Random Matrix Products

Alex Rutar*
University of St Andrews

Spring 2021†

*alex@rutar.org
†Last updated: April 18, 2021

Contents

Preface

These lecture notes on random matrix products are prepared for the reading group on random matrix products for the analysis group in Spring 2021. Much of the content is based on Alex Gorodnik's lecture notes for his course "Random walks on matrix groups". Any errors or omissions can be sent to the author.

I. The Multiplicative Ergodic Theorem

1 Random matrix products and the subadditive ergodic theorem

1.1 The Birkhoff ergodic theorem

Let Ω be a separable, second-countable metric space equipped with its Borel σ-algebra $\mathcal{B}$, and let μ be a Borel probability measure on Ω. Suppose we are given a measurable function $\theta : \Omega \to \Omega$. We denote the *pushforward* of μ by θ to denote the Borel probability measure defined by the rule

$$\theta_* \mu(E) = \mu(\theta^{-1}(E))$$

for Borel sets $E \subset \Omega$. We say that the function θ is *measure preserving* if $\theta_* \mu = \mu$. In this situation, we call the information (Ω, μ, θ) a *measure-preserving dynamical system*.

Given a Borel set $E \subset \Omega$, we say that E is θ-invariant if $\theta^{-1}(E) = E$, and denote the set of θ-invariant sets by $\mathcal{B}_\theta$. More generally, we say that a measurable function $f : \Omega \to K$ where K is a topological space is θ-invariant if $f(\omega) = f \circ \theta(\omega)$ for μ-a.e. ω. One can verify that $\mathcal{B}_\theta$ is a Borel σ-subalgebra of $\mathcal{B}$. In particular, f is θ-invariant if and only if f is $\mathcal{B}_\theta$-measurable. We say that (Ω, μ, θ) is *ergodic* if each θ-invariant set $E \in \mathcal{B}_\theta$ either has $\mu(E) = 0$ or $\mu(E) = 1$.

We will denote by T^n the n-fold composition $T \circ \cdots \circ T$. Given a function f, we write $f = f^+ + f^-$ where $f^+ \geq 0$ and $f^- \leq 0$. A standard result is the following.

1.1 Theorem (Birkhoff Pointwise Ergodic). *Let (Ω, μ, θ) be an ergodic measure-preserving dynamical system and let $f = f^+ + f^-$ satisfy $f_+ \in L^1(\Omega, \mu)$. Then for μ-a.e. $\omega \in \Omega$, we have*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(\omega)) = \int_\Omega f \, \mathrm{d}\mu$$

where the limit may be attained at $-\infty$.

We have written Theorem 1.1 in additive notation, but it can be easily rephrased in multiplicative notation. Denote by $\log^+(x) = \max(0, \log x)$. Write $g = \exp(f)$ and note that $f_+ = \log^+(g)$. Then for μ-a.e. $\omega \in \Omega$,

$$\lim_{n \to \infty} (g(T^{n-1}\omega) \cdots g(\omega))^{1/n} = \exp\left(\int_\Omega \log g \, \mathrm{d}\mu\right).$$

Of course, here, the group written in product notation is still commutative. In the following section, we consider a more general setting where this is no longer the case.

1.2 Random matrix products

The setting of Theorem 1.1 is nice, but in these notes we are interested in a somewhat more general situation. First consider the following example. Let Ω denote the compact product space $\mathrm{GL}_n(\mathbb{C})^{\mathbb{N}}$ equipped with the left-shift map $\sigma : \Omega \to \Omega$ given by

$$\sigma((M_n)_{n=1}^{\infty}) = (M_n)_{n=2}^{\infty}$$

for a sequence of matrices $(M_n)_{n=1}^{\infty} \in \Omega$. Let ν be a probability measure on $\mathrm{GL}_n(\mathbb{C})$ and let $X_i : \Omega \to \mathrm{GL}_n(\mathbb{C})$ for $i \in \mathbb{N}$ be independent random matrices with distribution ν. Asymptotic behaviour of random products of the form $X_n \cdots X_1$ can be interpreted as a matrix-valued generalization of the law of large numbers.

