Dynamics and time series: theory and applications

Stefano Marmi
Scuola Normale Superiore

Lecture 2, Nov 9 , 2011

Dynamical systems

- A dynamical system is a couple (X phase space, time evolution law: either a map $T:X\to X$ or a flow $g_t :X\to X$, here t is time)
- The phase space X is the set of all possible states (i.e. initial conditions) of our system
- Each initial condition uniquely determines the time evolution (determinism)
- The system evolves in time according to a fixed law (iteration of a map T, flow g_t for example arising from solving a differential equation, etc.)
- Often (but not necessarily) the evolution law is not linear
- Observables are simply scalar functions $\phi:X\to\mathbf{R}$
- Time series naturally arise from the time evolution of the observables:

$\phi(x)$, $\phi(T(x))$, $\phi(T\circ T(x))$, $\phi(T^3 (x))$, ….. Here $T^{n+1}(x)=T \circ T^n (x)$

Measure-preserving transformations

X phase space, μ probability measure

$\Phi: X \to \mathbf{R}$ observable (a measurable function, say L^2).

Let A be subset of X (event).

$\mu(\Phi) = \int_X \Phi \, d\mu$ is the expectation of Φ

$T: X \to X$ induces a time evolution

on observables: $\Phi \to \Phi \circ T$

on events: $A \to T^{-1}(A)$

T is measure preserving if $\mu(\Phi) = \mu(\Phi \circ T)$ i.e. $\mu(A) = \mu(T^{-1}(A))$

Law of large numbers

$\{\mathbf{X_i}\}$ independent identically distributed random variables

$$\mathrm{E}(\mathrm{X_i}) = \mu < +\infty$$

Then $$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \to \mu$$

Weak form:

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} \mathrm{P}(|\overline{X}_n - \mu| < \varepsilon) = 1$$

Strong form:

$$\overline{X}_n \to \mu$$ almost surely

Birkhoff theorem and ergodicity

Birkhoff theorem: if T preserves the measure μ then with probability one the time averages of the observables exist (statistical expectations). The system is ergodic if these time averages do not depend on the orbit (statistics and a-priori probability agree)

$$\frac{1}{N}\sum_{0}^{N-1}\varphi\circ T^i(x) := \frac{1}{N}S_N\varphi(x) \longrightarrow \int_X \varphi(t)d\mu(t)$$

$$\frac{1}{N}\#\{i\in[0,N), T^i(x)\in A\} \longrightarrow \mu(A)$$

Law of large numbers: Statistics of orbits = a-priori probability

Law of large numbers vs Birkhoff theorem

Random setting

$\{\mathbf{X_i}\}$ i.i.d. random variables

$$\mathrm{E}(\mathrm{X_i}) = \mu < +\infty$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to \mu$$

almost surely

Deterministic setting

$$T: X \to X$$

$f \in L^1(X, d\mu)$ observable

$$X_i := \{f \circ T^i\}$$

are not necessarily independent

If T ergodic

$$\frac{1}{n}\sum_{i=1}^{n} f \circ T^i \to \int f\, d\mu$$

almost surely

“Historia magistra vitae” or the mathematical foundation of backtesting

- Without assuming ergodicity, Birkhoff theorem shows that:

- Time averages exist and they give rise to an experimental statistics to compare with theory
- Past and future time averages agree almost everywhere

Recurrence times

- A point is recurrent when it is a point of accumulation of its future (and past) orbit
- Poincarè recurrence: given a dynamical system T which preserves a probability measure μ and a set of positive measure E a point x of E is almost surely recurrent
- First return time of x in E:

$$R(x,E)=\min\{n>0, T^n x \in E\}$$

- E could be an element of a partition of the phase space (symbolic dynamics): this point of view is very important in applications (e.g. the proof of optimality of the Lempel-Ziv data compression algorithm)

Kac's Lemma

- If T is ergodic and E has positive measure then

$$\int_E R(x,E)\,d\mu(x)=1 ,$$

i.e. $R(x,E)$ is of the order of $1/\mu(E)$: the average length of time that you need to wait to see a particular symbol is the reciprocal of the probability of a symbol. Thus, we are likely to see the high-probability strings within the window and encode these strings efficiently.

The ubiquity of “cycles” (as long as they last…)

Furstenberg’s recurrence: If E is a set of positive measure in a measure-preserving system, and k is a positive integer, then there are infinitely many integers n for which

$$\mu(E \cap T^{-n}(E) \cap \ldots \cap T^{-(k-1)n}(E) > 0$$

Strong vs. weak mixing: on events

- Strongly mixing systems are such that for every E, F we have

  $$\mu(T^n(E) \cap F) \to \mu(E)\,\mu(F)$$

  as n tends to infinity; the Bernoulli shift is a good example. Informally, this is saying that shifted sets become asymptotically independent of unshifted sets.

- Weakly mixing systems are such that for every E, F we have

  $$\mu(T^n(E) \cap F) \to \mu(E)\,\mu(F)$$

  as n tends to infinity *after excluding a set of exceptional values of n of asymptotic density zero*.

- Ergodicity does not imply $\mu(T^n(E) \cap F) \to \mu(E)\,\mu(F)$ but says that this is true for Cesaro averages:

  $$1/n \sum_0^{n-1} \mu(T^j(E) \cap F) \to \mu(E)\,\mu(F)$$

Mixing: on observables

Order n correlation coefficient:

$$c_n(\varphi,\psi) := \int \varphi.\psi \circ T^n d\mu - \int \varphi d\mu \int \psi d\mu$$

Ergodicity implies

$$\frac{1}{N}\sum_{0}^{N-1} c_n(\varphi,\psi) \longrightarrow 0$$

Mixing requires that

$$c_N(\varphi,\psi) \longrightarrow 0$$

namely φ and $\varphi \circ T^n$ become independent of each other as $n \to \infty$

Mixing of hyperbolic automorphisms of the 2-torus (Arnold’s cat)