UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Teorema de Furstenberg sobre o produto aleatório de matrizes

Herbert Milton Ccalle Maquera

Dissertação de Mestrado do Programa de Pós-Graduação em Matemática (PPG-Mat)

ICMC USP
SÃO CARLOS

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

Herbert Milton Ccalle Maquera

Teorema de Furstenberg sobre o produto aleatório de matrizes

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Matemática. *VERSÃO REVISADA*

Área de Concentração: Matemática

Orientador: Prof. Dr. Ali Tahzibi

USP – São Carlos
Setembro de 2018

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

M297t Maquera, Herbert Milton Ccalle
Teorema de Furstenberg sobre o produto aleatório de matrizes / Herbert Milton Ccalle Maquera; orientador Ali Tahzibi. -- São Carlos, 2018.
91 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Matemática) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2018.

1. Teorema de Furstenberg. 2. Expoente de Lyapunov. 3. Teorema de Oseledets. 4. Cociclos. I. Tahzibi, Ali , orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Herbert Milton Ccalle Maquera

Furstenberg theorem on the random product of matrices

Master dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Mathematics. *FINAL VERSION*

Concentration Area: Mathematics

Advisor: Prof. Dr. Ali Tahzibi

USP – São Carlos
September 2018

AGRADECIMENTOS

Agradeço eternamente a minhas mães Juana e Rosa e meu pai Mauro, pelo carinho, amor e pelo apoio interminável durante todos estos anos de minha vida e que com seguridade seguira sendo assim.

A meus tios Jose, Carlos e Manuel, meus segundos pais, sem cada um de eles eu não estaria hoje aqui finalizando esta etapa de minha vida, obrigado pelo carinho, pela compreensão, pelo constante apoio e por não permitir que tome um caminho errado em minha vida, estarei sempre muito agradecido com vocês.

Em geral agradeço a toda minha familia, sou muito afortunado de ter a vocês.

A minha namorada Xiomara pelo apoio, compreensão e amor em todo esse tempo.

Ao meu orientador Ali Tahzibi, pelos momentos de discussão frutífera e pelo enriquecimento da minha experiencia matemática.

À CAPES pelo suporte financeiro.

RESUMO

MAQUERA, H. M. C. **Teorema de Furstenberg sobre o produto aleatório de matrizes**. 2018. 91 p. Dissertação (Mestrado em Ciências – Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2018.

Nesta dissertação estudamos de um ponto de vista probabilístico, o comportamento assintótico de sistemas dinâmicos. Um exemplo simples de formular e profundo é o estudo de produto aleatório de matrizes (FURSTENBERG; KESTEN, 1960). Utilizaremos como ferramenta o estudo dos cociclos lineares, posteriormente mediante o Teorema de Furstenberg-Kesten definiremos o expoente de Lyapunov do cociclo, em seguida enunciamos e provamos o Teorema Ergódico Multiplicativo de Oseledets o qual nos permite entender o comportamento das órbitas típicas para um cociclo dado $F : M \times \mathbb{R}^2 \longrightarrow M \times \mathbb{R}^2$. O Teorema de Fusrtenberg-Kesten fornece informações sobre o crescimento das matrizes $A^n(x)$, enquanto o Teorema de Oseledets descreve o comportamento assintótico dos vetores $A^n(x) \cdot v$. Finalmente provamos o teorema principal desta dissertação, o Teorema de Furstenberg o qual diz que na maioria dos casos o maior expoente de Lyapunov é positivo (FURSTENBERG, 1963).

Palavras-chave: Produto aleatório de matrizes, medidas estacionárias, expoentes de Lyapunov.

ABSTRACT

MAQUERA, H. M. C. **Furstenberg theorem on the random product of matrices**. 2018. 91 p. Dissertação (Mestrado em Ciências – Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2018.

In this thesis we study from a probabilistic point of view, the asymptotic behavior of dynamic systems, a deep and simple example is the random product of matrices (FURSTENBERG; KESTEN, 1960). We will use as a tool, the study of linear cocycles, later using the Furstenberg-Kesten Theorem we will define the Lyapunov exponent of the cocycle, then we enunciate and prove the Multiplicative Ergodic Theorem of Oseledets which allow us to understand the behavior of the typical orbits for a given cocycle $F : M \times \mathbb{R}^2 \longrightarrow M \times \mathbb{R}^2$. The Fusrtenberg-Kesten theorem provides information on the growth of the matrices $A(x)$, while the theorems of Oseledets describe the asymptotic behavior of the vectors $A^n(x) \cdot v$. Finally we prove our main theorem, Furstenberg's Theorem which states that in most cases the greatest exponent of Lyapunov is positive (FURSTENBERG, 1963).

Keywords: Random product of matrices, stationary measures, Lyapunov exponent.

LISTA DE ILUSTRAÇÕES

Figura 1 – Vetores ortonormais $s(x)$ e $u(x)$ 69
Figura 2 – Imagem por $A^{n+1}(x)$ da base ortonormal $\{s_{n+1}, u_{n+1}\}$. 70
Figura 3 – Vetores unitários $s_n(x)$ e $s_{n+1}(x)$. 71
Figura 4 – Interpretação geométrica da desigualdade 75
Figura 5 – Ação projetiva. 76

SUMÁRIO

1 INTRODUÇÃO 15

2 PRELIMINARES 19
2.1 Decomposição de $\mathrm{SL}(2,\mathbb{R})$ 20
2.2 Cociclos lineares 21
2.3 Convolução de medidas 25
2.4 Espaços de probabilidade 26
2.5 Variáveis aleatórias 26
2.6 Indepêndencia 28
2.7 Esperança condicional 29
2.8 Martingales 30

3 MEDIDAS ESTACIONARIAS 35
3.1 Medidas estacionárias 38

4 TEOREMA DE FURSTENBERG 43
4.1 O maior expoente de Lyapunov 43

5 TEOREMA ERGÓDICO SUBADITIVO DE KINGMAN 53
5.1 Lemas Fundamentais 55
5.2 Desigualdade Inferior 59
5.3 Desigualdade Superior 61

6 TEOREMA ERGÓDICO MULTIPLICATIVO DE OSELEDETS . . . 67

7 TEOREMA PRINCIPAL 79
7.1 Medidas não atômicas 80
7.2 Convergencia do Push-forward de medidas 83
7.3 Demostração do Teorema de Furstenberg 89

REFERÊNCIAS 91

CAPÍTULO 1

INTRODUÇÃO

A lei clássica dos grandes números afirma que para variáveis aleatórias independentes e identicamente distribuídas Y_i com $\mathbb{E}(|Y_i|) < \infty$ temos que

$$\lim_n \frac{1}{n} \sum_{i=1}^{n} Y_i = \mathbb{E}(Y_1) \quad \text{quase sempre.} \tag{1.1}$$

A primeira versão foi provada por Jacob Bernoulli por volta de 1700 em seu *Ars Conjectandi* onde se dá a primeira versão da lei de grandes números. A versão moderna é consequência dos esforços de Chebyshev, Markov, Borel, Cantelli e Kolmogorov.

Posteriormente no início da década de 1930, Von Neumann provou o primeiro teorema ergódico (após o teorema de recorrência de Poincaré) e imediatamente depois Birkhoff provou seu teorema ergódico pontual. Considere (Ω, μ) um espaço de probabilidade e $T : \Omega \to \Omega$ uma transformação mensurável que preserva medida, isto é $\mu(T^{-1}A) = \mu(A)$, e ergódica (isto é $\mu(T^{-1}A \triangle A) = 0$ implica $\mu(A) = 0$ ou $\mu(A) = 1$). O teorema de Birkhoff diz que dado $f : \Omega \to \mathbb{R}$ integrável em relação à medida μ, isto é

$$\int_\Omega |f| \, d\mu < \infty, \tag{1.2}$$

então

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} (f \circ T^i)(\omega) = \int_\Omega f \, d\mu \quad \text{quase sempre.} \tag{1.3}$$

Por outro lado a lei dos grandes números é um caso especial. Considere $\Omega = \mathbb{R}^{\mathbb{N}}$ e $\mu = \nu^{\mathbb{N}}$. Além disso, considere a transformação $T : \Omega \to \Omega$ como sendo $(\omega_n)_{n\in\mathbb{N}} \mapsto (\omega_{n+1})_{n\in\mathbb{N}}$ o qual é ergódica. A função f é a projeção na primeira coordenada de modo que para $\omega = (\omega_1, \omega_2, \dots)$ temos que $Y_i(\omega) = (f \circ T^{i-1})(\omega) = \omega_i$.

Uma versão não-comutativa do teorema ergódico significa que a função f agora assume valores em um grupo mais geral G ao invés da reta real. Seja G um grupo topológico. Dado uma aplicação mensurável $f : \Omega \to G$, escreva

$$S_n(\omega) = f(\omega)f(T(\omega))\dots f(T^{n-1}(\omega))$$

que é a extensão da soma do lado esquerdo em (1.3) para um grupo geral, possivelmente não comutativo. A primeira questão surge aquí: assumindo alguma condição de integrabilidade que seja apropriada, o que se pode dizer sobre S_n quando n tende ao infinito?. As primeiras pessoas a tentar responder a essa pergunta foram Bellman, Kesten e Furstenberg, da década de 1950 em diante. Para fixar a terminologia, no caso geral chamamos de um cociclo ergódico a S_n e no caso especial das variáveis aleatórias independentes e identicamente distribuídas chamamos a S_n de passeio aleatório.

