Mini-course PUC-UEFRJ: Limit theorems in dynamical systems using transfer operator methods

Douglas Coates

June 20, 2023

Abstract

In this mini-course we will discuss the so called Nagaev spectral method: a method which exploits properties of the Ruelle-Perron-Frobienius transfer operator in order to establish limit theorems (such as the central limit theorem) for dynamical systems. The course will consist of three sessions. The first two sessions will be devoted to motivating the problem, introducing the required mathematical tools and discussing some first properties of the transfer operator. In the final session we will give explicit examples of proving the central limit theorem using the Nagaev method.

Contents

Disclaimer These are rough notes that likely contain some errors. One can find a more comprehensive review of the topic in [Sar20; Gou15] and the references therein.

1 Session 1: Introduction

1.1 The Central Limit Theorem for iid random variables

Setup

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. Recall that random variables $X, Y : \Omega \to \mathbb{R}$ are ***identically distributed*** if their ***distributions*** $\mathbb{P}_X(A) \coloneqq X_*\mathbb{P} = \mathbb{P}(X \in A)$ and $\mathbb{P}_Y(A) \coloneqq Y_*\mathbb{P} = \mathbb{P}(Y \in B)$ are equal. Recall that X, Y are independent if their joint distribution $\mathbb{P}_{X,Y}(A, B) \coloneqq \mathbb{P}(X \in A, Y \in B)$ is equal to the product distribution $\mathbb{P}_X(A) \cdot \mathbb{P}(B)$.

The strong law of large numbers

If we take a sequence of *independent and identically distributed* (iid) random variables $X_0, X_1, \ldots$ we know from the *Strong Law of Large numbers* (SLN) that

$$\overline{X}_n := \frac{1}{n}\sum_{k=0}^{n-1} X_k \to \overline{X} = \int X_0 \, d\mathbb{P}, \text{ almost surely.} \tag{1}$$

So, the SLN tells us that if we repeat a random experiment n times independently (e.g. dice roll, coin toss...), then the average outcome over time approaches the average over all possible outcomes of a single experiment. A natural question to ask is whether it is possible to understand how quickly $\overline{X}_n$ approaches $\overline{X}$, and also if we can understand how $\overline{X}_n$ is distributed for large n. The *central limit theorem* (CLT) Wall provide us with a partial answer in the case that $X_k \in L^2$.

The Central Limit Theorem

Recall that a sequence of measures (μ_n) *converges weakly to measure* μ, and write $\mu_n \to^w \mu$ if $\int \varphi \, d\mu_n \to \int \varphi \, d\mu$ for every continuous bounded function $\varphi : \Omega \to \mathbb{R}$; and recall that a sequence of random variables $X_0, X_1, \ldots : \Omega, \to \mathbb{R}$ converges in distribution to some $Y : \Omega \to \mathbb{R}$, and write $X_n \to^d Y$, if their distributions converge weakly: $(X_n)_*\mathbb{P} \to^w Y_*\mathbb{P}$.

Theorem 1.1 (CLT). *Suppose that $X_0, X_2, \ldots \in L^2$ are iid, then for every $t \in \mathbb{R}$*

$$\frac{\sum_{k-0}^{n-1} X_k - n\overline{X}}{\sqrt{n}} \to^d Z,$$

where $Z \sim \mathcal{N}(0, \sigma^2)$ and $\sigma^2 = \int X_0^2 \, d\mathbb{P}$.

Remark 1.2. Notice that

$$\frac{\sum_{k-0}^{n-1} X_k - n\overline{X}}{\sqrt{n}} = \frac{1}{1/\sqrt{n}}(\overline{X}_n - \overline{X}),$$

and we know $\overline{X}_n - \overline{X} \to^{a.s.} 0$ thus $1/\sqrt{n}$, in some sense, captures the rate of convergence[1]. Moreover, the CLT tells us that the $\overline{X}_n$ is close to the normal distribution $\mathcal{N}(0, \sigma^2/\sqrt{n})$

Characteristic functions

Later, we will give a simple proof of the central limit theorem. The proof will make use of the properties of the *characteristic function* (or Fourier transform) of a random variable. The *characteristic function* (CF) of a random variable X is given by[2]

$$\phi_X : \mathbb{R} \to \mathbb{C}, \quad \phi_X(t) := \int e^{itX} \, d\mathbb{P} = \int e^{itx} \, dX_*\mathbb{P}.$$

Lemma 1.3. *1. The CF of a (real) random variable always exists*

2. If a random variable X admits a density f_X (i.e. the Radon-Nikodym derivative $\mathrm{d}f_X/\mathrm{d}\,\mathrm{Leb}$) exists then, the CF is the Fourier transform of f_X

[1]More formally, If $\frac{1/\sqrt{n}}{a_n} \to 0$ then $\frac{\overline{X}_n - \overline{X}}{a_n} \to^{\mathbb{P}} 0$, so in fact $1/\sqrt{n}$ describes speed that $\overline{X}_n - \overline{X}$ converges to 0 in measure.

