Mathematical Notebook

Notation 1. We will write

1. $\mathcal{M}_{n\times m}(\mathbb{F})$ for the set of n by m matrices with entries in the field $\mathbb{F}$.
2. $\mathrm{GL}(n,\mathbb{R}) = \{M \in \mathcal{M}_{n\times n}(\mathbb{R}) : \det M \neq 0\}$.
3. $\mathrm{SL}_+(n,\mathbb{R}) = \{M \in \mathcal{M}_{n\times n}(\mathbb{R}) : \det M = 1\}$.
4. $\mathrm{SL}_-(n,\mathbb{R}) = \{M \in \mathcal{M}_{n\times n}(\mathbb{R}) : \det M = -1\}$.
5. The transpose of $M \in \mathcal{M}_{n\times m}(\mathbb{R})$ is denoted by $M^\intercal$.
6. $f^+ = \sup(f, 0)$.

1 Review

1.0.1 Algebra

Definition 1. A set S together with a binary operation $\cdot$ is called a *semigroup* if, for every $a, b, c \in S$:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

Definition 2. A set G together with a binary operation $\cdot$ is called a *group* if it satisfies the following three axioms:

1. $\forall a, b, c \in G : (a \cdot b) \cdot c = a \cdot (b \cdot c)$;
2. $\exists e \in G, \forall a \in G : e \cdot a = a \cdot e = a$;
3. $\forall a \in G, \exists b \in G : a \cdot b = b \cdot a = e$.

Additionally, if the operation is commutative, then G is called an *Abelian* or *commutative* group.

1.0.2 Topology

Definition 3. Let X be a set and $\mathcal{O} \subseteq \mathcal{P}(X)$. The class $\mathcal{O}$ is a *topology* on X if the following conditions hold:

1. $\emptyset \in \mathcal{O}$;
2. $X \in \mathcal{O}$;

3. Given any sequence $(A_i \in \mathcal{O} : i \in I)$ we have

$$\bigcup_{i \in I} A_i \in \mathcal{O};$$

4. If $A, B \in \mathcal{O}$ then $A \cap B \in \mathcal{O}$.

If these axioms are satisfied, we say a set $A \subseteq X$ is *open* if $A \in \mathcal{O}$ and it is closed if $A^{\complement}$ is open. The pair $(X, \mathcal{O})$ is called a *topological space*.

Definition 4. Consider topological spaces $(X, \mathcal{O})$ and $(Y, \mathcal{T})$. A function $f : X \to Y$ is *continuous* if

$$\forall A \in \mathcal{T} : f^{-1}(A) \in \mathcal{O}.$$

Definition 5. Let $(S, \cdot)$ be a semigroup and $\mathcal{O}$ be a topology on S. If the map $(M_1, M_2) \mapsto M_1 \cdot M_2$ from S^2 to S is continuous, then $(S, \cdot, \mathcal{O})$ is a *topological semigroup.*

Definition 6. Let $(G, \cdot)$ be a group and $\mathcal{O}$ be a topology on G. If the map $\cdot$ and $g \mapsto g^{-1}$ are continuous then $(G, \cdot, \mathcal{O})$ is a *topological group.*

Note 1. In the previous definition, the domain of the binary operation is the Cartesian product $G \times G$. The topology on this set is the product topology. The same applies to the definition of topological semigroup.

Example 1. Since matrix multiplication is an associative operation, and the identity matrix acts as the multiplicative identity, the sets $\mathrm{GL}(n, \mathbb{R})$ and $\mathrm{SL}(n, \mathbb{R})$ are groups when considered together with the usual matrix multiplication operation.

1.0.3 Normed spaces and bounded linear operators

Definition 7. Let X be a vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on X. We say that $\|\cdot\|_2$ is *equivalent* to $\|\cdot\|_1$ if there exists $M, m > 0$ such that,

$$\forall x \in X : m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$$

Proposition 1. *Let X be a finite-dimensional vector space. If $\|\cdot\|_1$, $\|\cdot\|_2$ are norms on X, then they are equivalent.*

Note 2. The set $M_{n \times n}(\mathbb{R})$ together with the usual matrix multiplication and scalar multiplication is a vector space. Its dimension is clearly $\dim M_{n \times n}(\mathbb{R}) = n^2$. The previous proposition tells us that the norm we choose to work with is not that important. No matter the norm we fix, the resulting space has an equivalent notion of convergence. Now my question is: Is convergence all that matters for us? Since equivalent norms differ in *some* ways, how are we justified in ignoring those differences?