Um dos primeiros estudos para o caso de produtos de matrizes aleatórias é um artigo de Furstenberg-Kesten em 1960, onde eles em particular provam que se S_n é um cociclo ergódico, integrável no sentido que

$$\int \log \|f\| \, d\mu < \infty$$

onde $\|\cdot\|$ é alguma norma matricial, então o seguinte limite existe quase sempre:

$$\lim_n \frac{1}{n} \log \|S_n\| = \gamma.$$

Isso foi posteriormente generalizado pelo Teorema de Oseledets e o Teorema ergódico subaditivo de Kingman. Em um artigo publicado em 1968, Kingman em resposta a uma pergunta de Hammersley, provou o que hoje se conhece como o Teorema ergódico subaditivo, que é uma extensão do teorema de Birkhoff o qual é de fundamental importância em diversos contextos. Em termos simples este teorema garante a existência em quase todos os pontos das médias associadas a sucessões subaditivas, isto é: seja $f_n(\omega)$ uma sequência de funções mensuráveis tais que para $m, n \geqslant 0$ vale a seguinte desigualdade

$$f_{m+n}(\omega) \leqslant f_n(\omega) + f_m(T^n(\omega))$$

e

$$\int_\Omega f_1(\omega) \, d\mu(\omega) < \infty,$$

então o seguinte limite existe quase sempre:

$$\lim_n \frac{1}{n} f_n(\omega) = \gamma.$$

Mais tarde, demonstrações mais simples do referido resultado surgiram tais como B. Weiss, J. Steele, K. Schüger e A. Ávila com J. Bochi cuja demonstração é independente do teorema ergódico o qual é uma questão de interesse para este trabalho. Assim, o teorema de Furstenberg-Kesten é um caso particular considerando

$$f_n(\omega) = \log \|S_n(\omega)\|.$$

No ano de 1968, ainda jovem o matemático V. Oseledets aluno de doutorado do Y. Sinai publicou o Teorema Ergódico Multiplicativo que foi resultado de sua tese de doutorado, a motivação deste resultado deve-se aos trabalhos no contexto da teoria das equações diferenciais ordinárias mais precisamente à estabilidade das soluções das referidas equações estudadas por A. Lyapunov. Existem também para este resultado diferentes versões tais como M. Raghunathan, o qual é baseado no teorema ergódico subaditivo e os valores singulares de uma matriz, para espaços de Hilbert também temos una versão dada por D. Ruelle; para produtos de operadores em espaços de Banach dado por P. Thieullen também a versão dada por R. Mañé para isomorfismos de fibrados de dimensão finita e P. Walters para cociclos lineares projetivos. Também de um ponto de vista mais geométrico temos ao matemático V. Kaimanovich para certos espaços de curvatura não positiva e F. Ledrappier para certos grupos de isometrías.

CAPÍTULO

2

PRELIMINARES

Neste capítulo fazemos um breve estudo de algumas propriedades e resultados elementares sobre o grupo de matrizes. Seja $M(d,\mathbb{R})$ o conjunto de matrizes com entradas reais. Denotamos o conjunto das matrizes reais $d \times d$ com determinante 1 por $SL(d,\mathbb{R})$, o qual é chamado de grupo linear especial. O grupo linear especial $SL(d,\mathbb{R})$ é subgrupo de $GL(d,\mathbb{R})$, chamado de grupo linear geral (conjunto das matrizes invertíveis). Além disso, M^* denota a matriz transposta de M. A maioria dos resultados que apresentamos não depende de uma norma particular em $\mathbb{R}$ e $M(d,\mathbb{R})$ (BOUGEROL *et al.*, 2012). Por conveniência, consideramos

$$\|x\| = \left(\sum_{i=1}^{d} x_i^2\right)^{\frac{1}{2}} \quad \text{e} \quad \|A\| = \sup\{\, \|Ax\| \; ; \; x \in \mathbb{R}^d, \|x\| = 1\}$$

para $x \in \mathbb{R}^d$ e $A \in M(d,\mathbb{R})$ respectivamente.

Primeiro, vamos formular um lema com propriedades básicas das matrizes em $SL(2,\mathbb{R})$.

Lema 1. Seja $M \in SL(2,\mathbb{R})$. Então

(i) $M^*JM = J$ onde $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

(ii) $\|M\| = \|M^{-1}\| \geqslant 1$.

(iii) Se λ é um autovalor de M, então também λ^{-1} é um autovalor de M.

Demonstração. (i) A equação resulta de um cálculo direto.

(ii) Uma vez que $\|J\| = \|J^{-1}\| = 1$ e M é inversível, logo pelo item i),

$$\|M^{-1}\| = \|J^{-1}M^*J\| \leqslant \|M\|$$

e

$$\|M\| = \|J^{-1}(M^{-1})^* J\| \leqslant \|M^{-1}\|$$

Além disso,

$$1 = \|MM^{-1}\| \leqslant \|M\|\|M^{-1}\| = \|M\|^2$$

e com isso

$$\|M\| = \|M^{-1}\| \geqslant 1.$$

(iii) Sejam λ e μ os valores próprios de M. Uma vez que $1 = \det M = \lambda\mu$, $\mu = \lambda^{-1}$.

□

2.1 Decomposição de $SL(2,\mathbb{R})$

Apresentamos nesta seção as noções básicas e notações relativas ao grupo $SL(2,\mathbb{R})$, muitas das quais são bem conhecidas. Por conveniência, notamos as formas explícitas de K, A e N.

Vamos primeiro identificar alguns subgrupos padrões de $SL(2,\mathbb{R})$:

$$K = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in [0,2\pi] \right\};$$

$$N = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} : x \in \mathbb{R} \right\};$$

$$A = \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} : a > 0 \right\}.$$

Há duas importantes decomposições para $SL(2,\mathbb{R})$ por meio dos subgrupos K, A e N. Aqui utilizaremos apenas a decomposição de Cartan.

Teorema 1. *[Decomposição de Cartan] Qualquer matriz g no grupo $SL(2,\mathbb{R})$ é escrita como $g = ha(g)k$, com*

(i) $k, h \in K$

(ii) $a(g) = \mathrm{diag}(a_1(g), a_2(g)) \in A$ *e* $a_1(g) \geqslant a_2(g) > 0$.

Além disso, $a(g)$ é único e os $a_i(g)$ são as raízes quadradas dos autovalores de gg^t. Em particular, $a_1(g) = \|g\|$ e $a_2(g) = \|g^{-1}\|^{-1}$. Essa decomposição é chamada de decomposição de Cartan da matriz g.

Demonstração. Se $g = ha(g)k$, então

$$gg^t = (ha(g)k)(ha(g)k)^t = ha(g)kk^ta^t(g)h^t = ha^2(g)h^t = ha^2(g)h^{-1}$$

daí a segunda parte do teorema segue. Além disso, gg^t é simétrico e positivo definido. Essa matriz admite, portanto, uma base ortonormal de autovetores ordenados pela diminuição dos autovalores. De modo matricial, isso é escrito como $gg^t = ha^2h^{-1}$ com $h \in K$ e a como desejado. Logo, defina $k = a^{-1}h^{-1}g$ e temos que $g = hak$ e $k \in K$ pois $k^tk = Id$. □

2.2 Cociclos lineares

Dado um sistema dinâmico (T, M, μ). Um *Skew product* sobre T é um sistema dinâmico F agindo no espaço produto $M \times N$ tal que o seguinte diagrama comuta

$$\begin{array}{ccc} M \times N & \xrightarrow{F} & M \times N \\ \pi \downarrow & & \downarrow \pi \\ M & \xrightarrow{T} & M \end{array}$$

isto é, $\pi \circ F = T \circ \pi$, onde $\pi : M \times N \to M$ é a projeção na primeira coordenada. Nosso interesse é focado em skew products que agem linearmente na segunda coordenada, assim N precisa ser um espaço vetorial. Tais aplicações serão chamadas de cociclos lineares.

No que segue apresentaremos alguns conceitos e propriedades de cociclos lineares seguindo Viana (VIANA, 2014).

Definição 1. Dado um sistema dinâmico (T, M, μ) e uma aplicação mensurável $A : M \to \mathrm{SL}(2,\mathbb{R})$, podemos definir um $\mathrm{SL}(2,\mathbb{R})$-*cociclo linear* sobre (T, M, μ) como sendo a transformação

$$\begin{array}{rccc} F: & M \times \mathbb{R}^2 & \longrightarrow & M \times \mathbb{R}^2 \\ & (x, v) & \mapsto & (Tx, A(x) \cdot v). \end{array}$$

Como o par (T, A) especifica $F = F_{T,A}$, nós também chamamos de cociclo. As vezes nós também chamamos o mapa A de cociclo. As iterações podem ser escritas como

$$F^n(x, v) = (T^n x, A^n_T(x) \cdot v), \tag{2.1}$$

onde

$$A^n_T(x) = A(T^{n-1}x)A(T^{n-2}(x))\ldots A(x) \quad \text{para } n > 0.$$

No caso que T é invertível, F também será, então podemos definir para $n \in \mathbb{N}$

$$A^{-n}_T(x) = [A(T^{-n}x)A(T^{-n+1}x)\ldots A(T^{-1}x)]^{-1} \quad \text{para } n > 0.$$

Logo temos que

$$\begin{aligned} A^m_T(T^n x)A^n_T(x) &= A(T^{m+n-1}x)A(T^{m+n-2}(x))\ldots A(T^n x)A(T^{n-1}x)A(T^{n-2}(x))\ldots A(x) \\ &= A(T^{m+n-1}x)A(T^{m+n-2}(x))\ldots A(x) \\ &= A^{m+n}_T(x). \end{aligned}$$

Portanto, temos a seguinte identidade que vamos chamar de *identidade de cociclos:*

$$A_T^m(T^n x)A_T^n(x) = A_T^{m+n}(x) \quad \text{para todo } x \in M,\ m, n \in \mathbb{Z}.$$

Observação 1. A maior parte do tempo T será fixado e nós escrevemos por simplicidade $A^n(x)$ ao invés de $A_T^n(x)$.

O teorema de Oseledets se aplicará a cociclos em que a base preserva uma medida de probabilidade μ. Então, sempre iremos supor que na base temos um sistema dinâmico (T, M, μ), e uma vez que só estamos interessados no comportamento assintótico de $A^n(x)$ e $A^{-n}(x)$ no caso que T seja invertível, tudo isto apenas num conjunto de medida total, nós apenas assumiremos que a aplicação A é mensurável.

