Products of random matrices

Jean-Baptiste Boyer (FMJH)

September 7, 2016

Let ρ be a (borelian) probability measure on
$\mathbf{G} := \mathrm{SL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| ad - bc = 1 \right\}$.
For any $x \in \mathbb{R}^2 \setminus \{0\}$ we consider the random walk defined by

$$\begin{cases} X_0 & = & x \\ X_{n+1} & = & g_{n+1} X_n \end{cases}$$

where (g_n) is an iid sequence of law ρ.

Let ρ be a (borelian) probability measure on
$\mathbf{G} := \mathrm{SL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \middle| ad - bc = 1 \right\}$.
For any $x \in \mathbb{R}^2 \setminus \{0\}$ we consider the random walk defined by

$$\begin{cases} X_0 & = & x \\ X_{n+1} & = & g_{n+1} X_n \end{cases}$$

where (g_n) is an iid sequence of law ρ.

If $\rho = \frac{1}{2}\delta_A + \frac{1}{2}\delta_B$ with $A, B \in \mathbf{G}$, then at each step, we toss a coin, if we get heads, we go in Ax and if we get tails, we go in Bx and so on.

In dimension 1, we have that

$$\ln|g_n \dots g_1 x| = \sum_{i=1}^{n} \ln|g_i| + \ln|x|$$

So, under moment assumptions, the law of large numbers proves that $\frac{1}{n}\ln|g_n \dots g_1 x|$ converges to $\int_{\mathbb{R}} \ln|y| \mathrm{d}\rho(y)$.

In dimension 1, we have that

$$\ln |g_n \dots g_1 x| = \sum_{i=1}^{n} \ln |g_i| + \ln |x|$$

So, under moment assumptions, the law of large numbers proves that $\frac{1}{n} \ln |g_n \dots g_1 x|$ converges to $\int_{\mathbb{R}} \ln |y| \mathrm{d}\rho(y)$.

This is why, in dimension 2, we expect $\frac{1}{n} \ln \|X_n\|$ to converge to something.

If the support of ρ is a subset of

$$\left\{ E_b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \middle| b \in \mathbb{R} \right\}$$

then, as

$$E_{b_1} E_{b_2} = E_{b_1+b_2}$$

what we have is a random walk on $\mathbb{R}$ rather than a random walk on $\mathrm{SL}_2(\mathbb{R})$.

If the support of ρ is a subset of

$$\left\{ E_b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \middle| b \in \mathbb{R} \right\}$$

then, as

$$E_{b_1} E_{b_2} = E_{b_1 + b_2}$$

what we have is a random walk on $\mathbb{R}$ rather than a random walk on $\mathrm{SL}_2(\mathbb{R})$.

So, we have to make assumptions on $\mathrm{supp}\rho$.

Definition

We say that a closed subgroup $\mathbf{H} < \mathbf{G}$ is *strongly irreducible* if it doesn't fix any finite union of lines in $\mathbb{R}^2$.

Example

$$\left\{ \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \right\} \bigcup \left\{ \begin{pmatrix} 0 & -a^{-1} \\ a & 0 \end{pmatrix} \right\},$$

are not strongly irreducible (but the last one is irreducible)

If the group Γ_ρ spanned by the support of ρ is compact then $\{gx | g \in \Gamma_\rho\}$ is bounded. So for any $x \in \mathbb{R}^2 \setminus \{0\}$,

$$\frac{1}{n} \ln \|X_n\| \to 0 \text{ a.e.}$$

So, in the sequel, we will assume that Γ_ρ is non-compact.

Lemma

A compact subgroup of $\mathbf{G}$ *is a subgroup of* $\mathrm{SO}(2)$ *(up to conjugacy) and in particular, it is abelian.*

Proof.

