left(\frac{\lambda_1}{\lambda_k}\right)^{2n}\alpha_1^2+\cdots+\left(\frac{\lambda_{k-1}}{\lambda_k}\right)^{2n}\alpha_{k-1}^2+\alpha_k^2}{\alpha_1^2+\cdots+\alpha_k^2}\right\}^{1/2n}\end{aligned}$$

$$= \lambda_k \left\{ \frac{\left(\frac{\lambda_1}{\lambda_k}\right)^{2n} \alpha_1^2 + \cdots + \left(\frac{\lambda_{k-1}}{\lambda_k}\right)^{2n} \alpha_{k-1}^2 + \alpha_k^2}{\alpha_1^2 + \cdots + \alpha_k^2} \right\}^{1/2n}$$

$$\to \lambda_k \text{ as } n \to \infty.$$

We can interpret the last equation above as follows: Project a vector into each eigenspace of A. The largest eigenvalue of the eigenspaces with nonzero projections is the asymptotic value of the geometric mean rate of expansion of the vector under the differential map—in our simple example. Almost all (in terms of Lebesgue measure on $\mathbb{R}^n \cong T_xM$) vectors will have $r_n(x, v) \to \lambda_d$, the largest eigenvalue of A.

An obvious question at this point is whether for an arbitrary flow an asymptotic value for $r_n(x, v)$ exists.The perhaps surprising answer is that it does, at least for almost all x in the manifold. This result follows from Oseledec's multiplicative ergodic theorem, which says that things work much the same as they do in our simple example.

But before continuing on, let us return to our function r_n and see what more we can say. Define

$$\lambda(x, v) = \lim_{n\to\infty} \log r_n(x, v),$$

when the limit exists. We take the logarithm here for historical reasons: when $\lambda(x, v)$ exists it is called the *Lyapunov* or *characteristic exponent* of φ for (x, v). We can calculate,

$$\begin{aligned}
\lambda(x, v) &= \lim_{n\to\infty} \log \left\{ \frac{\|(d\varphi_n)_x v\|}{\|v\|} \right\}^{1/n} \\
&= \lim_{n\to\infty} \frac{1}{n} \log \|(d\varphi_n)_x v\| - \lim_{n\to\infty} \frac{1}{n} \log \|v\| \\
&= \lim_{n\to\infty} \frac{1}{n} \log \|(d\varphi_n)_x v\|,
\end{aligned}$$

which because the differential is linear depends only on the direction of v, not on its length. (We could have seen this directly by noting that the same applies to r.)

The Lyapunov exponent is able to determine the rate of exponential growth or decay of $\|(d\varphi_n)_x v\|$ along an orbit, but is not able to detect the presence of sub-exponential growth or decay. If $\lambda(x, v) < 0$, then the orbits of x and of a nearby point y will converge toward each other; if $\lambda(x, v) > 0$, then the orbits of x and y will diverge (at least initially); if $\lambda(x, v) = 0$ then we can conclude nothing.

Also, if we let $\tau = d\varphi_1$, then

$$\begin{aligned}\lambda(x,v) &= \lim_{n\to\infty} \log \left\{ r(x,v) r(\tau(x,v)) r(\tau^2(x,v)) \cdots r(\tau^{n-1}(x,v)) \right\}^{1/n} \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log r(\tau^{k-1}(x,v)).\end{aligned}$$

Put a measure on the tangent bundle, TM, that is equal, in the local trivializations, to $dV dm$, where dV is the measure inherited from the Riemannian manifold and dm is Lebesgue measure on $\mathbb{R}^d$ (the tangent spaces). Then $\tau = d\varphi_1$ is measure-preserving since φ_1 is measure-preserving and $\det(d\varphi_1)_x = 1$ for all $x \in M$ so $(d\varphi_1)_x$ preserves measure in $\mathbb{R}^d$ (the volume of $(d\varphi_1)_x$ applied to a cube is equal to the determinant of $(d\varphi_1)_x$ times the volume of the cube).

So we have the exact setup we need to apply Birkhoff's ergodic theorem except for one critical item—$\log r(x,v)$, the function we are averaging, is not in $L^1(TM)$ (it does not have a finite integral). We might argue that since $r(x,v)$ depends only upon the direction of v, not on its magnitude, we could get away with using the sphere bundle SM rather than the tangent bundle (that is, use only the unit sphere in each tangent space rather than the entire tangent space), since $r(x,v)$ *is* in $L^1(SM)$. The problem with this, is that $d\varphi_1$ is not measure-preserving on SM.

This approach, so hopeful looking, has met an insurmountable obstacle.

What we have been looking at is how the length of an initial vector grows or shrinks as we iterate the differential map of φ_1. We did this by following the orbit of a point (x,v) in the tangent bundle and evaluating a real-valued function, r, at each point. But the natural measure on the tangent bundle that makes the differential map invariant is an infinite measure. The way around our difficulty is, paradoxically, to ask for more information about the behavior of the initial vector under iteration.

Let us ask what happens to the direction of the vector as well. One approach would be to examine what happens to an appropriately defined vector-valued function of points on the orbit in the tangent bundle. But this would suffer from the same problem as before: the unboundedness of the natural measure on the tangent bundle.

Instead, we look at the behavior of the differential map, viewed as a $d \times d$ real matrix, as it varies over the orbit of the initial point in the manifold itself. (Actually, we look at another $d \times d$ matrix formed from the differential map, as we shall see.) We never introduce an initial vector; if we can get enough information about the long term behavior of the differential map, then we can answer any question about an initial vector using that information.

