RANDOM MATRIX PRODUCTS

SEMINAR IN HOMOGENEOUS DYNAMICS, TAU 2024

1. Introduction

A sequence $\{Y_n \cdot \dots \cdot Y_1\}_n$ of products of i.i.d. random invertible matrices, or its action on a $\mathbb{R}^d$ vector $\{Y_n \cdot \dots \cdot Y_1 \cdot x\}_n$, possibly in the reversed order of multiplication $\{X_1 \cdot \dots \cdot X_n \cdot x\}_n$, give rises to many mathematical questions. Some are questions analogous to the classical probabilistic theorems regarding growth and (normalized) convergence of sum of i.i.d. random variables - the Law of Large Numbers, Central Limit Theorem, and more refined rate of convergence theorems such as the Law of the Iterated Logarithm; Some questions are of geometric flavor - spectrum of the products, special degenerated subspaces; and some are of dynamical flavor, about the relations between the probabilistic measure μ on $GL_d(\mathbb{R})$ and its continuous μ-invariant measures on $P\left(\mathbb{R}^d\right)$ the projective space of directions in $\mathbb{R}^d$.

Since the completion of the foundation of classical probability theory in the 60's, there have been lots of (pure and applied) researches on Random Matrices (and their multiplication), up to nowadays.

Last week we have seen Furstenberg-Kesten and Oseledets theorems, which well analyzes the growth of $\|Y_n \cdot \dots \cdot Y_1\|$, or $\|Y_n \cdot \dots \cdot Y_1 \cdot x\|$, $x \in \mathbb{R}^d$. For our purposes, to model the multiplication $Y_n \cdot \dots \cdot Y_1$ of random i.i.d. invertible matrices, we consider the probability preserving system with space $GL_d(\mathbb{R})^{\mathbb{N}}$, measure $\mu^{\otimes\mathbb{N}}$ and the shift operator. We conclude that $\mu^{\otimes\mathbb{N}}$-a.s. there exists $\lambda_+(\omega), \lambda_-(\omega) \in \mathbb{R}$ s.t.

$$\lambda_+ = \lim \frac{1}{n} \log \|Y_n \cdot \dots \cdot Y_1\| \ , \ \lambda_- = \lim \frac{1}{n} \log \left\|(Y_n \cdot \dots \cdot Y_1)^{-1}\right\|$$

and for the $d = 2$ case, when restricted to $Y_n \in SL_2^{\pm}(\mathbb{R})$, we either have $\lambda_+ = \lambda_- = 0$, for isometries, or $\lambda_+ > 0 > \lambda_- = -\lambda_+$ and for all but a proper subspace of $\mathbb{R}^2$,

$$\lim \frac{1}{n} \log \|Y_n \cdot \dots \cdot Y_1 x\| = \lambda_+ > 0$$

which guarantees an exponential rate of growth. Moreover, from ergodicity $\lambda_+(\omega)$ is a.e. a constant γ (the upper Lyapunov exponent). The growth rate analysis of the case $X_1 \dots X_n$ (where the matrices are multiplied in the reversed order) is the same, by use of the identity $\|Y_n \cdot \dots \cdot Y_1\| = \left\|(Y_n \cdot \dots \cdot Y_1)^t\right\|$.

Through the proof of the $d = 2$ case we have also seen that the rate of convergence in direction to the eigenspaces of $Y_n \cdot \dots \cdot Y_1$ is exponentially fast:

$$\overline{\lim} \frac{1}{n} \log \delta(s_n, s_{n+1}) \leqslant -2\lambda_+$$

Today we will be interested in geometric questions somewhat complementary to the growth rate question. We will only care for the direction of vectors, essentially reducing $x \in \mathbb{R}^2 \setminus \{0\}$ to its direction $\overline{x} \in P(\mathbb{R}^2)$. In the the scenarios $\{Y_n \cdot \dots \cdot Y_1\}_n$, $\{X_1 \cdot \dots \cdot X_n\}_n$, we ask about the direction of $X_1 \cdot \dots \cdot X_n \cdot \overline{x}$ and whether it converges, and about the angular distances such as $\delta(Y_n \cdot \dots \cdot Y_1 \overline{x}, Y_n \cdot \dots \cdot Y_1 \overline{y})$.