A compact subgroup of $\mathbf{G}$ fixes an inner product. □

Remark

If **H** is a non-compact subgroup of **G**, it cannot fix a finite union of 3 lines or more.
Indeed, if we have an unbounded sequence $(h_n) \in \mathbf{H}^{\mathbb{N}}$ we can write $h_n = k_n a_n l_n$ with $k_n, l_n \in \mathrm{SO}(2)$,

$$a_n = \begin{pmatrix} t_n & 0 \\ 0 & t_n^{-1} \end{pmatrix} \text{ and } \lim_n t_n = +\infty$$

We can assume without any loss of generality that $k_n \to k$ and $l_n \to l$.
If $X \in \mathbb{P}(\mathbb{R}^2)$ is such that $lX \notin \mathbb{R}e_2$, then $a_n l_n X \to \mathbb{R}e_1$ and so, $g_n X \to \mathbb{R}ke_1$. So $\mathbb{R}l^{-1}e_2$ and $\mathbb{R}ke_1$ are the only possible fixed lines in $\mathbb{R}^2$.

Definition

We call *non-elementary* any non compact subgroup of **G** that doesn't fix a line or the union of two lines.

Example

If $\rho = \frac{1}{2}\delta_A + \frac{1}{2}\delta_B$ with

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

then $\mathrm{supp}\rho$ generates a non-elementary subgroup : it is non-compact since none of the matrix has eigenvalues of modulus 1 and strongly irreducible since the matrices are diagonalisable and have different eigenvectors.

In the sequel, we will note $\mathbf{G}_\rho$ the closure of the subgroup of $\mathbf{G}$ generated by the support of ρ (it is still a group and it is non-elementary if and only if so is the group generated by the support of ρ)

Theorem (Furstenberg-Kesten, ...)

Let ρ be a borelian probability measure on **G** *such that $\int_{\mathbf{G}} |\ln \|g\|| \mathrm{d}\rho(g)$ is finite and $\mathbf{G}_\rho$ is non-elementary.*
Then, there is $\lambda_1 > 0$ such that

$$\lim_n \frac{1}{n} \ln \|g_n \dots g_1\| = \lim_n \frac{1}{n} \ln \lambda_1(g_n \dots g_1) = \lambda_1 \quad \rho^{\otimes \mathbb{N}}\text{-ae.}$$

where we noted $\lambda_1(g)$ the spectral radius of an element $g \in \mathbf{G}$.
Moreover, for any $x \in \mathbb{R}^2 \setminus \{0\}$,

$$\lim_n \frac{1}{n} \ln \|g_n \dots g_1 x\| = \lambda_1 \quad \rho^{\otimes \mathbb{N}}\text{-ae.}$$

In particular, the walk on $\mathbb{R}^2 \setminus \{0\}$ is transient :

$$\|g_n \dots g_1 x\| = e^{\lambda_1 n + o(n)} \text{ a.e.}$$

Remark

If we add moment assumptions on ρ, we can get the central limit theorem, the law of the iterated logarithm and large deviations inequalities.

Theorem

Let ρ be a probability measure on $\mathrm{SL}_2(\mathbb{R})$ whose support generates a non-elementary subgroup. Assume that for some $\varepsilon \in \mathbb{R}_+^*$,

$$\int_{\mathbf{G}} \|g\|^{\varepsilon} \mathrm{d}\rho(g) < +\infty$$

then, for any $\varepsilon \in \mathbb{R}_+^*$ there is $t \in \mathbb{R}_+^*$ such that for any $n \in \mathbb{N}$ large enough

$$\sup_{x \in \mathbb{R}^2 \setminus \{0\}} \rho^{\otimes n}\left(\left\{ g_1, \dots, g_n \in \mathbf{G} \middle| \left| \frac{1}{n} \ln \frac{\|g_n \dots g_1 x\|}{\|x\|} - \lambda_1 \right| > \varepsilon \right\} \right) \leqslant e^{-tn}$$

Corollary

Let ρ be a borelian probability measure on $\mathbf{G}$ such that $\int_{\mathbf{G}} |\ln \|g\||\mathrm{d}\rho(g)$ is finite and $\mathbf{G}_\rho$ doesn't fix a line or the union of two lines.
If there is $x \in \mathbb{R}^2 \setminus \{0\}$ such that

$$\lim_n \frac{1}{n} \ln \|g_n \dots g_1 x\| = 0 \ ae.$$

Then, $\mathbf{G}_\rho$ is compact. In particular, for any $x \in \mathbb{R}^2$, $\{gx | g \in \mathbf{G}_\rho\}$ is bounded.