[2]In the last inequality we make use of the change of variables formula: If $T : \Omega \to \Omega'$ is a measurable map between measrurable spaces then

3. *The CF of X uniquely determines $X_*\mathbb{P}$, (there exists some inversion formula)*

4. *If $X_0, X_1, \dots, X_{n-1}$ are independent then the CF of the sum is equal to the product of the CFs:*

$$\int e^{it\sum_{k=0}^{n-1}}\, d\mathbb{P} = \prod_{k=0}^{n-1} \int e^{itX_k}\, d\mathbb{P}.$$

Proof. 1. $x \mapsto e^{itx}$ is bounded and so integrable.

2. Follows from the definition.

3. One can check that if μ_X is the distribution of a random variable X and if φ_X is its distribution then

$$\mu_X(x_1, x_2) + \frac{1}{2}\mu_X(\{x_1\}) + \frac{1}{2}\mu_X(\{x_2\}) = \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-itx_1} - e^{-itx_2}}{it} \varphi_X(t)\, dt.$$

4. **Exercise**.

□

Example 1.4. If $Z \sim \mathcal{N}(0, \sigma^2)$ then

$$\int e^{itZ}\, d\mathbb{P} = e^{-\frac{1}{2}\sigma^2 t^2}.$$

Another important property of CFs is that they characterise converge in distribution.

Theorem 1.5 (Levy's continuity thoerem [Bil95, Theorem 26.3]). *A sequence of random varaibles $X_0, X_1, \dots : \Omega \to \mathbb{R}$ converges in distribution if and only if the corresponding sequence ϕ_{X_n} converges pointwise. In particular,*

$$X_n \to^d X \Leftrightarrow \int e^{itX_n}\, d\mathbb{P} \to \int e^{itX}\, d\mathbb{P} \text{ for every } t \in \mathbb{R}.$$

A proof of the CLT

Proof of Theorem 1.1. We will assume that $\int X_0\, d\mathbb{P} = 0$, the general case will then follow from considering $\tilde{X}_k = X_k - \int X_0\, d\mathbb{P}$. Let $Z_n = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} X_j$. As the X_k are independent, 4 of Lemma 1.3 gives

$$\int e^{itZ_n}\, d\mathbb{P} = \phi(t)^n \tag{2}$$

where $\phi(t) = \int e^{itX_0}\, d\mathbb{P}$. Since $X_0 \in L^2$, we have[3]

$$\phi(t) = \int \sum_{k=0}^{\infty} \frac{(itX_0)^k}{k!}\, d\mathbb{P} = 1 + it \int X_0\, d\mathbb{P} - t^2 \int X^2\, d\mathbb{P} + o(t^2) \quad (\text{as } t \to 0). \tag{3}$$

Combining (2) and (3) we find that

$$\int e^{itZ_n/\sqrt{n}}\, d\mathbb{P} = (\phi(t/sqrtn)^n = \left(1 - \frac{t^2}{2n}\sigma^2 + o\left(\frac{t^2}{n}\right)\right)^n \to e^{-\sigma^2 t^2/2}.$$

Theorem 1.5 together with 1.4 concludes the proof. □

[3]We write $f(t) = o(g(t))$ as $t \to c$ if $\lim_{t\to c} f(t)/g(t) = 0$, we will drop the "as $t \to c$" if it is clear from the context.

Important points form the proof

1. We used independence to write the CF of the sum as $\phi(t)^n$ for some ϕ
2. We found an asymptotic expansion of ϕ near 0
3. Used the continuity theorem to conclude

The deterministic case

Even though the above proof does not work for if X_k is a deterministic process it is very natural to still expect a central limit theorem to hold if the process is sufficiently chaotic. For example, the outcome of a coin toss is not random, just very sensitive to initial conditions, and experimentally we can see that successive coin tosses will satisfy the central limit theorem.

Let us suppose from now on that $(\Omega, \mathcal{B}, \text{Leb})$ is a ***standard probability space*** and that $f : \Omega \to \Omega$ is ***non-singular*** ($\text{Leb}(A) = 0 \Leftrightarrow \text{Leb}(f^{-1}A) = 0$) transformation. For every ***observable*** $\varphi : \Omega \to \mathbb{R}$ (i.e. real valued measurable function) we can consider the sequence $(\varphi \circ f^k)_{k\geq 0}$ and try to understand the statistical behaviour as we did with the iid sequence $X_0, X_1, \ldots$. The sequence $(\varphi \circ f^k)_{k\geq 0}$ is not iid, however:

- if μ is an invariant measure for f then the sequence $(\varphi \circ f^k)_{k\geq 0}$ is identically distributed (but not independent)
- if μ is invariant, ergodic and finite, then the strong law of large numbers holds: $\frac{1}{n}\sum_{k=0}^{n-1} \varphi \circ f^k \to^{a.s.} \int \varphi d\mu$
- if μ is mixing for f there is some form asymptotic independence: the events $A, f^{-n}B$ become independent over time, $\mu(f^{-n}A \cap B) \to \mu(A)\mu(B)$.

