Rigorous approximation of invariant measures for IFS

Joint work with S. Galatolo e I. Nisoli

Maurizio Monge

maurizio.monge@im.ufrj.br

Universidade Federal do Rio de Janeiro

April 8, 2016

Overview

- Invariant (stationary) measures.
- Iterate function systems.
- The problem of the computation of stationary measures.
- Tools (spectral approximation).
- Approximation strategy.
- Application to the IFS case.
- A priori contraction estimates.
- Notes on the implementation.
- Related result (on mixing time).

Invariant measure as a statistical invariant (1/2)

- Let $T : X \to X$ be a transformation (dynamical system), where X is a space equipped with a Borel σ-algebra and a Lebesgue measure $\mathcal{L}$.
- A probability measure μ is said *invariant measure* if we have

$$\mu(A) = \mu(T^{-1}(A)).$$

- That is, it is invariant applying the *transfer operator* L_T associated to T acting on the space of measures. $L_T(\mu)$ is defined as

$$L_T(\mu)(A) = \mu(T^{-1}(A)), \text{ for each } A \in \mathcal{B}.$$

- In this case we have

Theorem 1 (Birkhoff's ergodic)

For each μ-integrable function $f : X \to \mathbb{R}$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int f \mathrm{d}\mu, \ \mu\text{-almost every } x \in X.$$

Invariant measure as a statistical invariant (2/2)

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int f \mathrm{d}\mu, \ \ \mu\text{−almost every } x \in X.$$

- An invariant measure μ determines the statistics of an observable for μ-almost all points.
- In general a dynamical system admits several invariant measures, and many of them are supported on a set with Lebesge measure zero.
- An invariant measure μ is considered a satisfactory statistical invariant when it describes the statistics of the observables for a Lebsgue-non trival set of points.
- In such a case μ is said *physical measure*. Its support may still have Lebsegue measure zero (e.g. in the case of an attracting fixed point).

Iterated Function Systems - IFS

- An *Iterated Function System* (IFS) on X is the data of a family of functions $T_1, \ldots, T_n : X \to X$, and probabilities $p_1, \ldots, p_n$ (summing to 1). We have a stochastical dynamical system where at each step a funcion f is chosen and applied, where each f_i is chosen with independent probability p_i.
- The equivalent of the transfer operator for an IFS is defined as

$$L = \sum_i p_i L_{T_i},$$

  where L_{T_i} is the transfer operator corresponding to the transformation T_i.
- We have the following interpretation: if μ is a measure describing the probability distribution of the point x in the space X, $L(\mu)$ describes the probability distribution of the image under one application of the IFS.
- A measure invariant under L is said *stationary measure* for the IFS.

The problem of the rigorous approximation

- The purpouse of this project is developing programs to work concretely with different examples of dynamical systems, and allowing to compute the stationary measure up to a rigorous and certified error.
- The stationary measure is an invariant that it is worth approximating with certified error, as it allows to understand the behaviour of the observables, and approximate other invariants such as the entropy, Lyapunov exponents, and so on.
- We will also study the *speed of convergence to the equilibrium*, for its interest in the estimation of the "escape rates", and the variation under small perturbations ("linear response").
- The long term goal is developing instruments that may be useful in computer assisted proofs.

Computation of invariant measures

- There exist computable systems having non-computable invariant measure [Galatolo-Hoyrup-Rojas, 2011].
- The "naive" simulation appears to be very effective for approximating invariant measures but fails dramatically for the map of the interval $x \mapsto 2x$. This phenomenon is related to the representation of numbers in base 2 on the computer, and does not appear for the map $x \mapsto 3x$.
- Out approach uses the transfer operator L, approximated by a Markov chain L_δ on a finite number of states. We compute the stationary probability distribution, and relate such distribution to the stationary measure of the system.
- There exist powerful spectral stability results that allow to do this in a suitable functional context (Keller-Liverani's stability theorem), but they are hard to use in practice.

Stability of fixed points

Let $\mathcal{B}$ be Banach space of signed measure, which we assume to be preserved by L. Assume L_δ to be an approximation of L.