For convenience, unless specified otherwise, the norm on $\mathbb{R}^d$ that will be considered is the usual Euclidean norm, which, to every $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, assigns the quantity

$$\|x\| = \left(\sum_{i=1}^{d} x_i^2 \right)^{1/2}.$$

As for $\mathcal{M}_{n\times n}(d, \mathbb{R})$, the norm we'll consider by default is the operator norm, which maps $M \in \mathcal{M}_{n\times n}(d, \mathbb{R})$ to

$$\|M\| = \sup\{\|Mx^\intercal\| : x \in \mathbb{R}^d, \|x\| = 1\}.$$

One useful property of the operator norm is the following.

Proposition 2. *Let V, W and X be normed spaces and $T : V \to W$, $L : W \to X$ be bounded linear operators. Then*

$$\|LT\| \leq \|L\|\|T\|,$$

where $\|\cdot\|$ denotes the operator norm.

1.1 Probability Theory

1.1.1 What is a random matrix?

Note 3. No definition of random matrix is presented in [2]. In other sources, such as Wikipedia, it is stated that a random matrix is a matrix-valued random variable. I have not seen the general definition of r.v. before (usually the codomain is taken to be $\mathbb{R}$ or $\mathbb{R}^n$). So I fill these gaps next.

Definition 8. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(E, \mathcal{E})$ be a measurable space. A *random variable* is a $(\mathcal{F}, \mathcal{E})$-measurable map $X : \Omega \to E$.

Note 4. Speaking generally, clearly we would have $E = \mathcal{M}_{m\times n}(\mathbb{F})$, the question is, what is the suitable choice for $\mathcal{E}$? Since we usually consider the codomain of an r.v. to be $\mathbb{R}^n$ we typically work with the associated Borel σ-algebra. Is there an analogue here, perhaps generated by a "natural" topology on the set of matrices? Associated to every norm is a topology. So perhaps we should work with the σ-algebra generated by the topology induced by the operator norm?

Note 5. A potential alternative definition would be to consider a matrix with random variables as entries, but then this may present some problems when speaking of the "distribution" of a matrix. A random $m \times n$ matrix can be, maybe, viewed as a map $M : \{1, \ldots, m\} \times \{1, \ldots, n\} \to L_0(\Omega, \mathcal{F}, \mathbb{P})$ where L_0 denotes the set of r.v's on that probability space.

2 Ch1

Here I collect notes and rewrite chapter 1 of [2].

Definition 9. Let $(Y_i : i \in \mathbb{N})$ be a sequence of i.i.d. matrices with distribution μ such that

$$\mathbb{E}[\log^+(\|Y_1\|)] < \infty$$

The limit

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log\|Y_n \cdots Y_1\|$$

is called the *upper Lyapunov exponent.*

Note 6. In what ways does this differ from [1]? First of all, so far in [2], there's no restriction to $\mathrm{SL}_{\pm}(2, \mathbb{R})$, so it's not entirely clear what set is the expectation taken over.

Note 7. "Since all norms on the finite dimensional vector space $\mathcal{M}_{d\times d}(\mathbb{R})$ are equivalent, γ is independent of the chosen norm."

This is stated in [2]. Once again, issues with norm equivalence. Why is this true? The expression for γ depends explicitly on the values of

Example 2. (This is supposedly a "degenerate" example) Consider a sequence of diagonal matrices $(Y_n)_{n\in\mathbb{N}}$

$$Y_n = \begin{bmatrix} a_{11}^{(n)} & 0 & 0 & \cdots & 0 \\ 0 & a_{22}^{(n)} & 0 & \cdots & 0 \\ 0 & 0 & a_{33}^{(n)} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{mk}^{(n)} \end{bmatrix}.$$

Then,

$$\gamma = \sup_i \mathbb{E}\left[\log\left|a_{ii}^{(1)}\right|\right]$$

and

$$\gamma = \lim \frac{1}{n} \log \|Y_n \cdots Y_1\|.$$

This seems like a typo?

2.1 The examples in [1]

Example 3. We take up example (1) in [1]. The set $O(2)$ is defined by

$$O(2) = \left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} : \theta \in [0, 2\pi[\right\}.$$

Consider the bijection $A : [0, 2\pi[\to O(2)$ defined by

$$\theta \mapsto \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Given two matrices in $O(2)$, say $A(\theta_1)$ and $A(\theta_2)$, their product $A(\theta_1)A(\theta_2)$ is also in $O(2)$, because

$$\begin{aligned}
A(\theta_1)A(\theta_2) &= \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 & -(\cos\theta_1\sin\theta_2 + \sin\theta_1\cos\theta_2) \\ \sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2 & -\sin\theta_1\sin\theta_2 + \cos\theta_1\cos\theta_2 \end{bmatrix} \\
&= \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix} \\
&= A(\theta_1+\theta_2 \mod 2\pi).
\end{aligned}$$