We will show that in certain cases it is indeed possible to mimic the proof of the CLT given above for sufficiently chaotic dynamical systems.

1.2 Transfer operators

The main tool in proving a deterministic version of the central limit theorem will be the *Ruelle-Frobenius-Perron operator* or simply the *transfer operator* corresponding to a dynamical system.

Definition of The transfer operator

The transfer operator $\mathcal{L} : L^1 \to L^1$ is the operator which describes how densities evolve under the dynamics of f.

Question: Suppose that we distribute mas on Ω according to some density $h \in L^1$. How will mass be distributed after the space is transformed by f?

Well, our initial (signed) measure is $\mu_h = d\,\text{Leb}$, and so the measure after time 1 will be $f_*\mu_h$. Thus, we define

$$\mathcal{L}(h) := \frac{df_*\mu_h}{d\,\text{Leb}}.$$

2 Session 2: Transfer operators and spectral gap

2.1 Transfer operators continued

Basic properties

Lemma 2.1. *If* $h \in L^1$ *then* $\mathcal{L}(h)$ *is the unique element of* L^1 *such that*

$$\int \varphi \cdot \mathcal{L}h \, d\,\mathrm{Leb} = \int \varphi \circ f \cdot h \, d\,\mathrm{Leb},$$

for all $\varphi \in L^\infty$.

Proof. By definition

$$\int \varphi \cdot \mathcal{L}h \, d\,\mathrm{Leb} = \int \varphi \, df_*\mu_h = \int \varphi \circ f \, d\mu_h = \int \varphi \circ f \cdot h \, d\,\mathrm{Leb}\,.$$

Now, suppose that there exists another element $g \in L^\infty$ which satisfies the same identity. Then

$$\int \varphi \cdot (\mathcal{L}h - g) \, d\,\mathrm{Leb} = 0,$$

for every $\varphi \in L^\infty$. Taking $\varphi = \mathrm{sgn}(\mathcal{L}h - g)$ in the above we find that $\int |\mathcal{L}h - g| \, d\,\mathrm{Leb} = 0$, and so $\mathcal{L}h$ and g must be almost everywhere equal. □

Lemma 2.2. $\mathcal{L} : L^1 \to L^1$ *is a positive, bounded linear operator with norm equal to 1.*

Proof.

1. Linearity follows from the definition.
2. Positive: Suppose that $h \in L^1$ is almost everywhere positive, and so defines a positive measure μ_h. The measure $f_*\mu_h$ will also only assign positive measure to any Borel set. As $\mathcal{L}(h)$ is by definition the density of $f_*\mu_h$ it must be also almost everywhere positive as other wise there would be some A with $\mathrm{Leb}(A) > 0$ such that $f_*\mu_h(A) < 0$.
3. Bounded: Use: To see that $\mathcal{L}$ is bounded we can use Lemma 2.1 to obtain

$$\|\mathcal{L}h\|_1 = \int |\mathcal{L}h| \, d\Psi = \int \mathrm{sgn}(\mathcal{L}h) \cdot \mathcal{L}h \, d\,\mathrm{Leb} = \int \mathrm{sgn}(\mathcal{L}h) \circ f \cdot h \, d\,\mathrm{Leb} \leq \|h\|_1. \tag{4}$$

   Thus, the transfer operator is a weak contraction: $\|\mathcal{L}\|_1 \leq 1$.
4. Norm 1: take $h > 0$ in (4) above and use the fact that $\mathcal{L}$ is positive.

□

Definition 2.3. We will say that $f : [0,1] \to [0,1]$ is a *Markov map* if there exists a partition $\mathcal{P}$ of $[0,1]$ into sub-intervals I_n such that

1. $f : I_n \to f(I_n)$ is bijective
2. $f(I_n)$ is a union of elements in P
3. the sigma algebra $\sigma\{f^{-k}I_n : I_n \in \mathcal{P}\} = \mathcal{B}$ (up to sets of measure zero)