Lemma 2 (Variation on Galatolo-Nisoli '11)

Let μ, $\mu_\delta \in \mathcal{B}$ be probability measures invariant under L, L_δ respectively. Let $V = \{\mu \in \mathcal{B} \text{ s.t. } \mu(X) = 0\}$, and assume $L_\delta(V) \subseteq V$. Let's assume:

(A) $\|L_\delta \mu - L\mu\|_{\mathcal{B}} \leq \epsilon$ *(true when L_δ approximates L),*

(B) $\exists N$ *such that* $\|L_\delta^N|_V\|_{\mathcal{B}} < \frac{1}{2}$,

(C) *Let* $C = \|\sum_{i\in[0,N-1]} L_\delta^i|_V\|_{\mathcal{B}}$, *then*

$$||\mu_\delta - \mu||_{\mathcal{B}} \leq 2\epsilon C.$$

Other than condition (A), all other conditions only depend on L_δ, that we assume to be representable on a computer (up to a computable error).

Approximation in the case of expanding maps of the interval

- A transformation of the interval T is said *piecewise expanding* if the interval can be partitioned in a finite number of interval (c_i, c_{i+1}) such that T is C^2, $|T'| \geq 2$, and $T''/(T')^2$ is bounded.
- In the case of piecewise expanding maps we can apply the above strategy using *Ulam approximation* in the space of finite signed measures:

$$L_\delta = \pi_\delta L \pi_\delta,$$

  where $\pi_\delta(\mu) = E(\mu|\Pi)$, for a partition of the interval Π in intervals of size δ.
- The operator π_δ is a contration in the L^1-norm (assuming the L^1-norm of a finite signed measure to be the "total mass").
- Observe that $\|Id - \pi_\delta\|_{BV \to L^1} \leq \delta$.

Laota-Yorke inequality, and norm estimation

- A piecewise expanding maps satisfies the following theorem:

Theorem 3 (version in Liverani, 2004)

Let T be piecewise expanding, and μ be a finite measure on the interval $[0,1]$. Then

$$\|L_T\mu\|_{BV} \leq \lambda \cdot \|\mu\|_{BV} + B \cdot \|\mu\|_1,$$

for

$$\lambda = 2 \cdot \left\| \frac{1}{T'} \right\|_\infty, \qquad B = \frac{2}{min(c_i + c_{i+1})} + 2 \left\| \frac{T''}{(T')^2} \right\|_\infty.$$

- Iterating, if μ is as invariant measure, as $L_T\mu = \mu$ we obtain that

$$\|\mu\|_{BV} \leq \frac{B}{1-\lambda}.$$

Application of the theorem

- Consequently we can satisfy the point (A) of the approximation theorem with respect to the L_1 norm, because

$$\|(L_\delta - L)\mu\|_{L^1} \leq \|\mu\|_{BV} \cdot \|L_\delta - L\|_{BV\to L^1}$$

- At point (B), the estimation of $\|L_\delta^N|_V\|_{L^1} < \frac{1}{2}$ can be proved by the computer (and is what often requires most computing power!)
- At point (C), the term $\|\sum_{i\in[0,N-1]} L_\delta^i|_V\|_{L^1}$ can be estimated *a priori*, and possibly improved computationally.
- The theorem provides the error between the fixed point of L and the fixed point of L_δ.
- The fixed point of L_δ (that is representable as stochastic matrix) can be computed with certified error.
- The goodness of the approximation depends on B! The same L_δ could be the approximation for different systems, that satisfy Lasota-Yorke inequalities with very different B's.

Expanding IFS - Example of a rigorous computation (1/2)

- For different values of p_1 and $p_2 = 1 - p_1$, let's consider the transformations

$$T_2(x) = 4x + 0.01 \cdot sin(16\pi x), \qquad T_2(x) = 5x + 0.03 \cdot sin(5\pi x).$$

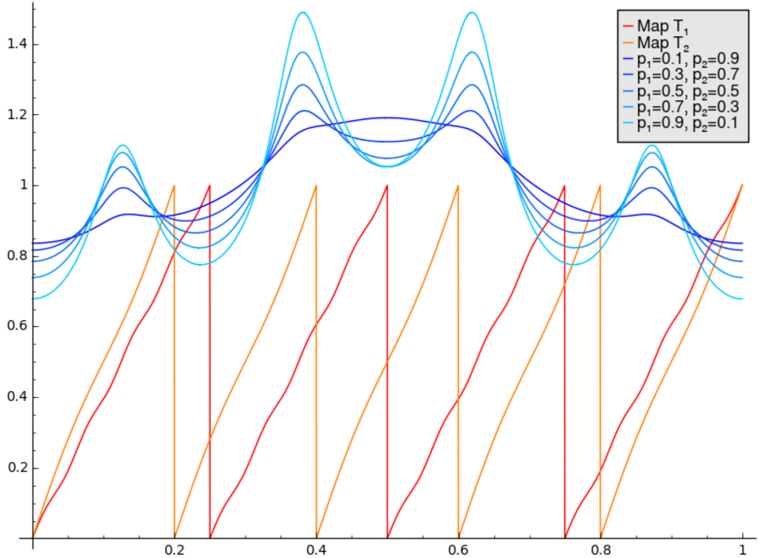
- The values of λ, B and μ_{BV} can be computed as