Exemplo 1. Um problema muito comum em probabilidade é o produto aleatório de matrizes. Considere um conjunto de matrizes com uma lei de probabilidade. Começamos a tomar matrizes desse conjunto de forma aleatória e fazemos o produto de elas. Em sistemas dinâmicos o objetivo é saber o que vai acontecer com os produtos matriciais. Isto pode ser modelado da seguinte maneira. Seja $M = \mathrm{GL}(d, \mathbb{R})^{\mathbb{Z}}$, a transformação de deslocamento

$$\begin{array}{rccc} T: & M & \longrightarrow & M \\ & (\alpha_k)_k & \mapsto & (\alpha_{k+1})_k, \end{array}$$

considere a aplicação

$$\begin{array}{rccc} A: & M & \longrightarrow & \mathrm{GL}(d, \mathbb{R}) \\ & (\alpha_k)_k & \mapsto & \alpha_0 \end{array}$$

e seja $F : M \times \mathbb{R}^d \to M \times \mathbb{R}^d$ o cociclo linear definido por A sobre T. Observe que a k-ésima iteração de F é dada por

$$F^n((\alpha_k)_k, v) = ((\alpha_{k+n})_k, \alpha_{n-1} \ldots \alpha_1 \alpha_0 v).$$

Dada uma medida de probabilidade ν no espaço $\mathrm{GL}(d, \mathbb{R})$, podemos considerar a medida produto $\mu = \nu^{\mathbb{Z}}$, que é caracterizada por

$$\mu(\{(\alpha_k)_k \ : \ \alpha_i \in E_i, \ldots, \alpha_j \in E_j\}) = \nu(E_i) \ldots \nu(E_j)$$

para cada $i \leqslant j$ e quaisquer conjuntos mensuráveis $E_i, \ldots, E_j \subset \mathrm{GL}(d, \mathbb{R})$. É claro que a medida μ é invariante pela transformação de deslocamento.

Exemplo 2. Chamamos de *cociclo linear localmente constante* à seguinte construção um pouco mais geral. Seja $(Y, \mathcal{Y}, \eta)$ um espaço de probabilidade, então considere $N = Y^{\mathbb{Z}}$ dotado com a σ-álgebra produto $\mathcal{C} = \mathcal{Y}^{\mathbb{Z}}$, a medida produto $\lambda = \eta^{\mathbb{Z}}$. Considere também a transformação de deslocamento $g : N \to N$, além disso, seja $B : N \to \mathrm{GL}(d, \mathbb{R})$ qualquer função mensurável, dependendo apenas da coordenada zero, isto é, da forma $B(y) = \beta(y_0)$ para alguma função mensurável $\beta : Y \to \mathrm{GL}(d, \mathbb{R})$. Considere o cociclo linear $G : N \times \mathbb{R}^d \to N \times \mathbb{R}^d$ definido por B sobre g.

Observação 2. Observe que os cociclos G e F nos exemplos anteriores são semi conjugados, com $\nu = \beta_* \eta$

$$\begin{array}{ccc} N \times \mathbb{R}^d & \xrightarrow{G} & N \times \mathbb{R}^d \\ \psi \big\downarrow & & \big\downarrow \psi \\ M \times \mathbb{R}^d & \xrightarrow{F} & M \times \mathbb{R}^d \end{array}$$

onde $\psi((x_k)_k, v) = ((\beta(x_k))_k, v)$.

Exemplo 3 (Cociclo derivado)**.** Considere um difeomorfismo $f : M \to M$ no toro $M = \mathbb{T}^2$ que preserva área e orientação. Sabemos que existem campos vetoriais suaves X_1, X_2 em $\mathbb{T}^2$ de modo que $\{X_1(x), X_2(x)\}$ é uma base para o espaço tangente $T_x M$ para cada $x \in M$. Então podemos definir o seguinte cociclo: seja μ a medida de Lebesgue em M, defina $A(x)$ como a transformação linear $Df_x : T_x M \to T_{f(x)} M$ e $A : M \to SL(2, \mathbb{R})$, e

$$F : M \times \mathbb{R}^2 \to M \times \mathbb{R}^2, \quad F(x, v) = (f(x), A(x) \cdot v).$$

Como f preserva área e orientação, então $Df_x \in SL(2, \mathbb{R})$ para todo $x \in M$ assim, A está bem definida. Portanto estudar a dinâmica do cociclo F é estudar a dinâmica da derivada Df_x. A regra da cadeia nos permite obter a identidade do cociclo.

Dada uma ação podemos definir um cociclo associado a ela.

Definição 2. Seja G um semigrupo topológico, dizemos que G *age* sobre um espaço topológico M se podemos associar continuamente a cada (g, x) em $G \times M$ um elemento $g \cdot x$ de M. Dito de outra maneira, se existe uma função contínua

$$\begin{aligned} \varphi \colon G \times M &\longrightarrow M \\ (g, x) &\longmapsto \varphi(g, x) = g \cdot x \end{aligned}$$

tal que

$$\varphi(g_1 g_2, x) = \varphi(g_1, \varphi(g_2, x)) \tag{2.2}$$

para todo $g_1, g_2 \in G$ e $x \in M$.

Se G for um grupo topológico com elemento identidade e e além disso $e \cdot x = x$ para todo $x \in M$, então dizemos que M é um G-espaço.

Exemplo 4. Considere o grupo $GL(d, \mathbb{R})$ de matrizes invertíveis e a esfera unitária $\mathbb{S} = \{x \in \mathbb{R}^d : \|x\| = 1\}$. Então podemos definir a seguinte aplicação

$$\begin{aligned} \varphi \colon GL(d, \mathbb{R}) \times \mathbb{S} &\longrightarrow \mathbb{S} \\ (g, x) &\longmapsto \varphi(g, x) = \frac{gx}{\|gx\|}. \end{aligned}$$

A aplicação φ é uma ação: de fato, sejam g e h matrizes em $GL(d,\mathbb{R})$ e x em $\mathbb{S}$, então

$$\begin{aligned}\varphi(gh,x) &= \frac{ghx}{\|ghx\|}\\ &= \frac{g\varphi(h,x)}{\|g\varphi(h,x)\|}\\ &= \varphi(g,\varphi(h,x)).\end{aligned}$$

Portanto $\varphi(gh,x)=\varphi(g,\varphi(h,x))$, isto é, φ é uma ação.

Definição 3. Seja G um semigrupo topológico agindo em um espaço M, uma aplicação contínua $\sigma\colon G\times M\to\mathbb{R}$ é dita de *cociclo aditivo* se

$$\sigma(g_1g_2,x)=\sigma(g_1,g_2\cdot x)+\sigma(g_2,x)$$

para todo g_1,g_2 em G e para todo x em M.

Exemplo 5. No exemplo 4 podemos definir a seguinte aplicação

$$\begin{aligned}\sigma\colon GL(d,\mathbb{R})\times\mathbb{S} &\longrightarrow \mathbb{R}\\ (g,x) &\longmapsto \sigma(g,x)=\log\|gx\|.\end{aligned}$$

É fácil verificar que σ é um cociclo aditivo. De fato, sejam g e h matrizes em $GL(d,\mathbb{R})$ e x em $\mathbb{S}$, então

$$\begin{aligned}\sigma(gh,x) &= \log\|ghx\|\\ &= \log\left\|\frac{ghx}{\|hx\|}\|hx\|\right\|\\ &= \log\left\|\frac{ghx}{\|hx\|}\right\|+\log\|hx\|\\ &= \log\|g\varphi(h,x)\|+\log\|hx\|\\ &= \sigma(g,\varphi(h,x))+\sigma(h,x).\end{aligned}$$

Definição 4. Seja G um semigrupo topológico agindo em um espaço M. Uma aplicação contínua $\sigma\colon G\times M\to\mathbb{R}$ é dita *cociclo multiplicativo* se

$$\sigma(g_1g_2,x)=\sigma(g_1,g_2\cdot x)\sigma(g_2,x)$$

para todo g_1,g_2 em G e para todo x em M.

Exemplo 6. A partir do exemplo 5 podemos definir a seguinte aplicação

$$\begin{aligned}\sigma\colon GL(d,\mathbb{R})\times\mathbb{S} &\longrightarrow \mathbb{R}\\ (g,x) &\longmapsto \sigma(g,x)=\|gx\|.\end{aligned}$$

Como no exemplo anterior, é fácil verificar que σ é um cociclo multiplicativo. Sejam g e h matrizes em $GL(d,\mathbb{R})$ e x em $\mathbb{S}$, então

$$\begin{aligned}
\sigma(gh,x) &= \|ghx\| \\
&= \left\| \frac{ghx}{\|hx\|}\|hx\| \right\| \\
&= \left\| \frac{ghx}{\|hx\|} \right\| \|hx\| \\
&= \|g\varphi(h,x)\|\|hx\| \\
&= \sigma(g,\varphi(h,x))\sigma(h,x).
\end{aligned}$$

Observação 3. É possível transformar um cociclo multiplicativo em um cociclo aditivo e vice-versa através da função logaritmo e exponencial respectivamente.

2.3 Convolução de medidas

Definição 5. Sejam μ e ν medidas de probabilidade no semigrupo topológico G e no espaço topológico M respectivamente. A *convolução das medidas* μ e ν é denotada por $\mu * \nu$ e é uma medida de probabilidade em M que satisfaz

$$\int_M f(x)\, d(\mu*\nu)(x) = \int_G \int_M f(g\cdot x)\, d\mu(g)\, d\nu(x),$$

para toda aplicação f mensurável e limitada.

Definição 6. Se G for um grupo topológico então a convolução de μ e ν é definida por

$$(\mu*\nu)(A) = \int_G \nu(g^{-1}A)\, d\mu(g)$$

onde $A \subset M$ mensurável e $g \in G$.

Observação 4. Considere f como sendo a função característica de $A \subset M$ um subconjunto mensurável, isto é $f(x) = \chi_A(x)$, logo

$$\begin{aligned}
\int_M f(x)\, d(\mu*\nu)(x) &= \int_M \chi_A(x)\, d(\mu*\nu)(x) \\
&= \int_A d(\mu*\nu)(x) \\
&= (\mu*\nu)(A)
\end{aligned}$$

Por outro lado:

$$\begin{aligned}
\int_G \int_M f(g\cdot x)\, d\mu(g)\, d\nu(x) &= \int_G \int_M \chi_A(g\cdot x)\, d\mu(g)\, d\nu(x) \\
&= \int_G \int_M \chi_{g^{-1}A}(x)\, d\mu(g)\, d\nu(x) \\
&= \int_G \int_{g^{-1}A} d\mu(g)\, d\nu(x) \\
&= \int_G \nu(g^{-1}A)\, d\mu(g).
\end{aligned}$$

Portanto, as definições acima são equivalentes.

