Dynamics and time series: introduction to ergodic theory

Giulio Tiozzo

January 18, 2010

Motivation

The goal is to study the asymptotic behaviour of orbits from a statistical point of view: of particular interest are deterministic systems whose distributions of orbits after a long period of time are the same as the ones obtained for random systems.

Setting

X phase space (endowed with some σ-algebra $\mathcal{A}$)
$T : X \to X$ the *time evolution*.
μ probability measure on $(X, \mathcal{A})$ which is *invariant*, i.e.

$$\mu(T^{-1}(A)) = \mu(A) \qquad \text{for every } A \subseteq X$$

$$\text{or equivalently } \int_X f d\mu = \int_X f \circ T d\mu \quad \forall f \in L^1(X, d\mu)$$

The quadruple $(X, \mathcal{A}, \mu, T)$ is called a *measurable dynamical system.*

Examples

- In mechanics, the phase space of N moving particles is

$$X = \{(r_1, v_1, \dots, r_N, v_N)\} \cong \mathbb{R}^{6N}$$

  where each r_i is a vector representing the position and of the i-th particle and v_i is a vector representing the velocity of the i-th particle. The time evolution is given by solving the differential equation of motion. In Hamiltonian mechanics, a classical invariant measure is the Liouville measure.

- Rotation on the circle : $X = S^1 = \mathbb{R}/\mathbb{Z}$,

$$T : x \mapsto x + \alpha \mod 1$$

  represents the rotation of angle $2\pi\alpha$; Lebesgue measure is invariant.

- Doubling map: $X = S^1 = \mathbb{R}/\mathbb{Z}$,

$$T : x \mapsto 2x \mod 1$$

  Lebesgue measure is again an invariant measure.

Observables

An observable is a function $f : X \to \mathbb{R}$. Typically, one considers only observables which belong to $L^1(X, d\mu)$.
Examples: kinetic energy of a particle, pressure of a gas, price of a stock.

Birkhoff's ergodic theorem

Given a measurable dynamical system $(X, \mathcal{A}, \mu, T)$, for every $f \in L^1(X, d\mu)$, the limit

$$\overline{f}(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$$

exists for almost all $x \in X$ with respect to the measure μ.

The function $\overline{f}(x)$ is called *Birkhoff average* (or *time average*) of the observable f along the (forward) orbit of x. If $f = \chi_A$, one has the *frequency of visit* of the event A

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_A(T^i(x)) = \frac{\#\{1 \leq i \leq n : T^i(x) \in A\}}{n}$$

Applications, backtesting and Bertrand Russell's turkey

In applications, the main way to infer the probability of an event A is to compute its frequency of visit by looking at historical data, i.e. observing a single orbit for some finite number of steps N, and compute

$$\mu(A) \cong \frac{\#\{1 \leq i \leq N : T^i(x) \in A\}}{N}$$

The reliability of such *backtesting* approach can be affected by the following issues:

1. The time average does not necessarily converge to the measure of A for ALL orbits, hence we could be considering an orbit which is not typical enough.

2. The approximation is made by considering a finite number N of steps, which may or may not be enough to predict the asymptotic behaviour.

To address the first issue we will need to define the concept of *ergodicity*; the second issue will be discussed in lecture 5.

Definition. *A measurable dynamical system* $(X, \mathcal{A}, \mu, T)$ *is* ergodic *if for every* $f \in L^1(X, d\mu)$*, the Birkhoff average is constant a.e. with respect to the measure* μ*, i.e.*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int_X f d\mu$$

Proposition. *For a measurable dynamical system, the following properties are equivalent:*

1. *Ergodicity.*

2. *Every event* $A \in \mathcal{A}$ *which is* T*-invariant (i.e. such that* $T^{-1}(A) = A$*) has either* $\mu(A) = 0$ *or* $\mu(A) = 1$*.*

3. *Every observable $f \in L^1(X, d\mu)$ which is T-invariant (i.e. such that $f \circ T = f$ a.e.) is constant a.e. with respect to μ.*

4. *For every pair of events $A, B \in \mathcal{A}$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}(A) \cap B) = \mu(A)\mu(B)$$

Example: rotation on the circle Let us consider the rotation on the circle of angle $2\pi\alpha$

$$x \mapsto x + \alpha \mod 1$$

If α is rational, every point is periodic of the same period. On the other hand, when α is irrational, every orbit is *dense* (i.e. it passes arbitrarily close to any point of the circle). For what α is the system ergodic?

