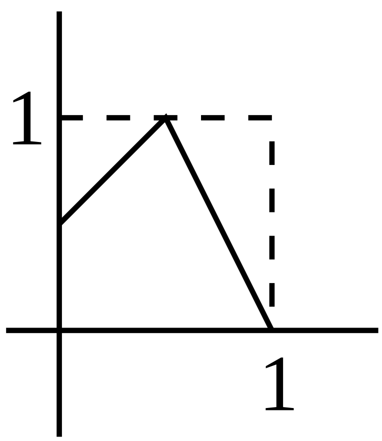
Lecture 3: Invariant Measures [1]

Finding the attracting set of a dynamical system tells us asymptotically what parts of the phase space are visited, but that is not the only thing of interest about the long-time behavior of a dynamical system. For instance, a "typical" trajectory of the system may spend a lot of time in one part of the attracting set, and very little in another part. An *invariant measure* on the attractor tells us how much time "on average", a trajectory spends in a given region.

Example 3.1 *Consider the dynamical system defined by the one-dimensional map*

$$f(x) = \begin{cases} x + 1/2 & ; \quad 0 \le x \le 1/2 \\ 2 - 2x & ; \quad 1/2 < x \le 1 \end{cases}$$



If we compute the Lyapunov exponent of the dynamical numerically, we find that after 1000 *iterations one has the following approximation to the Lyapunov exponents for three choices of initial point* x_0.

$$\begin{aligned} x_0 &= 0.3333 \qquad \gamma \approx 0.4588... \\ x_0 &= 0.4012 \qquad \gamma \approx 0.4637... \\ x_0 &= 0.7111 \qquad \gamma \approx 0.4644... \end{aligned}$$

Clearly, there seems to be a well defined value for the Lyapunov exponent (in fact, we knew that must be true by Oseledec's theorem). However, it is not so clear how to compute it. There is no tendency for nearby trajectories to separate when they are in the interval between 0 and 1/2, while in the interval from 1/2 to 1, the separation between two nearby trajectories doubles with each iteration. Thus to determine the Lyapunov exponent theoretically, we need to know how much of an orbit is spent in the interval $[0, 1/2]$ and how much in the interval from $[1/2, 1]$.

Intuitively, this means that we want to define a measure ρ on the phase space, so that

$$\rho(\mathcal{O}) = \{\text{fraction of points in an orbit that lie in } \mathcal{O}\}$$

Note that such a measure will be concentrated on the attractor, since *asymptotically* an arbitrarily large fraction of the points in the orbit lie in an arbitrarily small neighborhood of the attractor.

[1] These lecture notes are for the course MA 574, and are being prepared for the instructor's personal use. They are not for public distribution.

Let $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ be the orbits of a discrete dynamical system with initial condition $\mathbf{x}_0$. We define the invariant measure associated with this orbit to be

$$\rho = \lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \delta_{\mathbf{x}_j} \tag{1}$$

if this limit exists. Here $\delta_{\mathbf{x}_k}$ is the Dirac δ measure concentrated at the point $\mathbf{x}_j$. Note that if $\mathcal{O}$ is any (measurable) set,

$$\begin{aligned}\rho(\mathcal{O}) &= \lim_{k\to\infty} \frac{1}{k} \int_{\mathcal{O}} \sum_{j=0}^{k-1} \delta_{\mathbf{x}_j}(\xi) d\xi \\ &= \lim_{k\to\infty} \frac{\text{number of points in } \{\mathbf{x}_0, \ldots, \mathbf{x}_{\mathrm{j}-1}\}}{k}\end{aligned}$$

Thus, if the limit in (1) exists, this measure accords exactly with our intuitive idea what the measure should be.

Above, I referred to this measure as an invariant measure. Let me now explain what I mean by that. Suppose that ϕ is any (measureable) function. Then assuming as always that the limit in (1) makes sense, the integral of ϕ with respect to the measure ρ is defined as

$$\int \phi(\xi) d\rho(\xi) = \lim_{k\to\infty} \frac{1}{k} \int \sum_{j=0}^{k-1} \phi(\xi)\delta_{\mathbf{x}_j}(\xi) d\xi = lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \phi(\mathbf{x}_j) \; . \tag{2}$$

We say that the measure ρ is *invariant* for the discrete dynamical system defined by the function $\mathbf{f}$ if for every continuous function ϕ,

$$\int \phi \circ \mathbf{f}(\xi) d\rho(\xi) = \int \phi(\xi) d\rho(\xi) \tag{3}$$

That is, composing the integrand with the function $\mathbf{f}$ that defines the dynamical system doesn't change the integral.