2.4 Espaços de probabilidade

Definição 7. Um *espaço de probabilidade* é una tripla $(\Omega, \mathcal{F}, \mathbb{P})$, onde Ω é um conjunto não vazio, chamado de espaço de amostra, $\mathcal{F}$ é uma σ-álgebra de subconjuntos de Ω, cujos elementos se denominam eventos e $\mathbb{P}$ é uma medida de probabilidade em $\mathcal{F}$, isto é, $\mathbb{P}$ é uma medida em $\mathcal{F}$ e $\mathbb{P}(\Omega) = 1$ (em particular, $\mathbb{P}$ é uma medida finita). Para qualquer evento $B \in \mathcal{F}$, $\mathbb{P}(B)$ é chamado de probabilidade de B.

2.5 Variáveis aleatórias

Dados dois espaços mensuráveis $(Z, \mathcal{F})$ e $(Y, \mathcal{G})$, dizemos que uma função $\psi : (Z, \mathcal{F}) \to (Y, \mathcal{G})$ é mensurável se preserva a estrutura de mensurabilidade. Em teoria de probabilidade, funções mensuráveis recebem o nome de variáveis aleatórias.

Definição 8. Dado um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, uma função $X : \Omega \to Y$ é chamada de *variável aleatória* a valores no espaço mensurável $(Y, \mathcal{G})$ se para todo $A \in \mathcal{G}$, temos que $X^{-1}(A) \in \mathcal{F}$.

Observação 5. Denotaremos por $\sigma\langle X\rangle$ a σ-álgebra gerada por X, isto é, a menor σ-álgebra que faz com que X seja mensurável, $\sigma\langle X\rangle = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$.

A mensurabilidade de variáveis aleatórias permite definir uma medida no espaço mensurável $(Y, \mathcal{G})$. Tal medida é conhecida como distribuição.

Definição 9. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de medida, $(Y, \mathcal{G})$ espaço mensurável e $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (Y, \mathcal{G})$ variável aleatória, então para todo $E \subset Y$, $E \in \mathcal{G}$ definimos

$$X_*\mathbb{P}(E) = (\mathbb{P} \circ X^{-1})(E) = \mathbb{P}\{X \in E\}.$$

Chamamos de *medida induzida*, *distribuição* ou ainda *lei* da variável aleatória X.

Definição 10. Seja X uma variável aleatória definida em $(\Omega, \mathcal{F}, \mathbb{P})$. A *esperança* ou *média* é a integral

$$\mathbb{E}(X) = \int_\Omega X \, d\mathbb{P}.$$

Observação 6. A esperança também é chamada de valor esperado da variável aleatória X. Note que o valor esperado representa o sentido mais intuitivo que temos de média , isto é, soma dividido pelo total. Aquí a soma é ponderada pelas medidas dos conjuntos mensuráveis e o total neste caso foi normalizado em um.

Teorema 2 (Teorema da covergência monótona)**.** *Seja* $f_n : \Omega \to \mathbb{R}$ *para todo* $n \geqslant 1$ *uma sequência de funções integráveis tal que para* μ*-quase todo ponto* $x \in \Omega$

(i) $f_n(x) \leqslant f_{n+1}(x)$ *para todo* $n \in \mathbb{N}$ *e*

(ii) $\sup_n \int f_n(x)\, d\mu(x) < \infty$.

Então $f(x) = \lim_n f_n(x)$ *é uma função integrável e*

$$\int f(x)\, d\mu(x) = \lim_n \int f_n(x)\, d\mu(x).$$

Demonstração. Ver (FOLLAND, 2013). □

Teorema 3 (Lema de Fatou)**.** *Seja* $\{f_n\}$ *uma sequência de funções não negativas e integráveis,* $f_n : \Omega \to \mathbb{R}$ *tal que*

(i) $\liminf_n \int f_n(x)\, d\mu(x) < \infty$

(ii) $\lim_n f_n(x) = f(x)$ μ*-quase todo ponto.*

Então f *é integrável e*

$$\int f(x)\, d\mu(x) \leqslant \liminf_n \int f_n(x)\, d\mu(x).$$

Demonstração. Ver (FOLLAND, 2013). □

Teorema 4 (Teorema da covergência dominada)**.** *Seja* $\{f_n\}$ *uma sequência de funções integráveis* $f_n : \Omega \to \mathbb{R}$ *e seja* $g : \Omega \to \mathbb{R}$ *uma função integrável tal que*

(i) $|f_n(x)| \leqslant g(x)$ μ*-quase todo ponto, para todo* $n \in \mathbb{N}$ *e*

(ii) $\lim_n f_n(x) = f(x)$ μ*-quase todo ponto.*

Então f *é integrável e*

$$\lim_n \int f_n(x)\, d\mu(x) = \int f(x)\, d\mu(x).$$

Demonstração. Ver (FOLLAND, 2013). □

Teorema 5. *Seja* $(\Omega, \mathcal{F}, \mu)$ *um espaço de medida,* $(Y, \mathcal{G})$ *um espaço mensurável,* $X : \Omega \to Y$ *e* $f : Y \to \mathbb{R}$*, funções mensuráveis, então*

$$\int_Y f\, dX_*\mu = \int_\Omega (f \circ X)\, d\mu.$$

Demonstração. Considere o caso de f ser uma função simples (no caso geral tomamos aproximações de uma função por funções simples). Considere f como sendo

$$f(y) = \sum_{j=1}^{n} a_j \chi_{A_j}(y)$$

onde $A_1, A_2, \ldots, A_n$ são $\mathcal{G}$-mensuráveis, assim temos que

$$\begin{aligned}
\int_Y f \; dX_*\mu = \int_Y \sum_{j=1}^{n} a_j \chi_{A_j}(y) \; dX_*\mu(y) &= \sum_{j=1}^{n} a_j \int_Y \chi_{A_j}(y) \; dX_*\mu(y) \\
&= \sum_{j=1}^{n} a_j X_*\mu(A_j) \\
&= \sum_{j=1}^{n} a_j \mu(X^{-1}(A_j)) \\
&= \int_\Omega (f \circ X)(x) \; d\mu(x).
\end{aligned}$$

□

2.6 Indepêndencia

Embora um espaço de probabilidade não seja nada mais que um espaço de medida com a medida de todo o espaço igual a um, a teoria da probabilidade não é meramente um subconjunto da teoria da medida. Uma característica distintiva e fundamental da teoria da probabilidade é a noção de independência.

Definição 11. Dado um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, dois eventos A e B em $\mathcal{F}$ são chamados de *independentes* em relação à medida $\mathbb{P}$ se $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Generalizando a definição acima, temos o seguinte.

Definição 12. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ espaço de probabilidade e $\{B_1, B_2, \ldots, B_n\} \subset \mathcal{F}$ uma coleção finita de eventos.

(i) $B_1, B_2, \ldots, B_n$ são chamados de *independentes* com relação à medida $\mathbb{P}$ se

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{i_j}\right) = \prod_{j=1}^{k} \mathbb{P}(B_{i_j})$$

para todo $\{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$, $1 \leqslant k \leqslant n$.

(ii) $B_1, B_2, \ldots, B_n$ são chamados de *pares independentes* com relação à medida $\mathbb{P}$ se $\mathbb{P}(B_i \cap B_j) = \mathbb{P}(B_i)\mathbb{P}(B_j)$, para todo $i \neq j$.

Definição 13. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade. Uma coleção de eventos $\{B_\alpha : \alpha \in A\} \subset \mathcal{F}$ é chamada independente com relação à medida $\mathbb{P}$ se para cada sub coleção finita $\{\alpha_1, \alpha_2, \ldots, \alpha_k\} \subset A$, $1 \leqslant k < \infty$, temos que

$$\mathbb{P}\left(\bigcap_{j=1}^{k} B_{\alpha_j}\right) = \prod_{j=1}^{k} \mathbb{P}(B_{\alpha_j}).$$

Definição 14. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade e seja A um conjunto não vazio. Para cada $\alpha \in A$, seja $\mathcal{G}_\alpha \subset \mathcal{F}$ uma coleção de eventos. Então a família $\{\mathcal{G}_\alpha : \alpha \in A\}$ é chamada independente com relação à medida $\mathbb{P}$ se para cada escolha de B_α em $\mathcal{G}_\alpha$, a coleção de eventos $\{B_\alpha : \alpha \in A\}$ é independente com relação à medida $\mathbb{P}$ como na definição 13.

Definição 15. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade e seja $\{X_\alpha : \alpha \in A\}$ uma coleção de variáveis aleatórias em $(\Omega, \mathcal{F}, \mathbb{P})$. Então a coleção $\{X_\alpha : \alpha \in A\}$ é chamada de *independente* com relação à medida $\mathbb{P}$ se a família de σ-álgebras $\{\sigma\langle X_\alpha\rangle : \alpha \in A\}$ é independente com relação à medida $\mathbb{P}$, onde $\langle X_\alpha\rangle$ é a σ-álgebra gerada por X_α, isto é $\sigma\langle X_\alpha\rangle = \{X_\alpha^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$.

2.7 Esperança condicional

Definição 16. Seja X uma variável aleatória integrável em $(\Omega, \mathcal{F}, \mathbb{P})$ e seja $\mathcal{G}$ uma sub σ-álgebra de $\mathcal{F}$. A *esperança condicional* $\mathbb{E}(X|\mathcal{G})$ é uma variável aleatória $\mathcal{G}$-mensurável que satisfaz

$$\int_G X \, d\mathbb{P} = \int_G \mathbb{E}(X|\mathcal{G}) \, d\mathbb{P} \tag{2.3}$$

para todo conjunto $G \in \mathcal{G}$.

Podemos interpretar $\mathbb{E}(X|\mathcal{G})$ como uma estimativa da variável aleatória X tendo em conta a informação dada pela σ-álgebra $\mathcal{G}$. A expressão em (2.3) diz que as médias de X e $\mathbb{E}(X|\mathcal{G})$ são as mesmas em todo conjunto de $\mathcal{G}$.