We have a correspondence between $O(2)$ and $[0,2\pi[$ thus we can draw random numbers $\theta_1,\dots,\theta_n \in [0,2\pi[$ and consider the product

$$A(\theta_n)\cdots A(\theta_1) = A\left(\sum_{i=1}^{n}\theta_i \mod 2\pi\right).$$

Now observe that given $\theta \in [0,2\pi[$ we have that

$$\begin{aligned}
A(\theta)(x,y)^\mathsf{T} &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}
\end{aligned}$$

this implies that

$$\begin{aligned}
\|A(\theta)(x,y)^\mathsf{T}\| &= \sqrt{(x\cos\theta - y\sin\theta)^2 + (x\sin\theta + y\cos\theta)^2} \\
&= \sqrt{x^2+y^2}.
\end{aligned}$$

Therefore, $\|A(\theta)\| = 1$. Applying this to $\|A(\theta_n)\cdots A(\theta_1)\| = 1$ we get

$$\gamma = \lim_{n\to\infty} \frac{1}{n}\log(1) = 0.$$

Example 4. Assume that only two matrices occur:

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \text{ and } R_{\pi/2} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

If we have a a probability space $(\{0,1\}, \mathcal{P}(\{0,1\}), \mu)$ with such that

$$\mu(X) = \begin{cases} p & \text{if } X = \{0\} \\ 1-p & \text{if } X = \{1\} \end{cases}$$

for some $0 < p < 1$. Let's compute the associated norms:

$$Ax^\mathsf{T} = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ 1/2y \end{bmatrix} \implies \|Ax^\mathsf{T}\| = \sqrt{4x^2 + 1/4y^2}$$

and

$$R_{\pi/2}x^\intercal = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix} \implies \left\| R_{\pi/2}x^\intercal \right\| = \sqrt{y^2 + x^2}$$

Therefore we get

$$\|A\| = \sup\left\{ \sqrt{4x^2 + {}^1\!/\!_4 y^2} : \sqrt{x^2+y^2} \leq 1 \right\} = 2$$

and

$$\left\| R_{\pi/2} \right\| = \sup\{ \sqrt{y^2+x^2} : \sqrt{x^2+y^2} \leq 1 \} = 1.$$

Now we need the following proposition:

Proposition 3. *If* $(X, \|\cdot\|)$ *is a normed space and* $k \in \mathbb{R}_{>0}$ *then the map* $\|\cdot\|_k$ *defined by*

$$\|x\|_k = k\|x\| \qquad (\forall x \in X)$$

is a norm on X.

Proof. We have that

$$\begin{aligned} \|x\|_k = 0 &\Leftrightarrow k\|x\| = 0 \\ &\Leftrightarrow \|x\| = 0 \\ &\Leftrightarrow x = 0. \end{aligned}$$

If α is a scalar, then

$$\begin{aligned} \|\alpha x\|_k &= k\|\alpha x\| \\ &= k|\alpha|\|x\| \\ &= |\alpha|\|x\|_k. \end{aligned}$$

The triangle inequality is satisfied since

$$\|x+y\|_k = k\|x+y\| \leq k(\|x\| + \|y\|) = \|x\|_k + \|y\|_k.$$

□

What we can do is thus define a new norm for the space of matrices. We would like $\|A\| = 1$ to be true, but since it isn't, we define the (equivalent) norm

$$\|X\|_0 = \frac{1}{2}\|X\|$$

for every two-by-two matrix X. Consequently, we now have

$$\|A\|_0 = 1$$

and

$$\|R\|_0 = \frac{1}{2}.$$

Now consider a sequence of random variables $Y_1, \ldots, Y_n$ such that for each $i = 1, \ldots, n$ the matrix Y_i can be equal to either A or R. We want to prove the following inequality

$$\|Y_n \cdots Y_1\|_0 \leq \|Y_n\|_0 \cdots \|Y_1\|_0.$$

Unfortunately, this is not true. We can check that $\left\|A^2\right\|_0 \geq \|A\|_0 \|A\|_0$, and thus this method of proof fails.

Note 8. This actually makes sense. If this proof worked, it would not draw upon any of the probabilistic aspects of the problem, and thus would work as well for the infinite product $\cdots A_3 A_2 A_1$, but for this specific sequence we have $\gamma = \log 2 \neq 0$. Thus, it seems likely that any proof of this example will necessarily include some probability.

References

[1] JAIRO BOCHI. Furstenberg's theorem on products of iid 2 ˆ 2 matrices, 2016.

[2] Philippe Bougerol et al. *Products of random matrices with applications to Schrödinger operators*, volume 8. Springer Science & Business Media, 2012.