Lemma 3.2 *The measure defined by (1) is invariant.*

Proof: By (1),

$$\begin{aligned}\int \phi \circ \mathbf{f}(\xi) d\rho(\xi) &= lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \phi \circ \mathbf{f}(\mathbf{x}_j) \\ &= lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \phi(\mathbf{x}_{j+1}) \\ &= lim_{k\to\infty} \frac{1}{k} \{ \left(\sum_{j=0}^{k-1} \phi(\mathbf{x}_j) \right) + \phi(\mathbf{x}_k) - \phi(\mathbf{x}_0) \} \\ &= lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \phi(\mathbf{x}_j) + lim_{k\to\infty} \frac{1}{k} \left(\phi(\mathbf{x}_k) - \phi(\mathbf{x}_0) \right) \\ &= lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} \phi(\mathbf{x}_j) \\ &= \int \phi(\xi) d\xi\end{aligned}$$

Remark 3.3 *It is often common to define an invariant measure in terms of the measure of a set $\mathcal{O}$ rather than the integral of a function ϕ. In this approach, one says that ρ is invariant if for every measurable set $\mathcal{O}$,*

$$\rho(\mathcal{O}) = \rho(\mathbf{f}^{-1}(\mathcal{O})) \ . \tag{4}$$

Although the appearance of the inverse of $\mathbf{f}$ *rather than* $\mathbf{f}$ *itself may be surprising, this definition is equivalent to the definition in (3). To see why, take the function ϕ in (3) to be the characteristic function of the set $\mathcal{O}$. (Strictly speaking, we defined (3) to hold only for continuous functions and the characteristic function of $\mathcal{O}$ is not continuous, however, it can be arbitrarily well approximated by continuous functions, and this suffices to show that (3) also holds for such functions.) Then*

$$\int \phi(\xi) d\rho(\xi) = \int_{\xi \in \mathcal{O}} d\rho(\xi) = \rho(\mathcal{O}) \ .$$

On the other hand,

$$\int \phi \circ \mathbf{f}(\xi) d\rho(\xi) = \int_{\mathbf{f}(\xi) \in \mathcal{O}} d\rho(\xi) = \int_{\xi \in \mathbf{f}^{-1}(\mathcal{O})} d\rho(\xi) = \rho(\mathbf{f}^{-1}(\mathcal{O})) \ .$$

Since these two expressions are equal by (3), we see that (4) follows. By a similar argument, one can show that (3) follows from (4)

Remark 3.4 *Note that the measure defined by (1) has the property that $\int d\rho(\xi) = 1$. We will require this of all the measures we consider, and hence refer to them as invariant* probability *measures. This normalization is quite natural in light of our interpretation of these measures as the fraction of the total time that an orbit spends in some part of the phase space.*

Thus far, we have begged the question of whether such invariant probability measures actually exist – we have simply assumed that the limit in (1) was well defined. We now show that there are in fact, lots of such measures in a typical dynamical system.

Example 3.5 *Let $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{k-1}\}$ be a periodic orbit of length k for a dynamical system defined by the function* $\mathbf{f}$. *Define $\rho = \frac{1}{k}\sum_{j=0}^{k-1} \delta_{\mathbf{x}_j}$. Note that since there is no limit here, the measure is well defined. Furthermore, it is clearly normalized so that $\int d\rho = 1$. All that remains is to check that it is invariant. Given any set $\mathcal{O}$,*

$$\rho(\mathcal{O}) = \frac{1}{k}\int_{\xi \in \mathcal{O}} \sum_{j=0}^{k-1} \delta_{\mathbf{x}_j}(\xi) d\xi = \frac{\#\{\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{k-1}\} \cap \mathcal{O}\}}{k} \tag{5}$$

while

$$\rho(\mathbf{f}^{-1}(\mathcal{O})) = \frac{1}{k}\int_{\mathbf{f}(\xi) \in \mathcal{O}} \sum_{j=0}^{k-1} \delta_{\mathbf{x}_j}(\xi) d\xi = \frac{\#\{\{\mathbf{f}(\mathbf{x}_0), \mathbf{f}(\mathbf{x}_1), \ldots, \mathbf{f}(\mathbf{x}_{k-1})\} \cap \mathcal{O}\}}{k} \tag{6}$$