Teorema 6. *Sejam X e Y variáveis aleatórias integráveis em $(\Omega, \mathcal{F}, \mathbb{P})$ e seja $\mathcal{G}$ uma sub σ-álgebra de $\mathcal{F}$. Então as seguintes propriedades são válidas.*

(i) *(Invariância da média)* $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$.

(ii) *(Propriedade de projeção) Se X é $\mathcal{G}$-mensurável, então* $\mathbb{E}(X|\mathcal{G}) = X$.

(iii) *(Linearidade) Se $a, b \in \mathbb{R}$, então*

$$\mathbb{E}(aX + bY|\mathcal{G}) = a\mathbb{E}(X|\mathcal{G}) + b\mathbb{E}(Y|\mathcal{G}).$$

(iv) *(Monotonicidade) Se $X \leqslant Y$ em quase todo ponto, então $\mathbb{E}(X|\mathcal{G}) \leqslant \mathbb{E}(Y|\mathcal{G})$ em quase todo ponto.*

(v) *($\mathcal{G}$-mensuráveis agem como constantes) Se* X *é* $\mathcal{G}$*-mensurável e* XY *é integrável, então* $\mathbb{E}(XY|\mathcal{G}) = X\mathbb{E}(Y|\mathcal{G})$.

(vi) *(Independência) Se* X *é independente de* $\mathcal{G}$*, então* $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$.

(vii) *(Filtração) Se* $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$*, então* $\mathbb{E}(\mathbb{E}(X|\mathcal{G}_1)|\mathcal{G}_2) = \mathbb{E}(\mathbb{E}(X|\mathcal{G}_2)|\mathcal{G}_1) = \mathbb{E}(X|\mathcal{G}_1)$.

Demonstração. Ver (RUFFINO, 2012). □

2.8 Martingales

Esta seção trata de uma classe de processos estocásticos chamados martingales. Os martingales surgem de maneira natural em muitos problemas de probabilidade e estatística. Eles fornecem uma estrutura mais geral do que o caso de variáveis aleatórias independentes na qual resultados, como a lei forte dos grandes números, o teorema central do limite e outros teoremas de convergência podem ser estabelecidos. Grande parte da teoria de martingale foi desenvolvida pelo grande matemático americano J. L. Doob.

Definição 17. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade e seja $N = \{1, 2, \ldots, n_0\}$ um subconjunto não vazio de $\mathbb{N}$, onde $n_0 \in \mathbb{N}$.

1. Uma coleção $\{\mathcal{F}_n : n \in N\}$ de sub σ-álgebras de $\mathcal{F}$ é chamada de *filtração* se $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ para todo $1 \leqslant n < n_0$.
2. Uma coleção de variáveis aleatórias $\{X_n : n \in N\}$ é chamada de *adaptada* à filtração $\{\mathcal{F}_n : n \in N\}$ se X_n é $\mathcal{F}_n$-mensurável para todo $n \in \mathbb{N}$.

Definição 18. Dada uma filtração $\{\mathcal{F}_n : n \in N\}$ e variáveis aleatórias $\{X_n : n \in N\}$, a coleção $\{(X_n, \mathcal{F}_n) : n \in N\}$ é chamada de *martingale* se

(i) X_n é $\mathcal{F}_n$-mensurável para todo $n \in N$;

(ii) $\mathbb{E}(|X_n|) < \infty$ para todo $n \in N$;

(iii) para todo $1 \leqslant n < n_0$,
$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n.$$

Se $\{(X_n, \mathcal{F}_n) : n \in N\}$ é um martingale, então por questão de simplicidade $\{X_n : n \in N\}$ também é chamado de martingale com respeito a filtração $\{\mathcal{F}_n : n \in N\}$. Além disso $\{X_n : n \in N\}$ é chamado de martingale se for um martingale em relação a alguma filtração.

Observação 7. Se $\{X_n : n \in N\}$ é um martingale com respeito a qualquer filtração $\{\mathcal{F}_n : n \in N\}$, então $\{(X_n, \mathcal{X}_n) : n \in N\}$ também é um martingale em relação à filtração natural $\{\mathcal{X}_n : n \in N\}$

onde $\mathcal{X}_n = \sigma(\{X_1, X_2, \ldots, X_n\})$. De fato, claramente cada X_n é $\mathcal{X}_n$-mensurável. Por outro lado, para todo $n \in N$ temos que $\mathcal{X}_n \subset \mathcal{F}_n$, logo

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{X}_n) &= \mathbb{E}(\mathbb{E}(X_{n+1}|\mathcal{F}_n)|\mathcal{X}_n)\\ &= \mathbb{E}(X_n|\mathcal{X}_n)\\ &= X_n.\end{aligned}$$

Assim, $\{(X_n, \mathcal{X}_n) : n \in N\}$ é um martingale.

Exemplo 7. Seja Y uma variável aleatória, integrável e $\{\mathcal{F}_n : n \geqslant 1\}$ uma filtração definidas em um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$. Defina para $n \geqslant 1$ a seguinte variável aleatória

$$X_n = \mathbb{E}(Y|\mathcal{F}_n).$$

Então claramente temos que para todo $n \geqslant 1$, X_n é integrável e $\mathcal{F}_n$-mensurável, logo

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(\mathbb{E}(Y|\mathcal{F}_{n+1})|\mathcal{F}_n)\\ &= \mathbb{E}(Y|\mathcal{F}_n)\\ &= X_n.\end{aligned}$$

Portanto $\{(X_n, \mathcal{F}_n) : n \geqslant 1\}$ é um martingale.

Exemplo 8. Seja $\{Y_n\}_{n \geqslant 1}$ uma coleção arbitrária de variáveis aleatórias integráveis e considere $\mathcal{F}_n = \sigma(\{Y_1, Y_2, \ldots, Y_n\})$. Para todo $n \geqslant 1$ defina

$$X_n = \sum_{i=1}^{n} (Y_i - \mathbb{E}(Y_i|\mathcal{F}_{i-1})).$$

Então, para cada $n \geqslant 1$, X_n é integrável e $\mathcal{F}_n$-mensurável, logo

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \sum_{i=1}^{n+1} \mathbb{E}([Y_i - \mathbb{E}(Y_i|\mathcal{F}_{i-1})]|\mathcal{F}_n)\\ &= \sum_{i=1}^{n} [Y_i - \mathbb{E}(Y_i|\mathcal{F}_{i-1})] + [\mathbb{E}(Y_{n+1}|\mathcal{F}_n) - \mathbb{E}(Y_{n+1}|\mathcal{F}_n)]\\ &= X_n.\end{aligned}$$

Portanto $\{(X_n, \mathcal{F}_n) : n \geqslant 1\}$ é um martingale. Isto mostra que pode-se construir uma sequência de martingale a partir de uma sequência arbitrária de variáveis aleatórias integráveis.

Definição 19. Seja X uma variável aleatória, em um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ e seja $\{\mathcal{F}_n\}_{n \geqslant 1}$ uma filtração em $\mathcal{F}$. Se $\mathbb{E}(|X|) < \infty$ e $X_n = \mathbb{E}(X|\mathcal{F}_n)$. Então $\{(X_n, \mathcal{F}_n) : n \geqslant 1\}$ é chamado de *Martingale de Doob*.

Teorema 7. *Seja* $\{(X_n, \mathcal{F}_n) : n \in N\}$ *um submartingale então* $\{X_n\}$ *converge quase sempre para um limite finito.*

Demonstração. Ver (CHUNG, 2001). $\square$

Observação 8. Todo martingale é também um submartingale e um supermartingale. Por outro lado, qualquer processo estocástico que seja tanto um submartingale quanto um supermartingale é um martingale.

Os seguintes resultados são muito importantes na demostração do Teorema de Kingman.

Teorema 8. *(Teorema Ergódico de Birkhoff) Seja* $f : M \to M$ *uma transformação mensurável e seja* μ *uma medida de probabilidade invariante por* f. *Dada qualquer função integrável* $\varphi : M \to \mathbb{R}$, *o limite*

$$\tilde{\varphi}(x) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x))$$

existe em μ*-quase todo ponto* $x \in M$. *Além disso, a função* $\tilde{\varphi}$ *definida desta forma é integrável e satisfaz*

$$\int \tilde{\varphi}(x) d\mu(x) = \int \varphi(x) d\mu(x).$$

Demonstração. Ver (VIANA; OLIVEIRA, 2014). $\square$

Lema 2. *(Fekete) Se* $(a_n)_n$ *é uma sequência subaditiva, então*

$$\lim_n \frac{a_n}{n} = \inf \frac{a_n}{n} \in [-\infty, +\infty).$$

Demonstração. Se $a_m = -\infty$ para algum $m \in \mathbb{N}$ então pela subaditividade temos que

$$a_n \leqslant a_{n-m} + a_m = -\infty \quad \text{para todo} \quad n > m.$$

Daí

$$\lim_n \frac{a_n}{n} = \inf \frac{a_n}{n} = -\infty.$$

Portanto, a dificultade do lema é dada no caso de sequências reais. Seja $a_n \in \mathbb{R}$ para todo $n \in \mathbb{N}$. Denotemos por $L = \inf \frac{a_n}{n} \in [-\infty, +\infty)$ e seja $\widehat{L}$ qualquer número real maior que L. Escolha $k \geqslant 1$ de tal forma que

$$\frac{a_k}{k} < \widehat{L}.$$

Para $n > k$ temos, pelo algoritmo da divisão, inteiros p e q tais que $n = pk + q$ e $0 \leqslant q \leqslant k - 1$. Aplicando a definição de subaditividade muitas vezes, obtemos

$$a_n = a_{pk+q} \leqslant a_{pk} + a_q \leqslant pa_k + a_q \leqslant pa_k + \alpha$$

onde $\alpha = \max\{a_i : 1 \leqslant i \leqslant k\}$, logo

$$\frac{a_n}{n} \leqslant \frac{pk}{n} \frac{a_k}{k} + \frac{\alpha}{n}. \tag{2.4}$$

Observemos que

$$\lim_n \frac{pk}{n} = \lim_n \left(1 - \frac{q}{n}\right) = 1 \quad \text{e} \quad \lim_n \frac{\alpha}{n} = 0. \tag{2.5}$$