More generally, we are interested in matrix-valued measurable functions, i.e. functions $X : \Omega \to \mathrm{GL}_n(\mathbb{C})$ on a measure-preserving space (Ω, μ, θ). This setting is a generalization of the setting in Theorem 1.1, where we considered a measurable function $f : \Omega \to \mathbb{R}$ satisfying an integrability criteria. Let $\|\cdot\| : \mathrm{GL}_n(\mathbb{C})$ be a matrix norm. We will assume that $\|\cdot\|$ is *submultiplicative* (i.e. $\|AB\| \leq \|A\| \|B\|$), but we do not lose any generality since all matrix norms are equivalent. We also assume that X satisfies the integrability condition

$$\int_{\Omega} \log^{+} \|X(\omega)\| \, \mathrm{d}\omega < \infty.$$

As in the prior section, we are interested in determining statistical information concerning the limit of the random matrix product

$$S_n(\omega) = X(T^{n-1}\omega) \cdots X(\omega).$$

We will investigate various statistical properties of the random products $S_n(\omega)$. Here are three such examples which we will focus on: **TODO: add links here once the sections are written up**

(i) the growth rate of $\|S_n(\omega)\| = \left\|X(\theta^{n-1}\omega) \cdots X(\omega)\right\|$ for large n and "typical" ω.

(ii) the growth rate from a fixed starting point $\left\|X(\theta^{n-1}\omega) \cdots X(\omega)v\right\|$ for some $v \in \mathbb{C}^n$

(iii) the behaviour of the directions $\left\|X(\theta^{n-1}\omega) \cdots X(\omega)v\right\| / \left\|X(\theta^{n-1}\omega) \cdots X(\omega)v\right\|$ for some $v \in \mathbb{C}^n$.

Here are a couple examples of settings

Example. 1. Given fixed matrices $M_1, \ldots, M_\ell \in \mathrm{GL}_n(\mathbb{C})$, generate a sequence $S_0 = I$ and $S_{n+1} = M_i \cdot S_n$ where we take matrix M_i with probability $1/\ell$. The products S_n can be interpreted as a random walk on $\mathrm{GL}_n(\mathbb{C})$ (or $\mathbb{C}^n$) where the "steps" are given by multiplication by a matrix M_i.

2. If $U \subset \mathbb{R}^d$ is an open set and $F : U \to U$ is smooth, by the chain rule, the Jacobian of F^n at a point u satisfies

$$D(F^n)_u = (DF)_{F^{n-1}u} \cdots (DF)_u.$$

Here, $DF : U \to \mathrm{GL}_d(\mathbb{R})$ is a matrix-valued measurable function. The growth rate of DF is related to the entropy of F and the dimension of invariant measures.

3. If $T_i(x) = A_i x + t_i$ where $A_1, \ldots, A_\ell \in \mathrm{GL}_n(\mathbb{R})$ have operator norms $\|A_i\| < 1$ for $i = 1, \ldots, \ell$ and $t_i \in \mathbb{R}^n$, then there is a unique *self-affine set* K satisfying

$$K = \bigcup_{i=1}^{\ell} T_i(K)$$

and, given probabilities $p_1, \dots, p_\ell$, a unique *self-affine measure*, which is a Borel probability measure ν satisfying

$$\nu = \sum_{i=1}^{\ell} p_i (T_i)_* \mu.$$

Here, dimensional properties of the measure ν are related to properties of random products of the matrices $\{A_1, \dots, A_\ell\}$.

1.3 LYAPUNOV EXPONENTS AND THE SUBADDITIVE ERGODIC THEOREM

A fundamental statistical property associated with the matrix-valued function X is the following.