Suppose $\alpha = \frac{1}{2}$ and compute some Birkhoff averages:

- The time average of the observable $f = \chi_{[0,1/2)}$ is always $1/2$.

- If $f = \chi_{[0,1/2]}$, the time average is

$$\overline{f}(x) = \begin{cases} 1/2 & \text{for } x \neq 0, 1/2 \\ 1 & \text{for } x = 0 \text{ or } x = 1/2 \end{cases}$$

  In this case the time average is still constant for ALMOST ALL orbits (since the set $\{0, 1/2\}$ has zero measure), so this does not violate the definition of ergodicity.

- If $f = \chi_{[0,1/4)} + \chi_{[1/2,3/4)}$, then the time average is

$$\overline{f}(x) = \begin{cases} 1 & \text{for } x \in [0, 1/4) \cup [1/2, 3/4) \\ 0 & \text{for } x \in [1/4, 1/2) \cup [3/4, 1) \end{cases}$$

  Since we have found SOME observable such that the time average is not constant almost everywhere, then the system is NOT ergodic.

A similar argument works for every $\alpha \in \mathbb{Q}$. For irrational α we can prove the system to be ergodic: we have to check that every $f \in L^1(X, d\mu)$ such that $f \circ T = f$ is constant almost everywhere. In order to do so, we can exploit the fact that measurable functions on the circle can be thought of as periodic functions on the real line, hence one can write the Fourier expansion of f

$$f(x) = \sum_{k\in\mathbb{Z}} a_k e^{2\pi i k x}$$

so

$$f \circ T = f(x + \alpha) = \sum_{k\in\mathbb{Z}} a_k e^{2\pi i k(x+\alpha)} = \sum_{k\in\mathbb{Z}} a_k e^{2\pi i k\alpha} e^{2\pi i k x}$$

If $f \circ T = f$ almost everywhere, by comparing the coefficients of the Fourier series

$$a_k = a_k e^{2\pi i \alpha} \qquad \forall k \in \mathbb{Z}$$

Since $\alpha \notin \mathbb{Q}$, one can divide by $e^{2\pi i\alpha} - 1 \neq 0$ and get $a_k = 0\ \forall k \in \mathbb{Z} \setminus \{0\}$ which amounts to saying that f is constant a.e.

Mixing and correlation functions

Given an observable f, one can consider the sequence of random variables on $(X, \mathcal{A}, \mu)$

$$F_i := f \circ T^i$$

Since μ is T-invariant, all these variables are identically distributed, but in general will not be independent. Mixing systems are precisely systems where these variables become less and less correlated as time goes by. Formally,

Definition. *A measurable dynamical system is* (strongly) mixing *if for every pair of events* $A, B \in \mathcal{A}$

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) \to \mu(A)\mu(B)$$

This is the same as saying that the *correlation function* of every pair of observables $\phi, \psi \in L^1(X, d\mu)$ goes to zero as $n \to \infty$, i.e.

$$\lim_{n\to\infty} \int_X \phi(\psi \circ T^n) d\mu - \int_X \phi d\mu \int_X \psi d\mu = 0$$

Example

Arnold's cat is the linear map of the two-dimensional torus

$$\begin{array}{rccc} T: & \mathbb{R}^2/\mathbb{Z}^2 & \mapsto & \mathbb{R}^2/\mathbb{Z}^2 \\ & (x,y) & \mapsto & (2x+y, x+y) \qquad \mod \mathbb{Z}^2 \end{array}$$

Remark. *Strong mixing* $\Rightarrow$ *ergodicity.*