Pela observação acima, temos em (2.4) que quando $n \to \infty$

$$L \leqslant \frac{a_n}{n} \leqslant \widehat{L}$$

para qualquer n suficientemente grande. Finalmente, fazendo $\widehat{L} \to L$, concluímos que

$$\lim_n \frac{a_n}{n} = L = \inf_n \frac{a_n}{n}. \tag{2.6}$$

□

Lema 3. *Seja* $(M, \mathcal{A}, \mu)$ *um espaço de probabilidade e* $T : M \to M$ *uma transformação mensurável que preserva a medida* μ*. Se* $g : M \to \mathbb{R}$ *é uma função mensurável tal que* $g(x) \leqslant g(T(x))$ *para todo* $x \in M$*, então* $g = g \circ T$ μ*-quase todo ponto.*

Demonstração. Dado qualquer $l \in \mathbb{R}$, considere o conjunto

$$\mathcal{U}_l = \{x \in M : g(x) \geqslant l\}.$$

Pela hipótese temos que para $p \in \mathcal{U}_l$, $(g \circ T)(p) \geqslant g(p) \geqslant l$, daí $T(p) \in \mathcal{U}_l$, isto é:

$$\mathcal{U}_l \subset T^{-1}(\mathcal{U}_l).$$

Como T é uma tranformação que preserva a medida μ, então para $\mathcal{U}_l \subset M$ temos que $\mu(\mathcal{U}_l) = \mu(T^{-1}(\mathcal{U}_l))$, logo $\mu(T^{-1}(\mathcal{U}_l) \setminus \mathcal{U}_l) = 0$. Note que

$$T^{-1}(\mathcal{U}_l) \setminus \mathcal{U}_l = \{x \in M : g(x) < l \leqslant (g \circ T)(x)\}. \tag{2.7}$$

Consideremos o conjunto

$$\mathcal{V} = \{x \in M : g(x) < (g \circ T)(x)\}.$$

O objetivo é provar que o conjunto $\mathcal{V}$ tem medida nula. Seja $(q_n)_n$ uma enumeração dos números racionais, então para cada $n \in \mathbb{N}$ defina

$$\mathcal{V}_n = \{x \in M : g(x) < q_n \leqslant (g \circ T)(x)\}.$$

Pela igualdade em (2.7) temos que $\mu(\mathcal{V}_n) = 0$. Observemos que $\mathcal{V} = \cup_{n \geqslant 1} \mathcal{V}_n$, então

$$\mu(\mathcal{V}) = \mu(\bigcup_{n=1}^{\infty} \mathcal{V}_n) \leqslant \sum_{n=1}^{\infty} \mu(\mathcal{V}_n) = 0.$$

Assim, $\mathcal{V}$ tem medida nula, isto é $\mu(\mathcal{V}) = 0$. Portanto $g = g \circ T$ μ-quase todo ponto. □

CAPÍTULO

3

MEDIDAS ESTACIONARIAS

Considere o espaço de probabilidade $(X, \mathscr{X}, p)$ e denote por M o espaço de sequências $X^{\mathbb{N}}$ (ou $X^{\mathbb{Z}}$), dotado com a σ-álgebra produto $\mathscr{A} = \mathscr{X}^{\mathbb{N}}$ (ou $\mathscr{X}^{\mathbb{Z}}$) e a medida produto $\mu = p^{\mathbb{N}}$ (ou $\mu = p^{\mathbb{Z}}$). Mais ainda, seja $f : M \to M$ o deslocamento de Bernoulli em M. Seja $(N, \mathscr{B})$ um espaço mensurável e tome o espaço produto $M \times N$ dotada com a σ-álgebra produto $\mathscr{A} \otimes \mathscr{B}$. Dado um conjunto mensurável $E \subset M \times N$ e $x \in M$ definimos a fibra de x como sendo

$$E^x = \{v \in N : (x, v) \in E\}.$$

De forma análoga para $v \in N$ definimos a fibra de v por

$$E_v = \{x \in M : (x, v) \in E\}.$$

Definição 20. Uma *transformação aleatória* sobre f é uma transformação mensurável da forma

$$\begin{array}{rccc} F: & M \times N & \longrightarrow & M \times N \\ & (x, v) & \mapsto & (f(x), F_x(v)) \end{array}$$

onde F_x depende apenas da coordenada zero de x em M, isto é, $F_x = F_{x_0}$.

Definição 21. Seja $F : M \times N \to M \times N$ uma transformação. *As probabilidades de transição* associadas a F são definidas por

$$p(v, B) = \mu(\{x \in M : F_x(v) \in B\}) = \mu(F^{-1}(M \times B)_v),$$

para cada v em N e cada conjunto mensurável B de N.

Observação 9. O conjunto do lado direito na definição 21 é mensurável e assim $p(v, B)$ está bem definida. Além disso, $p(v, \cdot)$ é uma medida de probabilidade pois ela é enumeravelmente aditiva na segunda variável. Mais ainda: a função $N \to \mathbb{R}$, $v \to p(v, B)$, é mensurável para qualquer $B \subset N$ mensurável.

Definição 22. Dada uma transformação aleatória F, o *operador de transição* associado a F é a transformação linear $\mathscr{P}$ agindo no espaço de funções mensuráveis limitadas $\varphi : N \to \mathbb{R}$ por

$$\mathscr{P}\varphi : N \to \mathbb{R}, \quad \mathscr{P}\varphi(\nu) = \int \varphi(F_x(\nu))d\mu(x).$$

Observação 10. A aplicação $M \to \mathbb{R}$, $x \to \varphi(F_x(\nu))$, é mensurável. Portanto, o operador de transição está bem definido, e é limitado.

Proposição 1. *O operador de transição associado com a transformação aleatória F é mensurável.*

Demonstração. Supondo que φ é a função característica de algum conjunto $B \subset N$, temos que

$$\begin{aligned} \mathscr{P}\varphi(\nu) &= \int \varphi(F_x(\nu))d\mu(x) \\ &= \int \chi_B(F_x(\nu))d\mu(x) \\ &= \mu(\{x : F_x(\nu) \in B\}) \\ &= p(\nu, B), \end{aligned}$$

o qual é uma função mensurável de ν. Agora considere φ como uma função simples, então pela linearidade do operador de transição temos que

$$\mathscr{P}\varphi(\nu) = \int \sum_{i=0}^{n} \chi_{A_i}(F_x(\nu))d\mu(x) = \sum_{i=0}^{n} p(\nu, A_i)$$

que é, uma função mensurável de ν. Em geral, existe uma sequência $(\varphi_n)_n$ de funções simples que convergem uniformemente para φ. Assim temos que

$$\begin{aligned} \lim_{n\to\infty} \mathscr{P}\varphi_n(\nu) &= \lim_{n\to\infty} \int \varphi_n(F_x(\nu))d\mu(x) \\ &= \int \lim_{n\to\infty} \varphi_n(F_x(\nu))d\mu(x) \\ &= \int \varphi(F_x(\nu))d\mu(x) \\ &= \mathscr{P}\varphi(\nu). \end{aligned}$$

Assim, o operador de transição $\mathscr{P}\varphi$ é o limite uniforme das funções mensuráveis $\mathscr{P}\varphi_n$, para todo $n \geqslant 1$, e portanto é mensurável. □

Definição 23. Dada uma transformação aleatória F, o *operador de transição autoadjunto* associado a F age no espaço de medidas de probabilidade η em N como sendo

$$\mathscr{P}^*\eta(B) = \int (F_x)_*\eta(B)d\mu(x) = \int \eta(F_x^{-1}(B))d\mu(x)$$

para qualquer conjunto mensurável $B \subset N$.

Observação 11. Observe que $\eta(F_x^{-1}(B)) = \eta(F^{-1}(M \times B)^x)$ é uma função mensurável de x, portanto o operador de transição autoadjunto está bem definido. Por outro lado, é claro que a partir da definição temos que

$$\begin{aligned}
\mathscr{P}^*\eta(\cup_{i=1}^{\infty} B_i) &= \int \eta(F_x^{-1}(\cup_{i=1}^{\infty} B_i))d\mu(x) \\
&= \int \eta(\cup_{i=1}^{\infty} F_x^{-1}(B_i))d\mu(x) \\
&= \int \sum_{i=1}^{\infty} \eta(F_x^{-1}(B_i))d\mu(x) \\
&= \sum_{i=1}^{\infty} \mathscr{P}^*\eta(B_i).
\end{aligned}$$

Assim $\mathscr{P}^*\eta$ é enumeravelmente aditiva, com $\mathscr{P}^*\eta(N) = 1$, logo $\mathscr{P}^*\eta$ é uma medida de probabilidade em N.

Em geral considere o espaço $\mathscr{T}(N) = \{g : N \to N, \text{transformação mensurável}\}$ dotado de alguma σ-álgebra de modo que a aplicação $\mathscr{F} : M \to \mathscr{T}(N)$, $x \mapsto \mathscr{F}(x) = F_x$ seja mensurável. Assim, considere o push-forward da σ-álgebra $\mathscr{A}$ pela aplicação $\mathscr{F}$. Denotamos por ν o push-forward $\mathscr{F}_*\mu$ da medida de probabilidade μ por $\mathscr{F}$. Portanto as definições do operador de transição e do operador de transição autoadjunto são respectivamente

$$\mathscr{P}\varphi(v) = \int \varphi(g(v))d\nu(g) \quad \text{e} \quad \mathscr{P}^*\eta(B) = \int \eta(g^{-1}(B))d\nu(g).$$

Assim, os operadores de transição são completamente caracterizados pela medida de probabilidade ν.

