Chapter 7

Lyapunov Exponents

Lyapunov exponents tell us the rate of divergence of nearby trajectories—a key component of chaotic dynamics. For one dimensional maps the exponent is simply the average $< \log |df/dx| >$ over the dynamics (chapter 4). In this chapter the concept is generalized to higher dimensional maps and flows. There are now a number of exponents equal to the dimension of the phase space $\lambda_1, \lambda_2 \ldots$ where we choose to order them in decreasing value. The exponents can be intuitively understood geometrically: line lengths separating trajectories grow as $e^{\lambda_1 t}$ (where t is the continuous time in flows and the iteration index for maps); areas grow as $e^{(\lambda_1+\lambda_2)t}$; volumes as $e^{(\lambda_1+\lambda_2+\lambda_3)t}$ etc. However, areas and volumes will become strongly distorted over long times, since the dimension corresponding to λ_1 grows more rapidly than that corresponding to λ_2 etc., and so this is not immediately a practical way to calculate the exponents.

7.1 Maps

Consider the map

$$U_{n+1} = F(U_n). \tag{7.1}$$

with U the phase space vector. We want to know what happens to a small change in U_0. This is given by the iteration of the "tangent space" given by the Jacobean matrix

$$K_{ij}(U_n) = \left.\frac{\partial F_i}{\partial U^{(j)}}\right|_{U=U_n}. \tag{7.2}$$

Then if the change in U_n is ε_n

$$\varepsilon_{n+1} = \mathbf{K}(U_n)\varepsilon_n, \tag{7.3}$$

or

$$\frac{\partial U_n^{(i)}}{\partial U_0^{(j)}} = M_{ij}^n = \left[\mathbf{K}(U_{n-1})\mathbf{K}(U_{n-2})\ldots\mathbf{K}(U_0)\right]_{ij}. \tag{7.4}$$

7.2 Flows

For continuous time systems

$$\frac{dU}{dt} = f(U) \tag{7.5}$$

a change $\varepsilon(t)$ in $U(t)$ evolves as

$$\frac{d\varepsilon}{dt} = \mathbf{K}(U)\varepsilon \quad \text{with} \quad K^{(ij)} = \left.\frac{\partial f_i}{\partial U^{(j)}}\right|_{U=U(t)}. \tag{7.6}$$

Then

$$\frac{\partial U^{(i)}(t)}{\partial U^{(j)}(t_0)} = M_{ij}(t, t_0) \tag{7.7}$$

with $\mathbf{M}$ satisfying

$$\frac{d\mathbf{M}}{dt} = \mathbf{K}(U(t))\mathbf{M}. \tag{7.8}$$

7.3 Oseledec's Multiplicative Ergodic Theorem

Roughly, the eigenvalues of $\mathbf{M}$ for large t are $e^{\lambda_i n}$ or $e^{\lambda_i(t-t_0)}$ for maps and flows respectively. The existence of the appropriate limits is known as Oseledec's multiplicative ergodic theorem [1]. The result is stated here in the language of flows, but the version for maps should then be obvious.

For almost any initial point $U(t_0)$ there exists an orthonormal set of vectors $v_i(t_0)$, $1 \le i \le n$ with n the dimension of the phase space such that

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t - t_0} \log \|\mathbf{M}(t, t_0) v_i(t_0)\| \tag{7.9}$$

exists. For ergodic systems the $\{\lambda_i\}$ do not depend on the initial point, and so are global properties of the dynamical system. The λ_i may be calculated as the log of the eigenvalues of

$$\left[\mathbf{M}^T(t, t_0)\mathbf{M}(t, t_0)\right]^{\frac{1}{2(t-t_0)}}. \tag{7.10}$$

with T the transpose. The $v(t_0)$ are the eigenvectors of $\mathbf{M}^T(t, t_0)\mathbf{M}(t, t_0)$ and are independent of t for large t.

Some insight into this theorem can be obtained by considering the "singular valued decomposition" (SVD) of $M = M(t, t_0)$ (figure 7.1a). Any real matrix can be decomposed

$$\mathbf{M} = \mathbf{WDV}^T \tag{7.11}$$

where D is a diagonal matrix with diagonal values d_i the square root of the eigenvalues of $\mathbf{M}^T\mathbf{M}$ and V, W are orthogonal matrices, with the columns v_i of V the orthonormal eigenvectors of $M^T M$ and the columns w_i of W the orthonormal eigenvectors of MM^T. Pictorially, this shows us that a unit circle of initial conditions is mapped by M into an ellipse: the principal axes of the ellipse are the w_i and the lengths of the semi axes are d_i. Furthermore the preimage of the w_i are v_i i.e. the v_i are the particular choice of orthonormal axes for the unit circle that are mapped into the ellipse axes. The multiplicative ergodic theorem says that the vectors v_i are *independent* of t for large t, and the d_i yield the Lyapunov exponents in this limit. The vector v_i defines a direction such that an initial displacement in this direction is asymptotically amplified at a rate given by λ_i. For a fixed *final point* $U(t)$ one would similarly expect the w_i to be independent of t_0 for most t_0 and large $t - t_0$. Either the v_i or the w_i may be called Lyapunov eigenvectors.