Since the manifold is compact and φ_1 preserves its natural measure, our problem with applying the ergodic theorem disappears. The tradeoff is that the ergodic theorem itself no longer even applies. But a natural, if difficult, generalization of it, Oseledec's multiplicative ergodic theorem, does.

Before moving on, let us summarize what has just been said. We wanted to apply Birkhoff's ergodic theorem to compute the average of a real-valued function over a measure space that is not finite. This not being possible, we reduce the domain to a finite measure space, but complicate the codomain by expanding it to matrix-valued functions.

One final comment is that because this new approach relies on the subadditive ergodic theorem, a variant of Birkhoff's ergodic theorem, we establish existence of Lyapunov exponents, but no effective procedure for calculating them.

Oseledec's Multiplicative Ergodic Theorem

The existence of Lyapunov exponents is assured, almost everywhere, by the multiplicative ergodic theorem. As it relates to the Lyapunov exponents of the previous section, the function T in the statement of this theorem corresponds to the mapping that assigns to each point in the manifold the differential, $d\varphi_1$, viewed as a real $d \times d$ matrix. The map τ is φ_1, a map from the manifold to the manifold (whereas in the previous section τ was that same map extended to be a map from the tangent bundle to the tangent bundle). The theorem is more general, though, than these specific interpretations would limit it to. For one thing, the matrices are $m \times m$, where m needn't be the dimension of the manifold.

Theorem 0.1 (Discrete-time Multiplicative Ergodic Theorem)**.** *Let T be a measurable function from M to the space of all real $m \times m$ matrices, such that*

$$\log^+ \|T(\cdot)\| \in L^1(M, \rho).$$

Let $\tau : M \to M$ be a measure-preserving map and let

$$T_x^n = T_{\tau^{n-1}(x)} \cdots T_{\tau(x)} T_x.$$

Then there is a $\Gamma \subseteq M$ with $\rho(\Gamma) = 1$ and such that $\tau\Gamma \subseteq \Gamma$, and the following holds for all $x \in \Gamma$:

(1) $\Lambda_x := \lim_{n\to\infty}((T_x^n)^* T_x^n)^{1/2n}$ *exists.*

(2) *Let* $\exp \lambda_x^{(1)} < \cdots < \exp \lambda_x^{(s)}$ *be the eigenvalues of Λ_x, where $s = s(x)$, the $\lambda_x^{(r)}$ are real, and $\lambda_x^{(1)}$ can be $-\infty$, and $U_x^{(1)}, \ldots, U_x^{(s)}$ the corresponding eigenspaces. Let $m_x^{(r)} = \dim U_x^{(r)}$. The functions $x \mapsto \lambda_x^{(r)}$ and $x \mapsto m_x^{(r)}$ are τ-invariant. Let $V_x^{(0)} = \{0\}$ and $V_x^{(r)} = U_x^{(1)} \oplus \cdots \oplus U_x^{(r)}$ for $r = 1, \ldots, s$. Then for $u \in V_x^{(r)} \setminus V_x^{(r-1)}$, $1 \le r \le s$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \|T_x^n u\| = \lambda_x^{(r)}.$$

These are the Lyapunov exponents.

By τ-invariant, we mean that a function f on M satisfies $f(\tau(x)) = f(x)$. Applying this relation repeatedly, it follows that the function is constant on the forward orbit of the point x under τ. If τ is invertible, then the function is the same on the entire orbit, forward and backward.

The norm on matrices we use in this theorem and throughout is the operator norm in Euclidean space, which is identical in value to the spectral norm (the highest-modulus of the eigenvalues of the matrix). A critical property of this norm is that it is a true matrix norm—that is, it is submultiplicative ($\|AB\| \le \|A\| \, \|B\|$).

For any $d \times d$ matrix A, the matrix A^*A is self-adjoint and positive-definite. The matrix $(A^*A)^{1/k}$ exists for all positive integers k. Also, $\left\|(A^*A)^{1/2}\right\| = \|A\|$. Thus, $((T_x^n)^* T_x^n)^{1/2n}$ is guaranteed to exist, and we would expect it to be about the same size (loosely speaking) as a typical value of T_x.

A few words on the proof of this theorem:

- The subadditive ergodic theorem is all that is needed to establish the existence of limiting values for the eigenvalues of $((T_x^n)^* T_x^n)^{1/2n}$. This is the "easy" part of the proof.

- Establishing the existence of a the limiting value, Λ_x, for the matrices $((T_x^n)^* T_x^n)^{1/2n}$ themselves is much harder. It involves proving the convergence in Grassman manifolds of subspaces constructed from the eigenvectors of $((T_x^n)^* T_x^n)^{1/2n}$.

- The existence of the limiting matrix Λ_x is required to prove that $\lim_{n\to\infty} \frac{1}{n} \log \|T_x^n u\| = \lambda_x^{(r)}$, thereby establishing the existence of the Lyapunov exponents. Thus, even though it is easy to show the existence of the $\lambda_x^{(r)}$, the hard part of the proof must be faced (it seems) to establish their connection to the Lyapunov exponents.

One final comment on the proof of the multiplicative ergodic theorem. The subadditive theorem only establishes the existence of a limit, unlike Birkhoff's ergodic theorem, which gives a formula for calculating the limit. Thus, we obtain the existence of the Lyapunov exponents, but no effective method for calculating them. Ultimately, this limitation is due to the fact that the matrix norm is submultiplicative—$\|AB\| \leq \|A\| \, \|B\|$—rather than multiplicative.