Lema 4. *Seja $\varphi : N \to \mathbb{R}$ uma função mensurável e limitada. Então*

$$\int \varphi d(\mathscr{P}^*\eta) = \int (\mathscr{P}\varphi)d\eta.$$

Demonstração. Suponha primeiro que φ é a função característica de algum conjunto $B \subset N$, então

$$\int \varphi d(\mathscr{P}^*\eta) = \int \chi_B d(\mathscr{P}^*\eta) = \mathscr{P}^*\eta(B) = \int \eta(F^{-1}(M \times B)^x)d\mu(x) = (\mu \times \eta)(F^{-1}(M \times B)).$$

Por outro lado,

$$\begin{aligned}
\int (\mathscr{P}\varphi)d\eta &= \int (\mathscr{P}\chi_B)d\eta \\
&= \int \mu(\{x : F_x(v) \in B\})d\eta(v) \\
&= \int \mu(F^{-1}(M \times B)_v)d\eta(v) \\
&= (\mu \times \eta)(F^{-1}(M \times B)).
\end{aligned}$$

Pela linearidade, o resultado é válido para funções simples. Cada função mensurável e limitada é limite uniforme de funções simples, então existe uma sequência $(\varphi_n)_n$ de funções simples tal que $\lim_n \varphi_n = \varphi$ uniformemente, logo

$$\begin{aligned}
\int \varphi d(\mathscr{P}^*\eta) &= \int \lim_{n\to\infty} \varphi_n d(\mathscr{P}^*\eta) \\
&= \lim_{n\to\infty} \int \varphi_n d(\mathscr{P}^*\eta) \\
&= \lim_{n\to\infty} \int \mathscr{P}\varphi_n d\eta \\
&= \int \lim_{n\to\infty} (\mathscr{P}\varphi_n) d\eta \\
&= \int (\mathscr{P}\varphi) d\eta.
\end{aligned}$$

Na última igualdade, notar que, como $\mathscr{P}$ é linear, ele é contínuo. □

3.1 Medidas estacionárias

Definição 24. Uma medida de probabilidade η em N é chamada de *medida estacionária* para a transformação F se $\mathscr{P}^*\eta = \eta$, isto é, se

$$\eta(B) = \int \eta(F_x^{-1}(B)) d\mu(x) = \int \eta(g^{-1}(B)) d\nu(g)$$

para todo conjunto mensurável $B \subset N$, onde $\nu = \mathscr{F}_*\mu$.

A seguinte caracterização de medidas estacionárias é específica para transformações aleatórias unilaterais; isto é, tal que $M = X^{\mathbb{N}}$ e sendo $f : M \to M$ o deslocamento unilateral.

Proposição 2. *Seja* $F : M \times N \to M \times N$ *transformação aleatória unilateral. Uma medida de probabilidade* η *em* N *é estacionária para* F *se e somente se a medida de probabilidade* $\mu \times \eta$ *em* $M \times N$ *é* F*-invariante.*

Demonstração. Seja η uma medida estacionária em N. Dada qualquer função mensurável e limitada $\psi : M \times N \to \mathbb{R}$, considere a função

$$\begin{aligned}
\varphi : \quad N &\longrightarrow \mathbb{R} \\
\nu &\mapsto \int \psi(x,\nu) d\mu(x).
\end{aligned}$$

Então

$$\begin{aligned}
\iint \psi(x,v)d\mu(x)d\eta(v) &= \int \varphi(v)d\eta(v) \\
&= \int \varphi(v)d(\mathscr{P}^*\eta)(v) \\
&= \int \mathscr{P}\varphi(v)d\eta(v) \\
&= \iint \varphi(F_x(v))d\mu(x)d\eta(v) \\
&= \iiint \psi(y,F_x(v))d\mu(y)d\mu(x)d\eta(v).
\end{aligned}$$

Considere $x=(x_0,x_1,\dots)$ e $y=(y_0,y_1,\dots)$. Como F_x depende apenas de x_0, a integral tripla pode ser escrita como

$$\begin{aligned}
\iint \psi(x,v)d\mu(x)d\eta(v) &= \iiint \psi(y,F_x(v))d\mu(y)d\mu(x)d\eta(v) \\
&= \iiint \psi(y,F_{x_0}(v))dp^{\mathbb{N}}(y)dp(x_0)d\eta(v).
\end{aligned}$$

Agora, considere $z=(x_0,y_0,y_1\dots)$. Assim temos que $f(z)=y$ e então

$$\begin{aligned}
\iint \psi(x,v)d\mu(x)d\eta(v) &= \iiint \psi(y,F_{x_0}(v))dp^{\mathbb{N}}(y)dp(x_0)d\eta(v) \\
&= \iint \psi(f(z),F_{z_0}(v))dp^{\mathbb{N}}(y)dp(x_0)d\eta(v) \\
&= \iint \psi(f(z),F_z(v))d\mu(z)d\eta(v) \\
&= \iint (\psi\circ F)(z,v)d\mu(z)d\eta(v).
\end{aligned}$$

Isso prova que

$$\int \psi d(\mu\times\eta) = \int \psi\circ F d(\mu\times\eta).$$

Como ψ é arbitrário, então $\mu\times\eta$ é F-invariante. Reciprocamente, seja $\varphi: N\to\mathbb{R}$ função mensurável e limitada, defina $\psi: M\times N\to\mathbb{R}$ como sendo $\psi(x,v)=\varphi(v)$. Então pelo Lema 4 temos

$$\begin{aligned}
\int \varphi(v)(d\mathscr{P}_*\eta)(v) &= \int \mathscr{P}\varphi(v)d\eta(v) \\
&= \iint \psi(x,F_x(v))d\mu(x)d\eta(v) \\
&= \iint \psi(f(x),F_x(v))d\mu(x)d\eta(v) \\
&= \iint \psi(x,v)d\mu(x)d\eta(v) \\
&= \int \varphi(v)d\eta(v).
\end{aligned}$$

Como φ é arbitrário, isso mostra que $\mathscr{P}_*\eta=\eta$, então η é estacionária. $\square$

Lema 5. *Seja* $\varphi : N \to \mathbb{R}$ *contínua. Então* $\mathscr{P}\varphi : N \to \mathbb{R}$ *também é contínua.*

Demonstração. Dado $\nu \in N$, $m \in \mathbb{N}$ e $\delta > 0$, então definamos

$$\mathscr{B}(\nu, m, \delta) = \left\{ x \in M \ ; \ F_x\left(B\left(\nu, \frac{1}{m}\right)\right) \subset B(F_x(\nu), \delta) \right\}$$

o conjunto de todos os elementos em M de modo que F_x seja contínua em $\nu \in N$. Fixe ν em N e $\epsilon > 0$. Pela continuidade, existe $\delta > 0$ tal que

$$d(z, y) < \delta \text{ então } |\varphi(z) - \varphi(y)| < \frac{\epsilon}{2}.$$

O fato que cada F_x seja contínuo garante que quando $m \to \infty$

$$\mu(\mathscr{B}(\nu, m, \delta)^c) \to 0.$$

Então fixe m suficientemente grande de modo que

$$\mu(\mathscr{B}(\nu, m, \delta)^c) < \frac{\epsilon}{4 \sup |\varphi|}.$$

Finalmente para provar a continuidade, seja $w \in N$ tal que $d(\nu, w) < \delta$, então

$$\begin{aligned}
|\mathscr{P}\varphi(\nu) - \mathscr{P}\varphi(w)| &= \left| \int \{\varphi(F_x(\nu)) - \varphi(F_x(w))\} d\mu(x) \right| \\
&\leqslant \int_{\mathscr{B}(\nu,m,\delta)} |\varphi(F_x(\nu)) - \varphi(F_x(w))| d\mu(x) \\
&\quad + \int_{\mathscr{B}(\nu,m,\delta)^c} |\varphi(F_x(\nu)) - \varphi(F_x(w))| d\mu(x) \\
&\leqslant \frac{\epsilon}{2} \int_{\mathscr{B}(\nu,m,\delta)} d\mu(x) + 2 \sup|\varphi| \int_{\mathscr{B}(\nu,m,\delta)^c} d\mu(x) \\
&= \frac{\epsilon}{2} \mu(\mathscr{B}(\nu, m, \delta)) + 2 \sup|\varphi| \mu(\mathscr{B}(\nu, m, \delta)^c) \\
&< \frac{\epsilon}{2} + 2 \sup|\varphi| \frac{\epsilon}{4 \sup|\varphi|} \\
&= \epsilon.
\end{aligned}$$

Portanto $\mathscr{P}\varphi$ é contínua. □

Lema 6. *O operador* $\mathscr{P}^*$ *é contínuo em relação à topologia fraca** *no espaço de medidas de probabilidade em* N.

Demonstração. Seja $(\eta_i)_i$ uma sequência de medidas de probabilidade em N convergindo para η na topologia fraca*. Considere $\varphi : N \to \mathbb{R}$ contínuo. Pelo Lema 5, a função $\mathscr{P}\varphi$ também é contínua. Usando o Lema 4 temos que

$$\lim_{i \to \infty} \int \varphi d(\mathscr{P}^* \eta_i) = \lim_{i \to \infty} \int (\mathscr{P}\varphi) d\eta_i = \int (\mathscr{P}\varphi) d\eta = \int \varphi d(\mathscr{P}^* \eta).$$

Então $\mathscr{P}^*\eta_i$ converge a $\mathscr{P}^*\eta$ na topologia fraca* portanto $\mathscr{P}^*$ é contínuo. □

Proposição 3. *Seja* F *uma transformação aleatória. Suponha que* N *é um espaço métrico compacto e* $F_x : N \to N$ *é contínuo para cada* x *em* M*. Então existe uma medida estacionária para* F.

Demonstração. Seja ζ uma medida de probabilidade arbitrária em N e defina

$$\eta_n = \frac{1}{n}\sum_{i=0}^{n-1} \mathscr{P}^{*i}\zeta.$$

Como o espaço de medidas de probabilidade em N é compacto na topologia fraca*, existe uma subsequência de modo que, na topologia fraca*, (η_{n_k}) converge a alguma probabilidade η. Pela definição do operador autoadjunto temos que

$$\mathscr{P}^*(\eta_{n_k}) = \frac{1}{n_k}\sum_{i=0}^{n_k-1} \mathscr{P}^*(\mathscr{P}^{*i}\zeta) = \frac{1}{n_k}\sum_{i=1}^{n_k} \mathscr{P}^{*i}\zeta = \eta_{n_k} + \frac{1}{n_k}(\mathscr{P}^{*n_k}\zeta - \zeta).$$

Logo

$$\lim_k \mathscr{P}^*(\eta_{n_k}) = \lim_k \eta_{n_k} = \eta.$$

Por outro lado, pelo Lema 6, $\mathscr{P}^*(\eta_{n_k})$ converge a $\mathscr{P}^*\eta$. Assim a medida η é estacionária, isto é $\mathscr{P}^*\eta = \eta$. $\square$

CAPÍTULO

4

TEOREMA DE FURSTENBERG

Neste capítulo, definimos o maior expoente de Lyapunov γ, que fornece a taxa exponencial de crescimento da norma de produtos de matrizes aleatórias independentes e identicamente distribuídas.