7.4 Practical Calculation

The difficulty of the calculation is that for any initial displacement vector v (which may be an attempt to approximate one of the v_i) any component along v_1 will

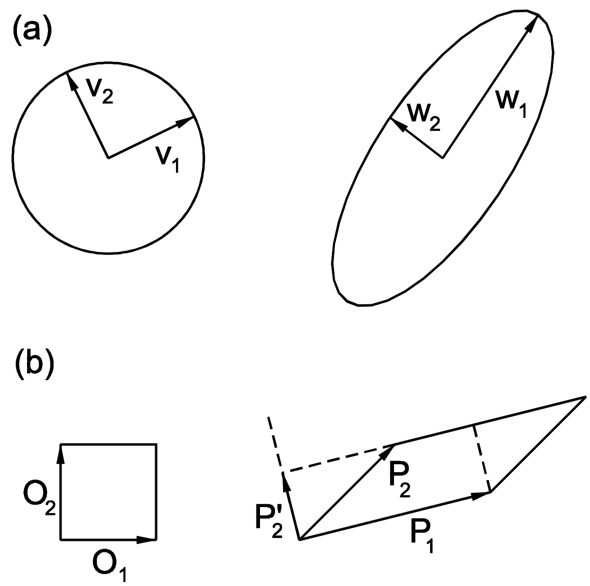

Figure 7.1: Calculating Lyapunov exponents. (a) Oseledec's theorem (SVD picture): orthonormal vectors v_1, v_2 can be found at initial time t_0 that $M(t, t_0)$ maps to orthonormal vectors w_1, w_2 along axes of ellipse. For large $t - t_0$ the v_i are independent of t and the lengths of the ellipse axes grow according to Lyapunov eigenvalues. (b) Gramm-Schmidt procedure: arbitrary orthonormal vectors O_1, O_2 map to P_1, P_2 that are then orthogonalized by the Gramm-Schmidt procedure preserving the growing area of the parallelepiped.

be enormously amplified relative to the other components, so that the iterated displacement becomes almost parallel to the iteration of v_0, with all the information of the other Lyapunov exponents contained in the tiny correction to this. Various numerical techniques have been implemented [2] to maintain control of the small correction, of which the most intuitive, although not necessarily the most accurate, is the method using Gramm-Schmidt orthogonalization after a number of steps [3] (figure 7.1b).

Orthogonal unit displacement vectors $O^{(1)}, O^{(2)}, \ldots$ are iterated according to the Jacobean to give, after some number of iterations n_1 (for a map) or some time Δt_1 (for a flow), $P^{(1)} = \mathbf{M}O^{(1)}$ and $P^{(2)} = \mathbf{M}O^{(2)}$ etc. We will use $O^{(1)}$ to calculate λ_1 and $O^{(2)}$ to calculate λ_2 etc. The vectors $P^{(i)}$ will all tend to align along a single direction. We keep track of the orthogonal components using Gramm-

Schmidt orthogonalization. Write $P^{(1)} = N^{(1)}\hat{P}^{(1)}$ with $N^{(1)}$ the magnitude and $\hat{P}^{(1)}$ the unit vector giving the direction. Define $P'^{(2)}$ as the component of $P^{(2)}$ normal to $P^{(1)}$

$$P'^{(2)} = P^{(2)} - \left(P^{(2)} \cdot \hat{P}^{(1)}\right) \hat{P}^{(1)}. \tag{7.12}$$

and then write $P'^{(2)} = N^{(2)}\hat{P}'^{(2)}$. Notice that the area $P^{(1)} \times P^{(2)} = P^{(1)} \times P'^{(2)}$ is preserved by this transformation, and so we can use $P'^{(2)}$ (in fact its norm $N^{(2)}$) to calculate λ_2. For dimensions larger than 2 the further vectors $P^{(i)}$ are successively orthogonalized to all previous vectors. This process is then repeated and the eigenvalues are given by (quoting the case of maps)

$$\begin{aligned} e^{n\lambda_1} &= N^{(1)}(n_1)N^{(1)}(n_2)\ldots \\ e^{n\lambda_2} &= N^{(2)}(n_1)N^{(2)}(n_2)\ldots \end{aligned} \tag{7.13}$$

etc. with $n = n_1 + n_2 + \ldots$.