p_1	0.1	0.3	0.5	0.7	0.9
λ	0.255202	0.272696	0.290190	0.307683	0.325177
B	2.74553	4.63969	6.53386	8.42802	10.32219
$\lVert\mu\rVert_{BV}$	3.68628	6.37931	9.20508	12.17366	15.29615

- The contraction rate and the errors in the L^1 norm are

p_1	0.1	0.3	0.5	0.7	0.9
N (contraction rate)	8	7	7	8	9
L^1 error	0.00180	0.00272	0.00393	0.00594	0.00840
N (*a priori* c. rate)	34	222	2135	314	37
a priori L^1 error	0.00766	0.0865	1.200	0.233	0.0345

Expanding IFS - Example of a rigorous computation (2/2)



a priori estimation of the contraction rate

In the systems considered obtained from the two maps T_0, T_1 and with corresponding operators L_0, L_1, working with a given norm $\|\cdot\|$, we have that:

- Any sequence of applications $L^\omega = L_{\omega_1} L_{\omega_2} \dots L_{\omega_k}$ has uniformly bounded norm $\|L^\omega\| \leq C$, for each sequence $\omega \in \{0,1\}^k$ of any length k k.
- L_0, L_1 are contractions, and we can assume that $\|L_0^{n_0}\| \leq \frac{1}{2C}$ and $\|L_1^{n_1}\| \leq \frac{1}{2C}$.

Theorem 4 (Galatolo, M., Nisoli)

For each $p \in [0,1]$, and putting $N = \max\{n_0, n_1\}$, then

$$\left\| (pL_0 + (1-p)L_1)^M \right\| < \frac{1}{2}, \quad M \geq N - 1 + N \frac{\log 2C}{-\log\left(1 - \frac{p^{n_0}}{2} - \frac{(1-p)^{n_1}}{2}\right)}.$$

Sketch of proof

- The contraction rate of L^n can be estimated expanding $L = pL_0 + (1-p)L_1$, and considering all the weighted terms $L^\omega = L_{\omega_1} L_{\omega_2} \dots L_{\omega_k}$ appearing in the expansion, for a certain n.
- Increasing the length n, we can estimate the contraction rate with a linear recurrence depending on the contraction rates of the previous n.
- The linear recurrence has order $N = \min\{n_0, n_1\}$, and the characteristic polynomial is of the form

$$X^N - p_{N-1}X^{N-1} - \cdots - p_1 X - p_0.$$

- The p_i are positive and have sum slightly smaller than 1, so we can prove that the biggest real root α has absolute value < 1.
- We obtain that L^n has contraction rate $\leq K\alpha^n$ for some K, and estimating K and α we can predict when $\|L^n\| \leq 1/2$.

Contracting IFS - Kantorovich-Wasserstein distance

- In the case of a IFS formed by contracting maps, we can apply the same strategy, but the functional spaces need to be completely different, because for a contraction T in $\mathbb{R}^n$ the corresponding L_T is not a contraction in L^p or BV.
- A space with this property is the dual of Lipschitz, that is the measures for which

$$\|\mu\|_W = \sup_{\phi \in C^0(X): Lip(\phi) \leq 1} \int_X \phi d\mu,$$

  is finite. Such a distance is also known as *Kantorovich-Wasserstein distance*, or *earth-moving distance*, well known in Transportation Theory.
- Such a norm is only defined for μ having zero average, but this is sufficient for us.
- If T contracts by α at least, then we have $\|L_T\|_W \leq \alpha$.

Contracting IFS - Discretization

- We will work assuming that X is a bounded domain in $\mathbb{R}^n$, equipped with the *Manhattan* distance (L^1 distance on the coordinates).
- Given a rectangular lattice of δ-spaced points p_i, the projection is given by

$$\pi_\delta(\mu) = \sum_i \left(\int h_{p_i} \mathrm{d}\mu \right) \cdot \delta_{p_i}$$

where h_{p_i} is a certain *hat function* centered in p_i.