But $\{\mathbf{f}(\mathbf{x}_0), \mathbf{f}(\mathbf{x}_1), \ldots, \mathbf{f}(\mathbf{x}_{k-1})\} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_0\}$, because $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{k-1}\}$ is a periodic orbit, so (5) and (6) are equal, and the measure is invariant. Thus, we see that at the very least, every dynamical system which has a periodic orbit has an invariant probability measure.

In fact, one has in general

Theorem 3.6 *If the function* $\mathbf{f}$ *defining the dynamical system maps some closed, bounded subset of* $\mathbf{R}^n$ *to itself, then* $\mathbf{f}$ *has at least one invariant probability measure.*

Indeed, the problem in general is not that there are too few invariant measures, but rather too many! Example 3.5 showed that there is an invariant probability measure for every periodic orbit in the system. The tent map, for example, has infinitely many periodic orbits and hence infinitely many invariant measures. This is typical of chaotic attractors. What's worse (or better, depending on your point of view) is that given any two distinct, invariant probability measures, ρ_1 and ρ_2, one can combine them to get a third by taking $\tilde{\rho} = \rho_1 + \rho_2$. In order to restrict the class of invariant measures we must study, we will rule out combinations of this type by focusing on *ergodic* measures. These are measures that are "indecomposable".

Definition 3.7 *An invariant probability measure* ρ *is* ergodic *if it* cannot *be written as* $\rho = \frac{1}{2}(\rho_1 + \rho_2)$, *for* ρ_1 *and* ρ_2 *two distinct, invariant, probability measures.*

Remark 3.8 *Concentrating our attention to ergodic measures is not such a great restriction since (subject to mild technical restrictions), any invariant probability measure can be decomposed into a sum (or integral) of ergodic, invariant probability measures.*

As an exercise, show that if $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{k-1}\}$ is a periodic orbit, the measure $\rho = \sum_{j=0}^{k-1} \delta_{\mathbf{x}_j}$ is an ergodic measure.

An alternative (but equivalent) definition of ergodicity which is frequently useful is that ρ is an ergodic invariant probability measure for the dynamical system defined by $\mathbf{f}$ if all invariant sets of $\mathbf{f}$ have measure either zero or one. Symbolically, we mean that if $\mathbf{f}^{-1}(\mathcal{O}) = \mathcal{O}$, then $\rho(\mathcal{O})$ is either zero or one. It's a very good exercise to show why this definition is equivalent to our previous one.

Aside from being the basic building blocks for invariant probability measures, ergodic measures have another extremely important property – they allow one to introduce averages over time and space. Many of the quantities that one measures in a dynamical system require one to evaluate something at each point along an orbit of the dynamical system, and then average the result over time. This is difficult for (at least) two reasons. First, it may take a long time in order to collect enough data to insure a good estimate of the average. Secondly, one would have to wonder whether or not one would get the same answer if one averaged over another orbit with different initial conditions. For a dynamical system with an ergodic, invariant probability measure, such questions are answered by the ergodic theorem.

Theorem 3.9 *(Ergodic Theorem) Suppose that* ρ *is an ergodic, invariant probability measure for* $\mathbf{x}_{m+1} = \mathbf{f}(\mathbf{x}_m)$. *Let* ϕ *be an integrable function. Then*

$$\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} \phi(\mathbf{x}_j) = \int \phi(\xi) d\rho(\xi) \ , \tag{7}$$

for almost every choice of initial condition $\mathbf{x}_0$.