Example 2.4. Suppose that $f : [0,1] \to [0,1]$ is Markov and that f is C^1 on each of its partition elements. Then, we can compute the transfer operator explicitly. Let $h \in L^1$ and $\psi \in L^\infty$ and write $f_n := f|_{I_n}$, then

$$\begin{aligned}\int h \cdot \psi \circ f \, d\,\mathrm{Leb} &= \sum_n \int_{I_n} h \cdot \psi \circ f \, d\,\mathrm{Leb} \\ &= \sum_n \int_{f(I_n)} h \circ f_n^{-1} \cdot \psi \circ f \circ f_n^{-1} \cdot (f_n^{-1})' \, d\,\mathrm{Leb} \\ &= \int \left(\sum_n \left[\frac{1}{f'} 1_{I_n} h \right] \circ f_n^{-1} \right) \cdot \psi \, d\,\mathrm{Leb} \,.\end{aligned}$$

As $h \in L^1$ and $\psi \in L^\infty$ were arbitrary we know from Lemma 2.1 that

$$\mathcal{L}h = \sum_n \left[\frac{1}{f'} 1_{I_n} h \right] \circ f_n^{-1}$$

Example 2.5. Suppose that $f(x) = 2x \mod 1$, then our observation above says that the corresponding transfer operator will be given by

$$\mathcal{L}h(x) = \frac{1}{2} h(x/2) + \frac{1}{2} h((x+1)/2).$$

Dynamical properties of transfer operators

Note from the definition of $\mathcal{L}$ that a function $h \geq 0$ in L^1 is the density of an absolutely continuous invariant measure if and only if $\langle = h$, i.e. h is an eigenvector corresponding to the eigenvalue 1. Moreover, we have the following result.

Proposition 2.6. *If* 1 *is a simple eigenvalue (i.e. the dimension of the corresponding eigenspace if* 1*) then* f *preserves an ergodic absolutely continuous measure.*

Proof. First we show that if 1 is a simple eigenvalue then f has an absolutely continuous invariant measure. It is enough to show that there exists a **non-negative** h such that $\mathcal{L}h = h$. As the eigenspace $\Lambda(1)$ is 1 dimensional we know there exists a fixed point h of $\mathcal{L}$. Write $h = h^+ - h^-$ where h^+, h^- are the positive and negative parts of h. Then, by the linearity and positivity of $\mathcal{L}$ (Lemma 2.2) we find that $\mathcal{L}h^\pm = h^\pm$. If h^+, h^- are linearly independent then we obtain a contradiction to $\dim \Lambda(1) = 1$; so, $\Lambda(1)$ must be the span of a non-negative function.

Now, let us fix $h \geq 0$ in $\Lambda(1)$ and suppose for contradiction that $\mu = h\,dm$ is ***not*** ergodic. Let A be an f invariant set with $\mu(A) \in (0,1)$ and let $B = A^c$. The measures $\mu_A := (\cdot \cap A)$ and $\mu_B := \mu(\cdot \cap B)$ are f invariant and their densities $h 1_A$, $h_B 1_B$ are linearly independent. This contradicts the fact that $\dim \Lambda(1) = 1$. □

Recall, a sequence φ_n converges weakly to φ in L^1 if

$$\int \varphi_n \psi \, d\,\mathrm{Leb} \to \int \varphi \psi \, d\,\mathrm{Leb} \,.$$

We state the following proposition and leave the proof as an exercise.

Proposition 2.7. *If there exist a $h \geq 0$ such that $\mathcal{L}^n(\varphi) \to h \int \varphi d\,\mathrm{Leb}$ weakly in L^1 for every $\varphi \in L^1$ then*

1. *f preserves the measure μ_h,*
2. *f is mixing with respect to μ_h (and in particular ergodic).*

2.2 Spectral gap

We have seen that the iterates of $\mathcal{L}$ determine some interesting dynamical properties.

Definition of spectral gap

Definition 2.8. For an operator $\mathcal{T}$ acting on a Banach space $\mathcal{B}$ we define the spectrum of $\mathcal{T}$ to be the set

$$\sigma(\mathcal{T}) = \{\lambda \in \mathcal{C} : (\mathcal{T} - \lambda\,\mathrm{Id}) \text{ has no bounded inverse.}\}$$

The spectral radius of $\rho(\mathcal{T}) = \sup_{\lambda \in \sigma(\mathcal{T})} |\lambda|$.

Definition 2.9 (spectral gap). We say that a bounded linear operator $\mathcal{T} : \mathcal{B} \to \mathcal{B}$ acting on a Banach space $(\mathcal{B}, \|\cdot\|)$ has a *spectral gap* if

$$\mathcal{T} = \lambda P + N, \tag{5}$$

where P is a projection onto a 1-dimensional subspace of $\mathcal{B}$, N is a bounded linear operator with spectral radius $\rho(N) < |\lambda|$ and $NP = PN = 0$.