Proposition 1 (Galatolo, M., Nisoli)

If $\|\mu\|_W \leq 1$, then $\|\pi_\delta \mu\| \leq 1$.

Proposition 2 (Galatolo, M., Nisoli)

Putting $L_\delta = \pi_\delta L \pi_\delta$, we have

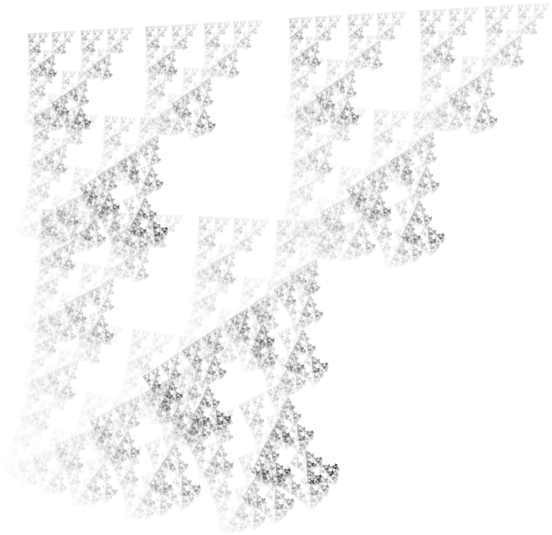
$$\|L - L_\delta\|_{L^1 \to W} \leq (\alpha + 1)\frac{n\delta}{2}.$$

Contracting IFS - Example of a rigorous computation (1/2)

- Computationally this case is easier, because the contraction rate is already know, and as a consequence of the approximation theorem we have $\|\mu - \mu_\delta\|_W \leq \frac{(1+\alpha)n\delta}{2(1-\alpha)}$.
- Let's consider the transformations $T_1, \ldots, T_4$ of the square $X = [0,1] \times [0,1]$ defined as
  - $T1$: scaling by 0.4 around $(0.6, 0.2)$ with rotation of $\pi/6$,
  - $T2$: scaling by 0.6 around $(0.05, 0.2)$ with rotation of $-\pi/30$,
  - $T3$: scaling by 0.5 around $(0.95, 0.95)$,
  - $T4$: scaling by 0.45 around $(0.1, 0.9)$.
- Let's take probabilities $p_1 = 0.18$, $p_2 = 0.22$, $p_3 = 0.3$, $p_4 = 0.3$, and a lattice of $2^{10} \times 2^{10}$ points, with $\delta = 2^{-10}$. The contraction rate α is ≤ 0.659430, and the error (in the $\|\cdot\|_W$ norm) can be estimated as

$$\|\mu - \mu_\delta\|_W \leq \frac{(1+\alpha)n\delta}{2(1-\alpha)} \leq 0.0047583.$$

Contracting IFS - Example of a rigorous computation (2/2)



Notes on the implementation

- Our framework is written in Python and uses the libraries from the computer algebra system Sage.
- A matrix approximating L_δ is computed with certified error using interval arithmetics, and interval Newton method for computing the Ulam approximation.
- The computationally intensive part is implemented via a program using OpenCL for computing on the GPU.
- In the contractive case, we can restrict the computation to a subset of the grid containing the attractor (on the line of what was explained by Kathrin Padberg-Gehle yesterday).

Convergence to the equilibrium

- Assume L to satisfy the inequality

$$\|L^n f\|_s \leq A\lambda_1^n \|f\|_s + B\|f\|_w.$$

- Let L_δ be an approximation satisfying

$$\|(L_\delta^n - L^n)f\|_w \leq \delta(C\|f\|_s + nD\|f\|_w).$$

- Assume that L preserves V, and $\|(L_\delta|_V)^{n_1}\| \leq \lambda_2$.

Theorem 5 (Galatolo, Nisoli, Saussol)

$$\begin{pmatrix} \|L^{in_1}(g)\|_s \\ \|L^{in_1}(g)\|_w \end{pmatrix} \preceq M \cdot \begin{pmatrix} \|g\|_s \\ \|g\|_w \end{pmatrix}, \text{ for } M = \begin{pmatrix} A\lambda_1^{n_1} & B \\ \delta C & \delta n_1 D + \lambda_2 \end{pmatrix},$$

for each $g \in V$, and in particular if ρ is the biggest eigenvalue of M then

$$\|L^{in_1} g\|_s \leq \frac{\rho^i}{a}\|g\|_s, \quad \|L^{in_1} g\|_w \leq \frac{\rho^i}{b}\|g\|_s \quad \text{for explicit } a,\ b.$$