Definition. With notation as above, we define the *top Lyapunov exponent* $\lambda : \Omega \to \mathbb{R}$ by

$$\lambda(\omega) = \lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)\| .$$

We now have the following fundamental result.

1.2 Theorem (Furstenburg-Kesten). *The function λ is θ-invariant and satisfies*

$$\int_\Omega \lambda(\omega)\, \mathrm{d}\omega = \lim_{n\to\infty} \frac{1}{n} \int_\Omega \log \|S_n(\omega)\|\, \mathrm{d}\omega.$$

This result can be thought of as interchanging the limit with the integral, i.e. averaging over space is the same as

In fact, we will prove Theorem 1.2 as a consequence of a more general result. We first make some observations about the average $a_n := \int_\Omega \log \|S_n(\omega)\|$. Observe by submultiplicative of the matrix norm that

$$\begin{aligned}
a_{n+m} &:= \int_\Omega \log \|S_{n+m}(\omega)\|\, \mathrm{d}\omega \\
&= \int_\Omega \log \left\|X(\theta^{n+m-1}\omega) \cdots X(\omega)\right\| \mathrm{d}\omega \\
&\le \int_\Omega \log \left\|X(\theta^{n+m-1}\omega) \cdots X(\theta^{m}\omega)\right\| \mathrm{d}\omega + \int_\Omega \log \left\|X(\theta^{m-1}\omega) \cdots X(\omega)\right\| \mathrm{d}\omega \qquad (1.1) \\
&= \int_\Omega \log \|S_n(\theta^m\omega)\|\, \mathrm{d}\omega + \int_\Omega \log \|S_m(\omega)\|\, \mathrm{d}\omega \\
&= a_n + a_m
\end{aligned}$$

where the last line follows by the integrability condition on X along with the fact that θ is measure preserving.

Definition. We say that the sequence $(a_n)_{n=1}^{\infty} \subset \mathbb{R}$ is *subadditive* if $a_{n+m} \le a_n + a_m$ for each $n, m \in \mathbb{N}$. More generally, we say that a sequence of functions $\varphi_n : \Omega \to \mathbb{R}$ is *subadditive* if

$$\varphi_{n+m}(\omega) \le \varphi_n(\theta^m\omega) + \varphi_m(\omega). \tag{1.2}$$

The following lemma is straightforward.

1.3 Lemma. *If $(a_n)_{n=1}^{\infty}$ is a subadditive sequence, then $\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 1} \frac{a_n}{n}$.*

In particular, implies that the limit

$$\lim_{n\to\infty} \frac{1}{n} \int_{\Omega} \log \|S_n(\omega)\| \,\mathrm{d}\omega$$

always exists. Moreover, if we set $\varphi_n(\omega) = \log \|S_n(\omega)\|$, we observed in (1.1) that the sequence of functions φ_n is subadditive. Thus Theorem 1.2 is a consequence of the following more general result.

Throughout the statement and the proof, note that many inequalities implicity hold for μ-a.e. $\omega \in \Omega$.

1.4 Theorem (Kingman's Subadditive Ergodic). *Let $\varphi_n : \Omega \to \mathbb{R}$ be a subadditive sequence with $\varphi_1^+ \in L^1(\Omega, \mu)$. Then the limit $\varphi(\omega) := \lim_{n\to\infty} \frac{\varphi_n(\omega)}{n}$ exists for almost every ω Moreover, φ is θ-invariant and*

$$\int_{\Omega} \varphi(\omega) \,\mathrm{d}\omega = \lim_{n\to\infty} \frac{1}{n} \int_{\Omega} \varphi_n(\omega) \,\mathrm{d}\omega =: L.$$

Set

$$\begin{aligned} \varphi_-(\omega) &= \liminf_{n\to\infty} \frac{\varphi_n(\omega)}{n} \\ \varphi_+(\omega) &= \limsup_{n\to\infty} \frac{\varphi_n(\omega)}{n}. \end{aligned}$$