Furstenberg and Kesten's studied the behaviour of $\|g_n \dots g_1 x\|$ and we would like to study now the direction :

$$\frac{g_n \dots g_1 x}{\|g_n \dots g_1 x\|}$$

Points x and $-x$ obviously have opposite behaviours so we consider the action on the projective space $\mathbb{P}(\mathbb{R}^2)$ rather than the one on the sphere and so, we are interested in the direction $\mathbb{R} g_n \dots g_1 x$.

It is clear that all the products $g_n \dots g_1 x$ cannot converge to the same limit (otherwise, this would give an invariant line and this contradicts the strong irreducibility assumption) and we can prove that there is no probability measure on $\mathbb{P}(\mathbb{R}^2)$ that is invariant by Γ_ρ. So we are at least looking for measures ν that are invariant on average : for any continuous function f,

$$\int_{\mathbb{P}(\mathbb{R}^2)} \int_{\mathbf{G}} f(gx) \mathrm{d}\rho(g) \mathrm{d}\nu(x) = \int_{\mathbb{P}(\mathbb{R}^2)} f(x) \mathrm{d}\nu(x)$$

Theorem

Let ρ be a borelian probability measure on $\mathbf{G}$ such that $\mathbf{G}_\rho$ is non-elementary. Then,

For any $X, Y \in \mathbb{P}(\mathbb{R}^2)$, with probability 1, we have that

$$\lim_{n \to +\infty} d(g_n \dots g_1 X, g_n \dots g_1 Y) = 0$$

There is a unique $P-$invariant probability measure ν on $\mathbb{P}(\mathbb{R}^2)$, for any $f \in \mathcal{C}^0(\mathbb{P}(\mathbb{R}^2))$ and any $X \in \mathbb{P}(\mathbb{R}^2)$,

$$\left| \mathbb{E} f(g_n \dots g_1 X) - \int_{\mathbb{P}(\mathbb{R}^2)} f(y) \mathrm{d}\nu(y) \right| \to 0$$

This theorem proves that the directions of $g_n \dots g_1 X$ are asymptotically distributed according to the measure ν.

Sketch of proof of the first point.

We use the same kind of ideas that in the proof of the fact that there cannot be a fixed finite union of more than three lines.

According to Furstenberg and Kesten's theorem, a generic element $g_n \dots g_1$ has a large spectral radius. In particular, it is diagonalisable and has an attractive fixed point and a repulsive one in the projective space. So, if we take two points X, Y in the projective space, the set of g such that X or Y is the repulsive point of g has a small measure (because of the strong irreducibility assumption) and so gX and gY are very close to the attractive point of g and in particular, they are close to each-other. □

Sketch of proof of the second point.

The first point proves that for any $X, Y \in \mathbb{P}(\mathbb{R}^2)$,

$$|\mathbb{E}f(g_n \dots g_1 X) - \mathbb{E}f(g_n \dots g_1 Y)| \to 0$$

The fact that the measure ν exists comes from general results about measures on compact sets (Banach-Alaoglu's theorem).
Taking the average against the measure $\mathrm{d}\nu(Y)$ and using the dominated convergence theorem, we have that for any $X \in \mathbb{P}(\mathbb{R}^2)$,

$$\left|\mathbb{E}f(g_n \dots g_1 X) - \int_{\mathbb{P}(\mathbb{R}^2)} \mathbb{E}f(g_n \dots g_1 Y)\mathrm{d}\nu(Y)\right| \to 0$$

Finally, as ν is $P-$invariant,

$$\int_{\mathbb{P}(\mathbb{R}^2)} \mathbb{E}f(g_n \dots g_1 Y)\mathrm{d}\nu(Y) = \int_{\mathbb{P}(\mathbb{R}^2)} f(Y)\mathrm{d}\nu(Y)$$

□

A question is whether the invariant measure has atoms ...