Comparing with the singular valued decomposition we can describe the Gramm-Schmidt method as following the growth of the area of parallelepipeds, whereas the SVD description follows the growth of ellipses.

Example 1: the Lorenz Model

The Lorenz equations (chapter 1) are

$$\begin{aligned} \dot{X} &= -\sigma(X - Y) \\ \dot{Y} &= rX - Y - XZ \;. \\ \dot{Z} &= XY - bZ \end{aligned} \tag{7.14}$$

A perturbation $\varepsilon_n = (\delta X, \delta Y, \delta Z)$ evolves according to “tangent space” equations given by linearizing (7.14)

$$\begin{aligned} \delta\dot{X} &= -\sigma(\delta X - \delta Y) \\ \delta\dot{Y} &= r\delta X - \delta Y - (\delta X\, Z + X\, \delta Z) \\ \delta\dot{Z} &= \delta X\, Y + X\, \delta Y - b\delta Z \end{aligned} \tag{7.15}$$

or

$$\frac{d\varepsilon}{dt} = \begin{bmatrix} -\sigma & \sigma & 0 \\ r - Z & -1 & -X \\ Y & X & -b \end{bmatrix} \varepsilon \tag{7.16}$$

defining the Jacobean matrix **K**.

To calculate the Lyapunov exponents start with three orthogonal unit vectors $t^{(1)} = (1, 0, 0)$, $t^{(2)} = (0, 1, 0)$ and $t^{(3)} = (0, 0, 1)$ and evolve the components of each vector according to the tangent equations (7.16). (Since the Jacobean depends on X, Y, Z this means we evolve (X, Y, Z) and the $t^{(i)}$ as a twelve dimensional coupled system.) After a number of iteration steps (chosen for numerical convenience) calculate the magnification of the vector $t^{(1)}$ and renormalize to unit magnitude. Then project $t^{(2)}$ normal to $t^{(1)}$, calculate the magnification of the resulting vector, and renormalize to unit magnitude. Finally project $t^{(3)}$ normal to the preceding *two* orthogonal vectors and renormalize to unit magnitude. The product of each magnification factor over a large number iterations of this procedure evolving the equations a time t leads to $e^{\lambda_i t}$.

Note that in the case of the Lorenz model (and some other simple examples) the trace of **K** is independent of the position on the attractor [in this case $-(1+\sigma+b)$], so that we immediately have the result for the sum of the eigenvalues $\lambda_1+\lambda_2+\lambda_3$, a useful check of the algorithm. (The corresponding result for a map would be for a *constant determinant* of the Jacobean: $\sum \lambda_i = \ln \det |K|$.)

Example 2: the Bakers' Map

For the Bakers' map, the Lyapunov exponents can be calculated analytically. For the map in the form

$$\begin{aligned} x_{n+1} &= \begin{cases} \lambda_a x_n & \text{if} \quad y_n < \alpha \\ (1-\lambda_b) + \lambda_b x_n & \text{if} \quad y_n > \alpha \end{cases} \\ y_{n+1} &= \begin{cases} y_n/\alpha & \text{if} \quad y_n < \alpha \\ (y_n-\alpha)/\beta & \text{if} \quad y_n > \alpha \end{cases} \end{aligned} \tag{7.17}$$

with $\beta = 1-\alpha$ the exponents are

$$\begin{aligned} \lambda_1 &= -\alpha \log \alpha - \beta \log \beta \quad > \quad 0 \\ \lambda_2 &= \alpha \ln \lambda_a + \beta \log \lambda_b \quad < \quad 0 \end{aligned} . \tag{7.18}$$

This easily follows since the stretching in the y direction is α^{-1} or β^{-1} depending on whether y is greater or less than α, and the measure is uniform in the y direction so the probability of an iteration falling in these regions is just α and β respectively. Similarly the contraction in the x direction is λ_a or λ_b for these two cases.

Numerical examples

Numerical examples on 2D maps are given in the demonstrations.