We first observe that φ_- (and by an analgous argument φ_+) is θ-invariant. By the subadditivity assumption (1.2) with $m = 1$,

$$\varphi_-(\omega) \leq \liminf_{n\to\infty} \frac{\varphi_n(\theta\omega) + \varphi(\omega)}{n+1} = \varphi_-(\theta\omega)$$

so with $X_a = \{\omega \in \Omega : \varphi_-(\omega) \geq a\}$ for any $a \in \overline{\mathbb{R}}$, we have $\theta^{-1}(X_a) \supset X_a$. But θ is measure-preserving, so this can forces $\mu(\theta^{-1}(X_a) \setminus X_a) = 0$, i.e. φ_- is θ-invariant.

We first prove two intermediate lemmas.

1.5 Lemma. *We have $\int_{\Omega} \varphi_-(\omega) \,\mathrm{d}\omega = L$.*

Proof. Let $\epsilon > 0$ be arbitrary. For $k \in \mathbb{N}$, define

$$E_k = \Big\{\omega \in \Omega : \frac{\varphi_j(\omega)}{j} \leq \varphi_-(\omega) + \epsilon \text{ for some } j = 1, \dots, k\Big\}.$$

Note that $E_k \subset E_{k+1}$ and $\bigcup_k E_k = \omega$. Now set

$$\psi_k(\omega) = \begin{cases} \varphi_-(\omega) + \epsilon & : \omega \in E_k \\ \varphi_1(\omega) & : \omega \in E_k^c \end{cases}$$

Observe that $\psi_k \geq \varphi_-(\omega) + \epsilon$ by definition of E_k. By subadditivity, $\varphi_n(\omega) \leq \sum_{i=0}^{n-1} \varphi(\theta^{i-1}\omega)$. We want to convert this into a statement using our desired limit φ_-. We first demonstrate such a bound in terms of the approximations ψ_k of φ_-. To be precise, we will prove that for all $n > k$ and almost every $\omega \in \Omega$,

$$\varphi_n(\omega) \leq \sum_{i=0}^{n-k-1} \psi_k(\theta^i\omega) + \sum_{i=n-k}^{n-1} \max\{\psi_k, \varphi_1\}(\theta^i\omega). \tag{1.3}$$

Since φ_- is θ-invariant, we may assume that $\varphi_-(\theta^n\omega) = \varphi_-(\omega)$ for all n.

We will inductively define a sequence $m_0 \leq n_1 < m_1 \leq n_2 < \cdots$ as follows. Let $m_0 = 0$. Inductively, let $n_j \geq m_{j-1}$ be the minimal integer such that $\theta^{n_j}\omega \in E_k$ (if it exists). By definition of E_k, there exists m_j such that $1 \leq m_j - n_j \leq k$ and

$$\varphi_{m_j-n_j}(\theta^{n_j}\omega) \leq (m_j - n_j)(\varphi_-(\theta^{n_j}\omega) + \epsilon). \tag{1.4}$$

Let ℓ be maximal such that $m_\ell \leq n$. By subadditivity, inductively applying the inequality

$$\varphi_i(\omega) \leq \varphi_1(\theta^i\omega) + \varphi_{i-1}(\omega)$$

if $i \neq m_j$ for some j and the inequality

$$\varphi_{m_j}(\omega) \leq \varphi_{n_j}(\omega) + \varphi_{m_j-n_j}(\theta^{n_j}\omega),$$

we obtain

$$\varphi_n(\omega) \leq \sum_{i\in I} \varphi_1(\theta^i\omega) + \sum_{j=1}^{\ell} \varphi_{m_j-n_j}(\theta^{n_j}\omega) \tag{1.5}$$

where $I = \bigcup_{j=0}^{\ell-1}[m_j, n_{j+1}) \cup [m_\ell, n)$. Now if $i \in I$ with $i < n_{\ell+1}$, we have

$$\varphi_1(\theta^i\omega) = \psi_k(\theta^i\omega)$$

since $\theta^i\omega \notin E_k^c$. Since $\varphi_-(\theta^n\omega) = \varphi_-(\omega)$ and $\psi_k \geq \varphi_- + \epsilon$ by definition, by (1.4),

$$\varphi_{m_j-n_j}(\theta^{n_j}\omega) \leq \sum_{i=n_j}^{m_j-1} (\varphi_-(\theta^i\omega) + \epsilon) \leq \sum_{i=n_j}^{m_j-1} \psi_k(\theta^i\omega).$$