For starter, we recall the definition of the angular distance $\delta(\cdot,\cdot)$ in $P(\mathbb{R}^2)$, and how $\delta(A_n\overline{x}, A_n\overline{y})$ relates to the magnitudes $\|A_n\overline{x}\|, \|A_n\overline{y}\|$.

Definition 1. For $\overline{x}, \overline{y} \in P(\mathbb{R}^2)$, x, y unit vectors, we define

$$\delta(\overline{x}, \overline{y}) = \sqrt{1 - \langle x, y \rangle^2} = |\sin(\angle \overline{x}, \overline{y})|$$

which for the $d = 2$ and $x, y \in \mathbb{R}^2 \setminus \{0\}$ (not necessarily unital) can also be computed by

$$\delta(\overline{x}, \overline{y}) = \left| \frac{\det(x|y)}{\|x\| \|y\|} \right| = \frac{|x_1 y_2 - x_2 y_1|}{\|x\| \|y\|}.$$

Claim 2. For $A_n \in SL_2^{\pm}(\mathbb{R})$, such as the determinant-normalized product $A_n = Y_n \dots Y_1$, we get

$$\delta(A_n\overline{x}, A_n\overline{y}) = \frac{\|x\| \|y\| \delta(\overline{x}, \overline{y})}{\|A_n\overline{x}\| \|A_n\overline{y}\|}$$

and hence a positive Lyapunov exponent ensures (a.s.) exponentially fast angular distance convergence to zero.

2. Intuition

The SVD decomposition, the effects of multiplying a "narrow" matrix by a more "rounded" matrix from the right or from the left.

3. Formal Analysis of the $d = 2$ Case

Lemma 3. *Let μ be a distribution on $GL_2(\mathbb{R})$, ν a μ-inv distribution on $P(\mathbb{R}^2)$. Then μ-a.s. the sequence of $P(\mathbb{R}^2)$-distributions*

$$\{X_1(\omega) \dots X_n(\omega) M_1 \dots M_k \nu\}_n$$

weakly converges to a distribution $\nu(\omega)$, same for any $M_1, \dots, M_k \in supp(\mu)$, $k \geqslant 0$.

Proof. For a bounded Borel function $f : P(\mathbb{R}^2) \to \mathbb{R}$ define $F : GL_2(\mathbb{R}) \to \mathbb{R}$ by

$$F(g) = \int f(g\overline{x}) \, d\nu(\overline{x}).$$

Then for $A_n = X_1 \dots X_n$, $n \geqslant 0$,

$$\begin{aligned} \mathbb{E}\left[F\left(A_{n+1}\right) \;\middle|\; A_n\right] &= \int F\left(A_n g\right) d\mu\left(g\right) \\ &= \iint f\left(A_n g \overline{x}\right) d\mu\left(g\right) d\nu\left(\overline{x}\right) \\ \left[\overline{y} = g\overline{x} \sim \mu * \nu = \nu\right] \quad &= \int F\left(A_n \overline{y}\right) d\nu\left(\overline{y}\right) \\ &= F\left(A_n\right) \end{aligned}$$

meaning $\left\{F\left(A_n\right)\right\}_n$ is a martingale. f being bounded, the martingale is bounded, and by martingale convergence theorem μ-a.s $F\left(A_n\right) \to \Gamma f$, and

$$\mathbb{E}\left[\Gamma f\right] = \mathbb{E}\left[F\left(A_0\right)\right] = \mathbb{E}\left[\int f\left(\overline{x}\right) d\nu\left(\overline{x}\right)\right] = \int f d\nu.$$

Considered as an integral, Γf defines a measure $\nu\left(\omega\right)$ and $X_1\left(\omega\right) \dots X_n\left(\omega\right) \nu \to v\left(\omega\right)$ holds. $\nu\left(\omega\right)$ must be a probabilistic measure.