A question is whether the invariant measure has atoms ...
It cannot have atoms because otherwise, there would be a maximal $t \in \mathbb{R}$ such that

$$X_t = \{x \in \mathbf{X} | \nu(\{x\}) = t\}$$

is finite and non empty.
But, ν is invariant, so for any $x \in X_t$,

$$t = \nu(\{x\}) = \int_{\mathbf{G}} \int_{\mathbf{X}} \mathbf{1}_{\{x\}}(gy) \mathrm{d}\nu(y) \mathrm{d}\rho(g) = \int_{\mathbf{G}} \nu(\{g^{-1}x\}) \mathrm{d}\rho(g) \leqslant t$$

So, for $\rho-$a.e. $g \in \mathbf{G}$, $\nu(\{g^{-1}x\}) = t$. So, X_t is a $\mathbf{G}_\rho-$invariant finite union of lines in $\mathbb{R}^2$ and this contradicts the strong irreducibility assumption.

The measure ν is actually very far from having atoms and one can prove that there are C, Δ such that for any $r \in]0,1]$ and any $x \in \mathbb{P}(\mathbb{R}^2)$,

$$\nu(B(x,r)) \leqslant Cr^{\Delta}$$

We studied the behaviour of $\|g_n \dots g_1 x\|$ and of the direction of $g_n \dots g_1 x$. We would like to put this together to describe the behaviour of $g_n \dots g_1 x$. To do so, we take a box B in $\mathbb{R}^2$ that doesn't contain 0 and we study the expectation of the number of times when $g_n \dots g_1 x \in B$. In other words, we study

$$\sum_{n=0}^{+\infty} \mathbb{E}\mathbf{1}_B(g_n \dots g_1 x)$$

We studied the behaviour of $\|g_n \dots g_1 x\|$ and of the direction of $g_n \dots g_1 x$. We would like to put this together to describe the behaviour of $g_n \dots g_1 x$. To do so, we take a box B in $\mathbb{R}^2$ that doesn't contain 0 and we study the expectation of the number of times when $g_n \dots g_1 x \in B$. In other words, we study

$$\sum_{n=0}^{+\infty} \mathbb{E}\mathbf{1}_B(g_n \dots g_1 x)$$

The idea is that when x, x' are two different starting points with x' much closer to 0 than x. Then, it will take a longer time to $g_n \dots g_1 x'$ to reach B than to $g_n \dots g_1 x$ so, when it does reach B, the distribution of the direction of $g_n \dots g_1 x'$ will be close to ν and the distribution of $\|g_n \dots g_1 x'\|$ will be controlled.

First of all, we need to go back from the projective space to the circle. There are actually only two cases

- $\mathbf{G}_\rho$ fixes a non trivial convex cone in $\mathbb{R}^2$ and, in this case, we have twice the situation on the projective space.
- $\mathbf{G}_\rho$ doesn't fix any non trivial convex cone in $\mathbb{R}^2$ and in this case, there is a unique invariant probability measure ν on the circle.

Theorem

Let ρ be a probability measure on $\mathbf{G}$ *such that for some $\varepsilon > 0$,* $\int_{\mathbf{G}} \|g\|^{\varepsilon} \mathrm{d}\rho(g)$ *is finite and* $\mathbf{G}_{\rho}$ *is non elementary and doesn't fix any non-trivial convex cone in $\mathbb{R}^2$.*
Then, for any compactly supported continuous function on $\mathbb{R}^2$ such that for some $\alpha \in]0,1]$, $C \in \mathbb{R}$ and any $x \in \mathbb{R}^2$, $|f(x)| \leqslant C\|x\|^{\alpha}$, we have that

$$\lim_{x \to 0} \sum_{n=0}^{+\infty} \mathbb{E} f(g_n \dots g_1 x) = \frac{1}{\lambda_1} \int_{\mathbb{R}_+} \int_{\mathbf{S}^1} f(uy) \mathrm{d}\nu(y) \frac{\mathrm{d}u}{u}$$

Remark

We cannot avoid the cone assumption.

Lemma (Ping-pong)

Let $\mathbf{G}$ be a group acting on a space $\mathbf{X}$ and $a, b \in \mathbf{G}$. Assume that there are 4 disjoint (non-empty) subsets A^+, A^-, B^+, B^- of $\mathbf{X}$ such that

a *maps* B^+, B^-, A^+ *to* A^+

a^{-1} *maps* B^+, B^-, A^- *to* A^-

b *maps* A^+, A^-, B^+ *to* B^+

b^{-1} *maps* A^+, A^-, B^- *to* B^-

Then a and b generate a free subgroup of $\mathbf{G}$.