Lemma 2.10. *[Sar20] We also note that as the name suggests $\mathcal{T}$ having a spectral gap implies that the spectrum $\sigma(\mathcal{T})$ of $\mathcal{T}$ consists of simple isolated eigenvalue at λ with remaining eigenvalues lying within a disc of radius strictly smaller that $|\lambda|$:*

$$\sigma(\mathcal{T}) = \{\lambda\} \cup A, \text{ where } \exists \gamma > 0 \text{ such that } A \subset \{z \in \mathbb{C} : |z| \leq e^{-\gamma}|\lambda|\}.$$

Moreover,

- *P is the projection onto the eigenspace corresponding to λ (i.e. $\mathrm{Im}(P) = \{\varphi \in \mathcal{B} : \mathcal{T}\varphi = \lambda\varphi\}$).*
- *So λ is a simple eigenvalue.*

Idea

- if an operator has a spectral gap we can say something about its iterates
- the iterates of the transfer operator tell us interesting things about the dynamics.
- problem, most systems do not have a spectral gap on $\mathcal{L}^1$

2.3 Gibbs-Markov maps

In this section we introduce the notion of a Gibbs-Markov map. For a detailed discussion of Gibbs-Markov maps see [Aar97, Chapter 4].

Definition 2.11 (Markov map). A non-singular map $f : \Omega \to \Omega$ of a standard probability space is *Markov* if there exits a measurable and at most countable partition $\mathcal{P}$ of Ω such that

1. $f : a \to f(a)$ is a bijection for every $a \in \mathcal{P}$
2. $f(a) \in \sigma(\mathcal{P})$
3. the partition $\mathcal{P}$ is *generating*: $\sigma\left(\bigvee_{k=0}^{\infty} f^{-n}(\mathcal{P}) = \mathcal{B}\right.$ (up to sets of zero Lebesgue measure).

Given a Markov map f with partition $\mathcal{P}$ there is natural measure of distance on the space Ω which comes from the notion of the ***separation time*** $s(x,y)$. Given tow points $x,y \in \Omega$ we define the separation time to be the smallest amount of time for two distinct points to lie in different elements of $\mathcal{P}$

$$s(x,y) \coloneqq \min\{n \geq 0 : f^n x, f^n y \text{ lie in different elements of } \mathcal{P}\}. \tag{6}$$

Then for $\theta \in (0,1)$ we may define the distance d_θ by putting

$$d_\theta(x,y) \coloneqq \theta^{s(x,y)}. \tag{7}$$

We note that the space (Ω, d_θ) is Polish and f is Lipschitz with respect to d_θ[4].

For a function $\varphi : \Omega \to \mathbb{R}$ and a partition element $a \in \mathcal{P}$ we denote by $D_\theta(\varphi)$ the least Lipschitz constant of $\varphi|_a$ with respect to the distance d_θ:

$$D_\theta(\varphi)(a) \coloneqq \sup_{x,y\in q} \frac{|\varphi(x) - \varphi(y)|}{d_\theta(x,y)}.$$

We define the semi-norm

$$|\varphi|_\theta \coloneqq \sup_{n\in\mathcal{P}} D_{\theta,n}(\varphi). \tag{8}$$

If $|\varphi|_\theta < \infty$ we say that φ is ***locally θ-Hölder***. We note that locally θ-Hölder functions may be unbounded. Let B_θ be the space of bounded locally θ-Hölder functions

$$L_\theta \coloneqq \{\varphi : \Omega \to \mathbb{R} : \|\varphi\|_\theta \coloneqq \|\varphi\|_{L^\infty(m)} + |\varphi|_\theta < \infty\}, \tag{9}$$

and remark that $(B_\theta, \|\cdot\|_\theta)$ forms a Banach space. By definition a Markov map is invertible on each partition element. Denoting by $\varphi_a : T^n a \to a$ the inverse of T^n on $a \in \mathcal{P}_n \coloneqq \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}$ we let φ_a' be the Radon-Nikodym derivatives

$$\varphi_a' \coloneqq \frac{dm \circ \varphi_a}{dm}.$$

Definition 2.12 (Gibbs-Markov). If T is Markov with partition $\mathcal{P}$ then we say that the tuple $(\Omega, \mathscr{B}, m, T, \mathcal{P})$ is ***Gibbs-Markov*** if two additional properties are satisfied

1. **big images**:
$$\inf_{q\in\mathcal{P}} m(Tq) > 0,$$
2. θ-**distortion**: there exists a $\theta \in (0,1)$ and there exists a $C > 0$ so that for all $n \geq 0$, all $a \in \mathcal{P}_n$ and almost every $x,y \in a$ we have that
$$\left|\frac{\varphi_a'(x)}{\varphi_a'(y)} - 1\right| \leq C d_\theta(x,y).$$

Exercise 1:

Check that the following are Gibbs-Markov maps of the interval

[4] To see that (Ω, d_θ) is Polish one quickly verifies that if $x_n \to x$ in (Ω, d_θ) then $x_n \to x$ in (Ω, d) and moreover one can easily check that sequences which are Cauchy in (Ω, d_θ) are also Cauchy in (Ω, d). To see that T is Lipschitz one simple observes that $d_\theta(Tx, Ty) = \theta^{s(Tx,Ty)} \leq \theta^{s(x,y)-1}$.