Thus (1.3) follows by (1.5) and the fact that $n - n_\ell < k$.

Now, suppose $\varphi_n/n \geq -C$ for some fixed constant $C > 0$. Then by Fatou's Lemma,

$$\int_\Omega \varphi_-(\omega)\,\mathrm{d}\omega \leq \liminf_{n\to\infty} \frac{1}{n}\int_\Omega \varphi_n(\omega)\,\mathrm{d}\omega = L.$$

To get the lower bound, by (1.3),

$$\frac{1}{n}\int_\Omega \varphi_n(\omega)\,\mathrm{d}\omega \leq \frac{n-k}{n}\int_\Omega \psi_k(\omega)\,\mathrm{d}\omega + \frac{k}{n}\int_\Omega \max\{\psi_k, \varphi_1\}(\omega)\,\mathrm{d}\omega.$$

Thus taking the limit as n goes to infinity, we have

$$L \leq \int_\Omega \psi_k(\omega)\,\mathrm{d}\omega.$$

Moreover, $\lim_{k\to\infty} \psi_k = \varphi_- + \epsilon$ by definition, so that $L \leq \int_\Omega \varphi_-(\omega)\,\mathrm{d}\omega + \epsilon$. But $\epsilon > 0$ was arbitrary, giving the desired equality.

More generally, let $\varphi_n^{(C)} = \max\{\varphi_n, -Cn\}$ and $\varphi_-^{(C)} = \max\{\varphi_-, -C\}$. Then by the Monotone Convergence Theorem,

$$\begin{aligned}\int_\Omega \varphi_-(\omega)\,\mathrm{d}\omega &= \inf_C \int_\Omega \varphi_-^{(C)}(\omega)\,\mathrm{d}\omega = \inf_C \inf_n \int_\Omega \frac{\varphi_n^{(C)}(\omega)}{n}\,\mathrm{d}\omega \\ &= \inf_n \int_\Omega \frac{\varphi_n(\omega)}{n}\,\mathrm{d}\omega = L\end{aligned}$$

as required. □

1.6 Lemma. *We have* $\limsup_{n\to\infty} \frac{\varphi_{nk}}{nk} = \varphi_+(\omega)$.

Proof. The upper bound follows since by subadditivity and invariance of φ_+,

$$\begin{aligned}\limsup_{n\to\infty} \frac{\varphi_{nk}(\omega)}{n} &\le \sum_{j=0}^{k-1} \limsup_{n\to\infty} \frac{\varphi_n(\theta^{nj}\omega)}{n} \\ &= k\varphi_+(\omega).\end{aligned}$$

Conversely, given $n \in \mathbb{N}$, write $n = kq_n + r_n$ where $r_n \in \{1, \ldots, k\}$. By subadditivity,

$$\varphi_n(\omega) \le \varphi_{kq_n}(\omega) + \varphi_{r_n}(\theta^{kq_n}\omega) \le \varphi_{kq_n}(\omega) + \psi(\theta^{kq_n}\omega)$$

where $\psi = \max\{\varphi_1^+, \ldots, \varphi_k^+\}$. By assumption, $\psi \in L^1$. Below, we will show that

$$\lim_{n\to\infty} \frac{\psi \circ \theta^{kq_n}}{q_n} = 0. \tag{1.6}$$

Assuming this result, we have

$$\limsup_{n\to\infty} \frac{\varphi_n}{n} \le \limsup_{n\to\infty} \frac{1}{n}\varphi_{kq_n} = \frac{1}{k} \limsup_{n\to\infty} \frac{1}{q_n}\varphi_{kq_n} \le \frac{1}{k} \limsup_{n\to\infty} \frac{\varphi_{nk}}{n}.$$