Proof.

Let w be a reduced word in a, b and there inverse. Assume it starts with a. Let $p \notin A^{+}$ and not in the repelling set of the last letter of w. Then, $wp \in A^{+}$ so, $wp \neq p$. And this proves that w doesn't act trivially on $\mathbf{X}$ and hence is non trivial in $\mathbf{G}$. □

Proof.

Let w be a reduced word in a, b and there inverse. Assume it starts with a. Let $p \notin A^+$ and not in the repelling set of the last letter of w. Then, $wp \in A^+$ so, $wp \neq p$. And this proves that w doesn't act trivially on $\mathbf{X}$ and hence is non trivial in $\mathbf{G}$. □

Corollary

A non elementary subgroup of $\mathbf{G}$ *contains a free subgroup.*

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We consider the two transformations $(x \mapsto 2x)$ and $(x \mapsto 3x)$ and we would like to know what are the closed invariant subsets of $\mathbb{T}$.
It is clear that 0 is invariant and that the orbit of rational points are finite (so they are closed and by definition, invariant). The circle itself is closed and invariant ... but are there other ones ?

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We consider the two transformations $(x \mapsto 2x)$ and $(x \mapsto 3x)$ and we would like to know what are the closed invariant subsets of $\mathbb{T}$.

It is clear that 0 is invariant and that the orbit of rational points are finite (so they are closed and by definition, invariant). The circle itself is closed and invariant ... but are there other ones ?

A result by Furstenberg proves that every orbit is either finite or dense but we don't know how it is distributed in the circle.

From now on, we note $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. This groups acts on the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$.

Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$

We have the same question than before : what are the closed subsets of the torus that are invariant by A and B ?

It is clear that $\{0\}$ is invariant. Moreover, for any $\frac{p}{q} \in \mathbb{Q}^2/\mathbb{Z}^2$, the orbit of p/q stays in $\frac{1}{q}\mathbb{Z}^2/\mathbb{Z}^2$ and so, is finite.

But are there other closed invariant subsets ? (except from the torus itself)

To study this, we define a random walk : we fix a probability measure ρ (eg. $\rho = \frac{1}{2}\delta_A + \frac{1}{2}\delta_B$) on $\mathrm{SL}_2(\mathbb{Z})$, and then, starting at some point $x \in \mathbb{T}^2$, we consider

$$\begin{cases} X_0 & = & x \\ X_{n+1} & = & g_{n+1}X_n \end{cases}$$

where (g_n) is an iid sequence of law ρ.

As we already said, the behaviour of the walk is easy to study when the staring point is rational but what happens if it is irrational ?

As we already said, the behaviour of the walk is easy to study when the staring point is rational but what happens if it is irrational ?

Theorem (Bourgain, Furmann, Lindenstrauss and Mozes)

Let ρ be a probability measure on Γ whose support generates a non-elementary subgroup and such that for some $\varepsilon \in \mathbb{R}_+^*$, $\sum_{\gamma\in\Gamma} \|\gamma\|^\varepsilon \rho(\gamma)$ is finite.
Note $\nu =$ Lebesgue's measure on $\mathbb{T}^2$.
Then, for any non rational point $x \in \mathbb{T}^2$ and any continuous function f,

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \to \int f \,\mathrm{d}\nu \;\; \rho^{\otimes\mathbb{N}}\text{-a.e.}$$

In particular, the only closed $\mathbf{G}_\rho-$invariant subsets are the orbits of rational points and the torus itself (because this result proves that if x is irrational then it's orbit is dense) and every orbit of an irrational point is equidistributed in the torus.

Many of the results we stated here are still valid in dimension d if one assumes that $\mathbf{G}_\rho$ doesn't fix any finite union of non-trivial subspaces of $\mathbb{R}^d$ and has an element g which has a locally attractive fixed point in the projective space $\mathbb{P}(\mathbb{R}^d)$.

Merci de votre attention !