- doubling map, $x \mapsto 2x \mod 1$
- Gauss map, $x \mapsto 1/x \mod 1$

Theorem 2.13. *[AD01] If f is a topologically mixing Gibbs-Markov map then there exists a $\theta \in (0,1)$ such that the corresponding transfer operator $\mathcal{L}$ acts on the space B_θ. Moreover, $\mathcal{L}$ has a spectral gap with $\lambda = 1$:*

$$\mathcal{L}|_{B_\theta}\varphi = P + N,$$

and the projection P is given by $P\varphi = h\int \varphi \, d\,\mathrm{Leb}$ for some $h \geq 0$.

3 Session 3: Central limit theorem for Gibbs-Markov maps

3.1 Dynamical interpretation of a spectral gap

Suppose that f is a Gibbs-Markov map of a standard probability space, then we know that from Theorem 2.13 that the transfer operator corresponding to f has a spectral gap and moreover

- 1 is a simple isolated eigenvalue of $\mathcal{L}$
- $h \geq 0$ is a fixed point (put $P1 = h$ and compute)
- this means that there exists a unique $\mu \ll \mathrm{Leb}$ which is ergodic for f with density h (recall Proposition 2.6).

We can also use the fact that $\mathcal{L}$ has a spectral gap in order to compute the iterates of $\mathcal{L}$,

$$\mathcal{L}^n = P^n + N^n$$

so $\|\mathcal{L} - P^n\| = \|N^n\| \leq (\rho(N) + \varepsilon)^n < \gamma^n$ for some $\gamma < 1$ by the spectral radius formula.

Note that $\|\varphi\| \geq \|\varphi\|_{L^1}$ so for every $\varphi \in B_\theta$

$$\left\| \mathcal{L}^n\varphi - h\int \varphi \, d\,\mathrm{Leb} \right\|_{L^1} \leq \left\| \mathcal{L}^n\varphi - h\int \varphi \, d\,\mathrm{Leb} \right\| < \gamma^n \to 0$$

exponentially fast. Recall that before we showed that $\mathcal{L}^n\varphi \to h\int \varphi \, d\,\mathrm{Leb}$ weakly in L^1 for every φ implies mixing. We do not quite have this B_θ does not contain the indicator functions.

However, we do have exponential decay of correlations: let $\psi \in L^\infty$ and $\varphi \in B_\theta$, then

$$\begin{aligned}
|\operatorname{Cor}(\varphi, \psi \circ f^n)| &= \left| \int \psi \circ f^n \cdot \varphi \, d\mu - \int \psi \, d\mu \int \varphi d\mu \right| \\
&= \left| \int \psi \mathcal{L}^n(\varphi h) \, d\,\mathrm{Leb} - \int \psi h \left(\int \varphi h \, d\,\mathrm{Leb} \right) \, d\,\mathrm{Leb} \right| \\
&\leq \left| \int \psi \left(\mathcal{L}^n(\varphi h) - h \int \varphi h \, d\,\mathrm{Leb} \right) \, d\,\mathrm{Leb} \right| \\
&\leq \|\psi\|_\infty \left\| \mathcal{L}^n(\varphi h) - h \int \varphi h \, d\,\mathrm{Leb} \right\|_{L^1} \leq \gamma^n \|\psi\|_\infty
\end{aligned}$$

3.2 Central limit theorem for Gibbs-Markov maps

Theorem 3.1. *Suppose that $(f, \Omega, \mathcal{B}, m, \mathcal{P})$ is a Gibbs-Markov map with θ-distortion. Let $\varphi : \Omega \to \mathbb{R}$ is bounded. Then a central limit theorem holds for φ*

$$\frac{\sum_{k=0}^{n-1} \varphi \circ f^k(x) - n \int \varphi d\mu}{\sqrt{n}} \to^d Z,$$

where:

- $Z \sim \mathcal{N}(0, \sigma^2)$ *for some* $\sigma^2 \geq 0$
- *the convergence in distribution occurs with respect to the acip* $\mu = h\, dm$ *(which exists by Theorem 2.13*

3.3 Proof of the CLT for Gibbs-Markov maps

Let us assume for the time being that $\int \varphi d\mu = 0$ and let us write $\varphi_n := \sum_{k=0}^{n-1} \varphi \circ f^k$. Recall: Levy's continuity theorem told us that to prove convergence in distribution it is in fact enough to show pointwise convergence of the characteristic functions. So, to prove the theorem it is enough to show that

$$\int \exp\left\{ it \frac{1}{\sqrt{n}} \varphi_n \right\} d\mu \to \exp\left\{ -\frac{1}{2} t^2 \sigma^2 \right\}, \quad \forall t \in \mathbb{R}.$$