7.5 Other Methods

7.5.1 Householder transformation

The Gramm-Schmidt orthogonalization is actually a method of implementing "QR decomposition". Any matrix **M** can be written

$$\mathbf{M} = \mathbf{QR} \tag{7.19}$$

with **Q** an orthogonal matrix

$$\mathbf{Q} = \begin{bmatrix} \vec{w}_1 & \vec{w}_2 & \cdots & \vec{w}_n \end{bmatrix}$$

and **R** an upper triangular matrix

$$\mathbf{R} = \begin{bmatrix} \nu_1 & * & * & * \\ 0 & \nu_2 & * & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \nu_n \end{bmatrix}, \tag{7.20}$$

where $*$ denotes a nonzero (in general) element. In particular for the tangent iteration matrix **M** we can write

$$\mathbf{M} = \mathbf{M}_{N-1}\mathbf{M}_{N-2}\ldots\mathbf{M}_0 \tag{7.21}$$

for the successive steps Δt_i or n_i for flows or maps. Then writing

$$\mathbf{M}_0 = \mathbf{Q}_1\mathbf{R}_0, \quad \mathbf{M}_1\mathbf{Q}_1 = \mathbf{Q}_2\mathbf{R}_1, \text{ etc.} \tag{7.22}$$

we get

$$\mathbf{M} = \mathbf{Q}_N\mathbf{R}_{N-1}\mathbf{R}_{N-2}\ldots\mathbf{R}_0 \tag{7.23}$$

so that $\mathbf{Q} = \mathbf{Q}_N$ and $\mathbf{R} = \mathbf{R}_{N-1}\mathbf{R}_{N-2}\ldots\mathbf{R}_0$. Furthermore the exponents are

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t - t_0} \ln R_{ii}. \tag{7.24}$$

The correspondence with the Gramm-Schmidt orthogonalization is that the $\mathbf{Q}_i$ are the set of unit vectors $P_1', P_2', \ldots$ etc. and the ν_i are the norms N_i. However an alternative procedure, known as the Householder transformation, may give better numerical convergence [1],[4].

7.5.2 Evolution of the singular valued decomposition

The trick of this method is to find a way to evolve the matrices $\mathbf{W}, \mathbf{D}$ in the singular valued decomposition (7.11) directly. This appears to be only possible for continuous time systems, and has been implemented by Kim and Greene [5].

7.6 Significance of Lyapunov Exponents

A positive Lyapunov exponent may be taken as the defining signature of chaos. For attractors of maps or flows, the Lyapunov exponents also sharply discriminate between the different dynamics: a fixed point will have all negative exponents; a limit cycle will have one zero exponent, with all the rest negative; and a m-frequency quasiperiodic orbit (motion on a m-torus) will have m zero eigenvalues, with all the rest negative. (Note, of course, that a fixed point on a map that is a Poincaré section of a flow corresponds to a periodic orbit of the flow.) For a flow there is in fact always one zero exponent, except for fixed point attractors. This is shown by noting that the phase space velocity satisfies the tangent equations:

$$\frac{d\dot{U}^{(i)}}{dt} = \frac{\partial F_i}{\partial U^{(j)}}\dot{U}^{(j)} \tag{7.25}$$

so that for this direction

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \log\left|\dot{U}(t)\right| \tag{7.26}$$

which tends to zero except for the approach to a fixed point.

7.7 Lyapunov Eigenvectors

This section is included because I became curious about the vectors defined in the Oseledec theorem, and found little discussion of them in the literature. It can well be skipped on a first reading (and probably subsequent ones, as well!).

The vectors v_i—the direction of the initial vectors giving exponential growth—seem not immediately accessible from the numerical methods for the exponents (except the SVD method for continuous time systems [5]). However the w_i are naturally produced by the Gramm-Schmidt orthogonalization. The relationship of these orthogonal vectors to the natural stretching and contraction directions seems quite subtle however.

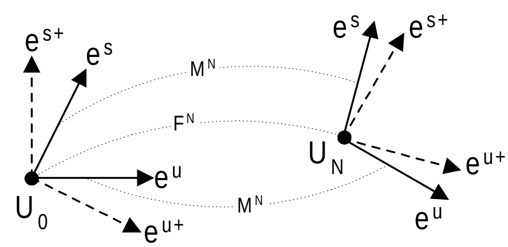

Figure 7.2: Stretching direction $\vec{e}^u$ and contracting direction $\vec{e}^s$ at points U_0 and $U_N = F^N(U_0)$. The vector $\vec{e}^u$ at U_0 is mapped to a vector along $\vec{e}^u$ at U_N by the tangent map $\mathbf{M}^N$ etc. The adjoint vectors $\vec{e}^{u+}$, $\vec{e}^{s+}$ are defined perpendicular to $\vec{e}^s$ and $\vec{e}^u$ respectively. An orthogonal pair of directions close to $\vec{e}^s$, $\vec{e}^{u+}$ is mapped by $\mathbf{M}^N$ to an orthogonal pair close to $\vec{e}^u$, $\vec{e}^{s+}$.