Remark 3.10 *The qualification "almost ever* $\mathbf{x}_0$*" means that if* $\mathcal{B} = \{\mathbf{x}_0 \mid (7) \text{ fails to hold}\}$ *then* $\rho(\mathcal{B}) = 0$.

Remark 3.11 *Note that two things are being asserted here. First that the time-average of ϕ exists for almost every orbit, and secondly that this time average equals the "space average", –i.e. the integral of ϕ with respect to ρ.*

Remark 3.12 *In fact, the ergodic theorem is considerably more general than I have stated here. It actually shows that for* any *invariant measure ρ, the time averages exist on almost all orbits. However, for non-ergodic measures we can't expect that these time averages will equal the spatial averages.*

Proof: The hard part of this theorem is showing that the time averages actually exist. However, this seems "intuitively" reasonable (I think!) so I will assume they exist and show that they equal the spatial average given on on the right-hand-side of (7). Let $T(\mathbf{x}_0) = \lim_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1}\phi(\mathbf{x}_j)$, and assume that this limit exists, at least for almost every $\mathbf{x}_0$. Define $M = \{\mathbf{x}_0 \mid (7)\ \textit{holds}\}$. That is, M is the set of points at which the space and time averages are equal. Note that because ρ is an invariant measure, M is also invariant – *i.e.* $M = f^{-1}(M)$. Therefore, since ρ is ergodic, $\rho(M)$ is either zero or one. If M has measure one, the theorem is true, so we assume that $\rho(M) = 0$ and show that this leads to a contradiction. Define $M^> = \{\mathbf{x}_0 \mid T(\mathbf{x}_0) > \int \phi d\rho\}$, and $M^< = \{\mathbf{x}_0 \mid T(\mathbf{x}_0) < \int \phi d\rho\}$. Both $M^>$ and $M^<$ are invariant sets, and hence each has measure either zero or one. If the measure of M is zero, either $M^>$ or $M^<$ must have measure one. Assume that $\rho(M^>) = 1$. Then $\int T d\rho = \int_{M^>} T d\rho(\xi) > \int \phi d\rho$. But $\int T d\rho = \lim_{k\to\infty}\frac{1}{k}\int \phi(\mathbf{f}^{(k)}(\xi)d\rho(\xi) = \int \phi d\rho$, using the fact that ρ is an invariant measure. This is a contradiction, and hence $\rho(M) = 1$.

Example 3.13 *Let's now return to the Example 3.1 with which we started this lecture. Note that for a one dimensional map, the Lyapunov exponent is*

$$\gamma = \lim_{k\to\infty}\frac{1}{k}(\log|Df(x_{k-1})| + \log|Df(x_{k-1})| + \ldots + \log|Df(x_0)|)$$

Thus, it is the time-average of $\log|Df(x)|$. *We will try to compute it by finding the physically relevant invariant measure. There will be lots of invariant measures concentrated on periodic orbits, but those are "atypical" because typical orbits for this dynamical system don't fall on a periodic orbit, but wander around all over the interval from* $[0,1]$. *Since the behavior of the function has constant slope on the interval* $[0,1/2]$, *we make the guess that the invariant measure has a constant density w_1 on the interval* $[0,1/2]$, *and similarly that it has some other density w_2 on the interval* $[1/2,1]$. *To figure out what w_1 and w_2 are we use the fact that the measure is invariant, so that $\rho(I) = \rho(f^{-1}(I))$ for any interval I. If I is a subinterval of* $[0,1/2]$, *it's preimage is an interval of half the length in* $[1/2,1]$. *Thus, we expect that $w_1 = (1/2)w_2$. On the other hand, since this is a probability measure, the total measure of the interval* $[0,1]$ *must be 1. Thus, $(1/2)w_1 + (1/2)w_2 = 1$. These two equations together imply that $w_1 = 2/3$, and $w_2 = 4/3$. I leave it as an exercise to show that this measure is ergodic, but assuming that it is, we see that by the ergodic theorem, the Lyapunov exponent (for almost all choices of the initial condition x_0 should be*

$$\gamma = \int \log|Df(x)|d\rho(x) = (1/2)w_1\log(1) + (1/2)w_2\log(2) = (2/3)\log 2 = 0.462...$$

which is in good agreement with our numerical calculations.