Let's prove (1.6). Let $\epsilon > 0$ be arbitrary. We first observe that

$$\begin{aligned}\sum_{n=1}^{\infty} \mu\big(\{\omega \in \Omega : |\psi(\theta^n\omega)| \ge \epsilon n\}\big) &= \sum_{n=1}^{\infty} \mu\big(\{\omega \in \Omega : |\psi(\omega)| \ge \epsilon n\}\big) \\ &= \sum_{k=1}^{\infty} k\mu\big(\{\omega \in \Omega : k\epsilon \le |\psi(\omega)| < (k+1)\omega\}\big) \\ &\le \int_\Omega \frac{|\psi(\omega)|}{\epsilon} < \infty.\end{aligned}$$

Thus the result follows by the Borel-Cantelli Lemma. □

Proof (of Theorem 1.4). We are now in position to complete the proof. As before, we first assume that $\varphi_n/n \ge -C$ for some fixed $C > 0$. Set

$$\phi_n = -\sum_{j=0}^{n-1} \varphi_k \circ \theta^{kj}.$$

By definition, $\phi_{n+m} = \phi_m + \phi_n \circ \theta^{km}$ and $\phi_1 = -\varphi_k \leq Ck$, so $\phi_1^+ \in L^1$. Let $\phi_- = \liminf_{n\to\infty} \frac{\phi_n}{n}\,\mathrm{d}\omega$. Then by Lemma 1.5 and the fact that μ is θ-invariant,

$$\int_\Omega \phi_-(\omega)\,\mathrm{d}\omega = \lim_{n\to\infty} \int_\Omega \frac{\phi_n(\omega)}{n}\,\mathrm{d}\omega = \int_\Omega \varphi_k(\omega)\,\mathrm{d}\omega.$$

Now by Lemma 1.6,

$$-\phi_- = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi_k \circ \theta^{kj} \geq \limsup_{n\to\infty} \frac{\varphi_{kn}}{n} = k\varphi_+.$$

Thus

$$\int_\Omega \varphi_+\,\mathrm{d}\omega \leq -\frac{1}{k}\int_\Omega \varphi_k(\omega)\,\mathrm{d}\omega.$$

But this holds for any $k \in \mathbb{N}$, so that $\int_\Omega \varphi_+ \leq L$.

In general, as in the proof of Lemma 1.5, set $\varphi_n^{(C)} = \max\{\varphi_n, -Cn\}$ and $\varphi_\pm^{(C)} = \max\{\varphi_\pm, -C\}$. We just showed that $\int_\Omega -\varphi_-^{(C)}\,\mathrm{d}\omega = \int_\Omega \varphi_+^{(C)}(\omega)\,\mathrm{d}\omega$. But $\varphi_-^{(C)} \leq \varphi_+^{(C)}$, so that $\varphi_-^{(C)} = \varphi_+^{(C)}$. Thus the result follows by the Monotone Convergence Theorem. □

Remark. This result generalizes Theorem 1.1 since, using the notation from that theorem, the function $\varphi_n(\omega) = \sum_{i=0}^{n-1} f(T^i\omega)$ is subadditive (since it is additive) and by invariance of T,

$$\int_\Omega f(T^i\omega)\,\mathrm{d}\omega = f(T^i\omega).$$

In fact, Theorem 1.1 follows directly from Lemma 1.5 since both $(\varphi_n)_{n=1}^\infty$ and $(-\varphi_n)_{n=1}^\infty$ are subadditive sequences of functions.

The argument in Lemma 1.6 can be interpreted as a "stability result" for subadditive sequences, which we then use to get control over φ_+ in the general case.