The relationship can be illustrated in the case of a map with one stretching direction $\vec{e}^u$ and one contracting direction $\vec{e}^s$ in the tangent space. These are unit vectors at each point on the attractor conveniently defined so that separations along $\vec{e}^s$ asymptotically contract exponentially at the rate e^{λ_-} per iteration for *forward* iteration, and separations along $\vec{e}^u$ asymptotically contract exponentially at the rate $e^{-\lambda_+}$ for *backward* iteration. Here λ_+, λ_- are the positive and negative Lyapunov exponents. The vectors $\vec{e}^s$ and $\vec{e}^u$ are tangent to the stable and unstable manifolds to be discussed in chapter 22, and have an easily interpreted physical significance. How are the orthogonal "Lyapunov eigenvectors" related to these directions? Since $\vec{e}^s$ and $\vec{e}^u$ are not orthogonal, it is useful to define the adjoint unit vectors $\vec{e}^{u+}$ and

$\vec{e}^{s+}$ as in Fig.(7.2) so that

$$\vec{e}^{s} \cdot \vec{e}^{u+} = \vec{e}^{u} \cdot \vec{e}^{s+} = 0. \qquad (7.27)$$

Then under some fixed large number of iterations N it is easy to convince oneself that orthogonal vectors $\vec{e}_1^{(0)}, \vec{e}_2^{(0)}$ asymptotically close to the orthogonal pair $\vec{e}^s, \vec{e}^{u+}$ at the point U_0 on the attractor are mapped by the tangent map $\mathbf{M}^N$ to directions $\vec{e}_1^{(N)}, \vec{e}_2^{(N)}$ asymptotically close to the orthogonal pair $\vec{e}^u, \vec{e}^{s+}$ at the iterated point $U_N = F^N(U_0)$, with expansion factors given asymptotically by the Lyapunov exponents (see Fig.(7.2)). For example $\vec{e}^s$ is mapped to $e^{N\lambda_-}\vec{e}^s$. However a small deviation from $\vec{e}^s$ will be amplified by the amount $e^{N\lambda_+}$. This means that we can find an $\vec{e}_1^{(0)}$ given by a carefully chosen deviation of order $e^{-N(\lambda_+ - \lambda_-)}$ from $\vec{e}^s$ that will be mapped to $\vec{e}^{s+}$. Similarly almost all initial directions will be mapped very close to $\vec{e}^u$ because of the strong expansion in this direction. Deviations in the direction will be of order $e^{-N(\lambda_+ - \lambda_-)}$. In particular an $\vec{e}_2^{(0)}$ chosen orthogonal to $\vec{e}_1^{(0)}$, i.e. very close to $\vec{e}^{u+}$, will be mapped very close to $\vec{e}^u$. Thus vectors very close to $\vec{e}^s, \vec{e}^{u+}$ at the point U_0 satisfy the requirements for the v_i of Oseledec's theorem and $\vec{e}^u, \vec{e}^{s+}$ at the iterated point $F^N(U_0)$ are the w_i of the SVD and the vectors of the Gramm-Schmidt procedure. It should be noted that for $2N$ iterations rather than N (for example) the vectors $\vec{e}_1^{(0)}, \vec{e}_2^{(0)}$, mapping to $\vec{e}^u, \vec{e}^{s+}$ at the iterated point U_{2N}, must be chosen as a very slightly *different* perturbation from $\vec{e}^s, \vec{e}^{u+}$— equivalently the vectors $\vec{e}_1^{(N)}, \vec{e}_2^{(N)}$ at U_N will *not* be mapped under a further N iterations to $\vec{e}^u, \vec{e}^{s+}$ at the iterated point U_{2N}.

It is apparent that even for this very simple two dimensional case neither the v_i nor the w_i separately give us the directions of both $\vec{e}^u$ and $\vec{e}^s$. The significance of the orthogonal Lyapunov eigenvectors in higher dimensional systems remains unclear.

Bibliography

[1] J-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985)

[2] K. Geist, U. Parlitz, and W. Lauterborn, Prog. Theor. Phys. **83**, 875 (1990)

[3] A. Wolf, J.B. Swift, H.L. Swinney, and J.A. Vastano, Physica **16D**, 285 (1985)

[4] P. Manneville, in “Dissipative Structures and Weak Turbulence” (Academic, Boston 1990), p280.

[5] J.M. Green and J-S. Kim, Physica **24D**, 213 (1987)