To strengthen the result and show that a.s. $X_1\left(\omega\right) \dots X_n\left(\omega\right) M_1 \dots M_k \nu \to \nu\left(\omega\right)$, consider the distribution $\lambda := \sum_{r=0}^{\infty} 2^{-r-1} \mu^{*r}$ on $GL_2\left(\mathbb{R}\right)$. Recall that as a martingale, we have for any $j, r \geqslant 0$,

$$\mathbb{E}\left[\left|F\left(A_{j+r}\right) - F\left(A_j\right)\right|^2\right] = \mathbb{E}\left[F\left(A_{j+r}\right)^2\right] - \mathbb{E}\left[F\left(A_j\right)^2\right]$$

Thus, using cancellation,

$$\begin{aligned} &\sum_{j=1}^{\infty} \mathbb{E}\left[\int \left|F\left(A_j g\right) - F\left(A_j\right)\right|^2 d\lambda\left(g\right)\right] = \\ &\sum_{j=1}^{\infty} \sum_{r=0}^{\infty} 2^{-r-1} \mathbb{E}\left[\int \left|F\left(A_j g\right) - F\left(A_j\right)\right|^2 d\mu^{*r}\left(g\right)\right] = \\ &\sum_{r=0}^{\infty} 2^{-r-1} \sum_{j=1}^{\infty} \mathbb{E}\left[\left|F\left(A_{j+r}\right) - F\left(A_j\right)\right|^2\right] \leqslant \\ &\sum_{r=0}^{\infty} 2^{-r-1} \cdot r \cdot \sup_g \left|F\left(g\right)\right|^2 < \infty \end{aligned}$$

so $\mathbb{P} \otimes \lambda$-a.s

$$\sum_{j=1}^{\infty} \left|F\left(A_j g\right) - F\left(A_j\right)\right|^2 < \infty$$

and

$$\lim F\left(A_j g\right) = \lim F\left(A_j\right) = \Gamma f = \int f d\nu\left(\omega\right).$$

$\square$

Proposition 4. *Let $\{X_i\}_{i=1}^{\infty} \subset SL_2(\mathbb{R})$ be iid with distribution μ, and ν be μ-invariant continuous*[1] *distribution on $P(\mathbb{R}^2)$. If $supp(\mu)$ is not contained in a compact subgroup*[2] *of $SL_2(\mathbb{R})$, then for μ-almost any $\omega \in \Omega$ there exists $\overline{Z(\omega)} \in P(\mathbb{R}^2)$ such that the measures $X_1(\omega) \dots X_n(\omega)\nu$ weakly converges to $\delta_{\overline{Z(\omega)}}$.*
Moreover, $\overline{Z} \sim \nu$ and this is the unique continuous μ-inv measure of $P\left(\mathbb{R}^d\right)$.

Proof. Let $A_n = X_1 \dots X_n$, and $B_n = \|A_n\|^{-\frac{1}{2}} A_n$ (of norm 1). Take a subsequence B_{n_j} converging to $A(\omega)$. μ-a.s., for any $M \in supp(\mu)$ we get from lemma 3 that

$$B_n(\omega) M\nu = A_n(\omega) M\nu \to \nu(\omega) \quad , \quad B_n(\omega)\nu = A_n(\omega)\nu \to \nu(\omega)$$

but we also have

$$B_n(\omega) M\nu \to A(\omega) M\nu \quad , \quad B_n(\omega)\nu \to A(\omega)\nu$$

hence $A(\omega) M\nu = A(\omega)\nu = \nu(\omega)$. Assuming (for contradiction) that $A(\omega)$ is invertible, then $M\nu = \nu$, meaning

$$supp(\mu) \subseteq H := \{M \in SL_2(\mathbb{R}) : M\nu = \nu\}$$

and observe that H is a subgroup of $GL_2(\mathbb{R})$ that cannot be closed, or else $supp(\mu)$ is contained in a compact subgroup of $SL_2(\mathbb{R})$. Thus there exist a sequence $M_j \in H$, $\|M_j\| \to \infty$, s.t. $\|M_j\|^{-\frac{1}{2}} M_j \to M$ where M is not invertible, yet by continuity $v = \lim M_j\nu = \lim \|M_j\|^{-\frac{1}{2}} M_j\nu = M\nu$, but then $M\nu$ is a Dirac measure on $P(\mathbb{R}^2)$, contradicting our assumption that ν is continuous.