To ease notation let us set

$$\Phi(t) := \int \exp\{it\varphi\}\, d\mu, \quad \text{and} \quad \Phi_n(t) := \int \exp\{it\varphi_n\}\, d\mu,$$

So, our aim is to show that

$$\Phi_n(t/\sqrt{n}) \to \Psi(t) := \exp\left\{ -\frac{1}{2} t^2 \sigma^2 \right\}, \quad \forall t.$$

If $X_1, X_2, \dots$ are iid random variables on some probability space $(\Omega', \mathcal{F}, \mathbb{P})$ with the same distribution as φ we know that $\int X_1^2\, d\mathbb{P} = \int \varphi^2\, d\mu$ and so Theorem 1.1 ensures that the $X_1, X_2, \dots$ will satisfy a central limit theorem. In particular,

$$\Psi(t) = \lim_{n\to\infty} \int e^{i \frac{t}{\sqrt{n}} \sum_{k=0}^{n-1} X_k}\, d\mathbb{P} = \lim_{n\to\infty} \left[\int e^{i \frac{t}{\sqrt{n}} X_1}\, d\mathbb{P} \right]^n = \lim_{n\to\infty} \Phi(t/\sqrt{n})^n. \tag{10}$$

The characteristic function operator

Given $t \in \mathbb{R}$ we define the *characteristic function operator* $\mathcal{L}_t : B_\theta \to B_\theta$ by setting

$$\mathcal{L}_t \psi := \mathcal{L}\left(e^{it\varphi} \psi \right).$$

Lemma 3.2. *The operator $\mathcal{L}_t$ acts on B_θ. Moreover, there exists a $C > 0$ such that*

$$\|\mathcal{L}_t - \mathcal{L}\| \leq C|t|. \tag{11}$$

Proof. Throughout the proof we will make repeated use of the following inequality

Sublemma 3.3. *There exists a $C > 0$ such that for any $a \in \mathbb{R}$*

$$|e^{ita} - 1| < C|a|. \tag{12}$$

Let $\psi \in B_\theta$ with $\|\psi\| < 1$.

By definition, $(\mathcal{L}_t - \mathcal{L})(\psi) = \psi(e^{it\varphi} - 1)$

Thus,

$$\|(\mathcal{L}_t - \mathcal{L})(\psi)\|_\infty \le \|\psi\|_\infty \|\varphi\|_\infty C|t| < C|t|.$$

Similarly

$$\begin{aligned} |(\mathcal{L}_t - \mathcal{L})(\psi)(x) - (\mathcal{L}_t - \mathcal{L})(\psi)(y)| &\le |\psi(x) - \psi(y)||e^{it\varphi(x)} - 1| + |\psi(y)||e^{it\varphi(y)}||e^{it\varphi(x) - \varphi(y)} - 1| \\ &\le C\|\varphi\|_\infty |\psi|_\theta |t| d_\theta(x,y) + C\|\psi\|_\infty |t||\varphi|_\theta d_\theta(x,y) \end{aligned}$$

and so

$$|(\mathcal{L}_t - \mathcal{L})(\psi)|_{B_\theta} < C|t|.$$

Finally, $\|\mathcal{L}_t\| \le C|t| + \|\mathcal{L}\|$. □

Basic properties of the CF operator

Notice that as $e^{it\varphi} \in L^\infty$

$$\int \mathcal{L}_t(h)\, d\mu = \int \mathcal{L}\left(e^{it\varphi} h\right)\, d\mu = \int e^{it\varphi} h dm = \int e^{it\varphi} d\mu = \Phi(t),$$

and moreover for $n \in \mathbb{N}$

$$\begin{aligned} \int \mathcal{L}_t^n(h)\, dm = \int \mathcal{L}\left[e^{it\varphi} h \mathcal{L}_t^{n-1}(h)\right]\, dm &= \int e^{it\varphi} \mathcal{L}_t^{n-1}(h)\, dm \\ &= \int e^{it\varphi} \circ f \cdot e^{it\varphi} \mathcal{L}_t^{n-2}(h)\, dm = \cdots \\ &= \int e^{it\varphi} \circ f^{n-1} \cdot e^{it\varphi} \circ f^{n-2} \cdots e^{it\varphi} h\, dm \\ &= \Phi_n(t) \end{aligned} \tag{13}$$

Perturbation of linear operators

Theorem 3.4 ([Gou15, Proposition 2.3]). *Suppose that $\mathcal{T}_t$ is a family of linear operators acting on a Banach space B and suppose further that $\mathcal{T}_0$ has a spectral gap with decomposition*

$$\mathcal{T}_0 = \lambda P_0 + N_t.$$

If $\|\mathcal{T}_t - \mathcal{T}_0\| < C|t|$ then, for all $|t|$ sufficiently small we have that

- *$\mathcal{T}_t$ has a spectral gap with decomposition $\mathcal{T}_t = \lambda_t P_t + N_t$*
- *$|\lambda_t - \lambda| = O(|t|)$*
- *$\|P_t - P\| = O(|t|)$.*

Thus, Lemma 3.2 yields allows us to use the theorem above to conclude that for all $|t|$ small enough $\mathcal{L}_t$ has a spectral gap with $\mathcal{L}_t = \lambda_t P_t + N_t$ with

$$|\lambda_t - 1| = O(|t|) \tag{14}$$

$$\|P_t \psi - h \int \psi d\mu\| = O(|t|) \text{ for all } \|\psi\| \le 1. \tag{15}$$

Expansion of the dominant eigenvalue λ_t

Proposition 3.5. *The dominant eigenvalue λ_t of $\mathcal{L}_t$ satisfies*

$$\lambda_t = \Phi(t) + O(|t|^2)$$

Proof. Let $h_t = \frac{P_t 1}{\int P_t 1\, dm}$ then

$$\mathcal{L}_t h_t = \lambda_t P_t\left(\frac{P_t 1}{\int P_t 1\, dm}\right) + N_t\left(\frac{P_t 1}{\int P_t 1\, dm}\right) = \lambda_t h_t.$$

So h_t is eigenvector of λ_t with $\int h_t\, dm = 1$. Thus $\int \mathcal{L}h - \mathcal{L}h_t\, dm = \int h\, dm - \int h_t\, dm = 0$ and we can use (14) and (15) to calculate that

$$\lambda_t = \int \mathcal{L}_t h_t\, dm = \int \mathcal{L}_t h\, dm + \int (\mathcal{L}_t - \mathcal{L})(h_t - h)\, dm = \Phi(t) + O(|t|^2).$$

□

Concluding the central limit theorem

Proof. Recall that since $\mathcal{L}_t$ has a spectral gap, we know from the spectral radius formula that there exists an ε such that for all n large enough and for all t

$$\|N_t^n\|_{L^1} \leq \|N_t^n\| \leq |\lambda_t - \varepsilon|^n.$$

So,

$$\frac{1}{\lambda^n_{t/\sqrt{n}}}\left|\int N^n_{t/\sqrt{n}} h\, dm\right| \leq \frac{1}{\lambda^n_{t/\sqrt{n}}}\|N^n_{t/\sqrt{n}}\| < \frac{|\lambda_{t/\sqrt{n}} - \varepsilon|^n}{\lambda^n_{t/\sqrt{n}}} \to 0 \text{ as } n \to \infty.$$

Using (15) we also have

$$\int P_{t/\sqrt{n}} h\, dm = \int Ph\, dm + \int P_{t/\sqrt{n}} h - Ph\, dm = 1 + o(1).$$

Finally, using (13), (10) and Proposition 3.5,

$$\begin{aligned}\Phi_n(t/\sqrt{n}) &= \lambda^n_{t/\sqrt{n}} \int P_{t/\sqrt{n}} h + N^n h\, dm \\ &= \lambda^n_{t/\sqrt{n}}(1 + o(1)) \\ &= (\Phi(t/\sqrt{n})^n(1 + o(1)), \\ &\to \Psi(t),\end{aligned}$$

concluding the proof. □

References

[Aar97] Jon Aaronson. *An introduction to infinite ergodic theory.* Vol. 50. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 1997, pp. xii+284. isbn: 0-8218-0494-4. doi: 10.1090/surv/050. url: https://doi-org.uoelibrary.idm.oclc.org/10.1090/surv/050.

[AD01] Jon Aaronson and Manfred Denker. "Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps". In: *Stoch. Dyn.* 1.2 (2001), pp. 193–237. issn: 0219-4937. doi: 10.1142/S0219493701000114. url: https://doi-org.uoelibrary.idm.oclc.org/10.1142/S0219493701000114.

[Bil95] Patrick Billingsley. *Probability and measure.* Third. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1995, pp. xiv+593. isbn: 0-471-00710-2. url: https://mathscinet-ams-org.uoelibrary.idm.oclc.org/mathscinet-getitem?mr=1324786.

[Gou15] Sébastien Gouëzel. "Limit theorems in dynamical systems using the spectral method". In: *Hyperbolic dynamics, fluctuations and large deviations.* Vol. 89. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, 2015, pp. 161–193. doi: 10.1090/pspum/089/01487. url: https://doi.org/10.1090/pspum/089/01487.

[Sar20] Omri Sarig. *Introduction to the transfer operator method.* Jan. 2020.