We conclude that a.s. $A(\omega)$ is not invertible, hence for the direction $\overline{Z(\omega)}$ of the range of $A(\omega)$ we get that $\nu(\omega) = A(\omega)\nu = \delta_{\overline{Z(\omega)}}$. Moreover, for any Borel function $f \colon P(\mathbb{R}^2) \to \mathbb{R}$

$$\int f(\overline{x})\, d\nu(x) = \mathbb{E}\left[\int f(\overline{x})\, d\nu_\omega(x)\right] = \mathbb{E}\left[f\left(\overline{Z}\right)\right]$$

hence $\overline{Z} \sim \nu$. We conclude that ν must be unique for μ. □

Remark 5. Because $P(\mathbb{R}^2)$ is compact, there must always be an μ-inv distribution ν on $P(\mathbb{R}^2)$, but it might not be continuous. The above proposition implies that there is up to one μ-inv continuous distribution, with the stated conditions.

Example 6. What are the μ-inv measures of $P(\mathbb{R}^2)$ for $\mu = \delta_M$, $M = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$?

Proposition 7. *Let $\{A_n\}_n \subset SL_2(\mathbb{R})$ and ν be a continuous distribution on $P(\mathbb{R}^2)$ s.t. $A_n\nu$ weakly converges to $\delta_{\overline{z}}$, $\overline{z} \in P(\mathbb{R}^2)$, $\|z\| = 1$. Then*

$$\lim \|A_n\| = \lim \left\|A_n^t\right\| = \infty$$

[1]For the $d > 2$ case we would require any hyperplane subspace of $P\left(\mathbb{R}^2\right)$ to be o ν-null measurable.

[2]For the $d > 2$ case we would require that $\langle supp(\mu)\rangle$ is of "index" 1, i.e. there exists a limit of (normalized) matrices from $supp(\mu)$ which is a matrix of rank 1.

and for any $x \in \mathbb{R}^2$,

$$\lim \frac{\|A_n^t x\|}{\|A_n^t\|} = \langle x, z \rangle .$$

Proof. Let A be a limit of a subsequence of $\left\{ \|A_n\|^{-1} A_n \right\}_n$. A must be of unit norm, and in particular $A \neq 0$, however A cannot be invertible (or else $A\nu = \lim A_{n_j}\nu = \delta_{\overline{z}}$, which implies that $\nu = A^{-1}\delta_{\overline{z}}$ is not continuous). Hence

$$0 = |\det(A)| = \lim \left| \|A_n\|^{-2} \det(A_n) \right| = \lim \frac{1}{\|A_n\|^2}.$$

The second claim holds since $\overline{z}$ must be the direction of the range of A. □

4. The $d > 2$ case

If time permits, we discuss the key differences between the $d = 2$ and $d > 2$ cases.

5. Further reading

To prepare for the talk, I've read Part A of the textbook "Products of Random Matrices with Application to Schrödinger Operators" by Bougerol & Lacroix, and a bit of the textbook "Random walks on reductive groups" by Benoist & Quint. The former is a very accessible book, yet the result are not presented in linear order, so the reader essentially has to "solve the puzzle" in the right order in the his/her mind. The latter is a more recent and more extensive textbook, written for somewhat more advanced audience, and is very well organized.

The content of the talk is originally based on the following three papers[3] :

- FURSTENBERG H. (1963). Non commuting random products. Trans. Amer. Soe. 108, 377-428.
- LE PAGE E. (1984). Repartition d'etat d'un operateur de Schrödinger aleatoire. Probability measures on groups VII, Proceedings Oberwolfach, Springer lectures notes series 1064, 309-367.
- GUIVARC'H, Y., and RAUGI, A. (1986). Products of random matrices: convergence theorems. In: Random Matrices and Their Applications, Workshop Brunswick Maine 1984.

[3]I didn't verified this myself. The reference numbers are erroneous in Bougerol & Lacroix's book, I hope I figured it out correctly.