4.1 O maior expoente de Lyapunov

Considere $(g_n)_{n\geqslant 1}$ uma sequência de matrizes aleatórias independentes e identicamente distribuídas com distribuição comum μ. Seja $M_n = g_n \dots g_1$, então

$$\|g_n \dots g_1\| = \|M_n\| \leqslant \|g_n\| \dots \|g_1\|.$$

Por outro lado

$$\begin{aligned}\log^+ \|M_n\| &\leqslant (\log\|g_n\| + \ldots + \log\|g_1\|)^+ \\ &\leqslant \log^+\|g_n\| + \ldots + \log^+\|g_1\|.\end{aligned}$$

Se $\mathbb{E}(\log^+\|g_i\|) < \infty$ (onde $f^+ = \sup\{f, 0\}$) então $\log^+\|M_n\|$ é integrável. Mais ainda, a sequência $(\mathbb{E}(\log\|M_n\|))_{n\geqslant 1}$ é subaditiva, de fato, sejam $n, p \geqslant 1$, então

$$\begin{aligned}\mathbb{E}(\log\|M_{n+p}\|) &\leqslant \mathbb{E}(\log\|g_{n+p}\| \dots \|g_{n+1}\|) + \mathbb{E}(\log\|g_n\| \dots \|g_1\|) \\ &= \mathbb{E}(\log\|M_p\|) + \mathbb{E}(\log\|M_n\|).\end{aligned}$$

Portanto em $[-\infty, +\infty)$ temos que

$$\lim_n \mathbb{E}(\log\|M_n\|) = \inf_m \frac{1}{m}\mathbb{E}(\log\|M_m\|).$$

Definição 25. Se $\mathbb{E}(\log^+\|g_1\|) < \infty$, o *maior expoente de Lyapunov* associado à medida μ é um elemento em $[-\infty, +\infty)$ definido como sendo o limite

$$\gamma = \lim_n \frac{1}{n}\mathbb{E}(\log\|g_n \dots g_1\|).$$

Observação 12. É fácil ver que a definição do maior expoente de Lyapunov é independente da norma escolhida uma vez que todas as normas em $M(d,\mathbb{R})$ são equivalentes.

Vamos provar que se as matrizes são invertíveis, então quase sempre

$$\gamma = \lim_n \frac{1}{n} \log \|g_n \dots g_1\|.$$

Embora isso se torne uma conseqüência fácil do teorema ergódico subaditivo de Kingman, devemos seguir a prova original do resultado. Há dois motivos para isso: tal prova, é uma boa ilustração das técnicas que usaremos, por outro lado, nos fornece algumas informações que serão necessárias posteriormente.

A seguinte proposição é importante na prova do resultado principal deste capítulo. Considere $(\Omega, \mathcal{A}, \mathbb{P})$ um espaço de probabilidade.

Proposição 4. *Sejam G um semigrupo topológico agindo no espaço topológico M e $\sigma : G \times M \to \mathbb{R}$ um cociclo aditivo, considere a sequência de variáveis aleatórias $(\omega_n)_n$ em G com distribuição comum μ. Se ν é uma medida em M μ-estacionária tal que*

$$\iint |\sigma(g,x)| d\mu(g) d\nu(x) < \infty$$

então o limite

$$\lim_{n\to\infty} \frac{1}{n} \sigma(\omega_n \cdots \omega_1, x) = \phi(\omega, x)$$

existe $\mathbb{P} \times \nu$-quase sempre e $\phi(\omega,x) \in L^1(\Omega \times M, \mathbb{P} \times \nu)$.

Demonstração. Considere $\Omega = G^{\mathbb{N}} = \{\omega = (\omega_1, \omega_2, \dots); \omega_i \in G\}$ e $\mathbb{P} = \mu^{\mathbb{N}}$ a medida de probabilidade em Ω e consideramos a σ-álgebra $\mathcal{A}$ como sendo a gerada pelas aplicações coordenadas ω_n.

Agora construímos o sistema dinâmico em $\Omega \times M$. Equipamos $\Omega \times M$ com a medida produto $\mathbb{P} \times \nu$, e definimos F como sendo

$$\begin{array}{rccc} F: & \Omega \times M & \longrightarrow & \Omega \times M \\ & ((\omega_n)_n, x) & \mapsto & ((\omega_{n+1})_n, \omega_1 x) \end{array}$$

como ν é uma medida em M μ-estacionária então pela Proposição 2 a medida $\mathbb{P} \times \nu$ é F-invariante, logo F é uma transformação mensurável que preserva a medida $\mathbb{P} \times \nu$. Defina a função

$\varphi : \Omega \times M \to \mathbb{R}$ tal que $\varphi((\omega_n)_n, x) = \sigma(\omega_1, x)$, então φ é $\mathbb{P} \times \nu$-integrável. Note que

$$\begin{aligned}
\sigma(\omega_n \cdots \omega_1, x) &= \sigma(\omega_n, \omega_{n-1} \cdots \omega_1.x) + \sigma(\omega_{n-1} \cdots \omega_1, x) \\
&= \sigma(\omega_n, \omega_{n-1} \cdots \omega_1.x) + \sigma(\omega_{n-1}, \omega_{n-2} \cdots \omega_1.x) + \sigma(\omega_{n-2} \cdots \omega_1, x) \\
&= \sigma(\omega_n, \omega_{n-1} \cdots \omega_1.x) + \sigma(\omega_{n-1}, \omega_{n-2} \cdots \omega_1.x) + \cdots + \sigma(\omega_1, x) \\
&= \sum_{j=0}^{n-1} \sigma(\omega_{j+1}, \omega_j \cdots \omega_1.x) \\
&= \sum_{j=0}^{n-1} (\varphi \circ F^j)(\omega, x).
\end{aligned}$$

Assim pelo Teorema de Birkhoff temos

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ F^j)(\omega, x) = \phi(\omega, x)$$

$\mathbb{P} \times \nu$-quase sempre. □

Teorema 9. *[(BOUGEROL et al., 2012)] Sejam $g_1, g_2, \ldots$ matrizes aleatórias independentes e identicamente distribuidas em $GL(d, \mathbb{R})$. Se $\mathbb{E}(\log^+ \|g_1\|) < \infty$, então com probabilidade um, temos que*

$$\lim_n \frac{1}{n} \log \|g_n \ldots g_1\| = \gamma.$$

Demonstração. Fixe um número inteiro m, para todo inteiro n considere $n = pm + q$ onde $0 \leqslant q < m$ e $p \geqslant 0$.

$$\begin{aligned}
\frac{1}{n} \log \|g_n \ldots g_{pm} \ldots g_1\| &\leqslant \frac{1}{n} \log(\|g_n \ldots g_{pm+1}\| \|g_{pm} \ldots g_1\|) \\
&= \frac{1}{n} \log \|g_n \ldots g_{pm+1}\| + \frac{1}{n} \|g_{pm} \ldots g_1\| \\
&\leqslant \frac{1}{n} \sum_{i=pm+1}^{n} \log \|g_i\| + \frac{1}{n} \sum_{j=0}^{p-1} \log \|g_{(j+1)m} \cdots g_{jm+1}\|. \qquad (4.1)
\end{aligned}$$

Note que o primeiro somatório à direita em (4.1) é limitado já que $\mathbb{E}(\log^+ \|g_1\|) < \infty$, o qual implica que $\log^+ \|g_1\| < \infty$ quase sempre e como $\log \|g_i\| \leqslant \log^+ \|g_i\|$, o somatório é limitado, assim, o limite vai para zero. Vamos considerar as seguintes variáveis aleatórias $(\log \|g_{(j+1)m} \cdots g_{jm+1}\|)_j$. Note que para $j = 0$ temos que

$$\begin{aligned}
\mathbb{E}(\log \|g_m \ldots g_1\|) &\leqslant \sum_{i=1}^{m} \mathbb{E}(\log \|g_i\|) \\
&\leqslant \sum_{i=1}^{m} \mathbb{E}(\log^+ \|g_i\|) \\
&< \infty.
\end{aligned}$$

Logo, como $\mathbb{E}(\log\|g_m \ldots g_1\|)$ é finito, pela lei dos grandes números temos que

$$\lim_p \frac{1}{p}\sum_{j=0}^{p-1} \log\|g_{(j+1)m}\cdots g_{jm+1}\| = \mathbb{E}(\log\|g_m \ldots g_1\|). \tag{4.2}$$

Agora, se analisamos o segundo somatório em (4.1), observamos em $n = pm+q$ que quando n tende ao infinito, p também tende ao infinito, logo

$$\begin{aligned}
\lim_p \frac{1}{n}\sum_{j=0}^{p-1} \log\|g_{(j+1)m}\cdots g_{jm+1}\| &= \left(\lim_p \frac{p}{n}\right)\left(\lim_p \frac{1}{p}\sum_{j=0}^{p-1} \log\|g_{(j+1)m}\cdots g_{jm+1}\|\right)\\
&= \frac{1}{m}\lim_p\left(\frac{1}{p}\sum_{j=0}^{p-1} \log\|g_{(j+1)m}\cdots g_{jm+1}\|\right)\\
&= \frac{1}{m}\mathbb{E}(\log\|g_m \ldots g_1\|).
\end{aligned}$$

Tomando o limite quando m tende ao infinito em (4.1), pela definição de γ, segue que

$$\lim_n \frac{1}{n}\log\|g_n \ldots g_{pm}\ldots g_1\| \leqslant \lim_m \frac{1}{m}\mathbb{E}(\log\|g_m \ldots g_1\|) = \gamma. \tag{4.3}$$

Portanto, no caso em que $\gamma = -\infty$ temos

$$\lim_n \frac{1}{n}\log\|g_n \ldots g_2 g_1\| = \gamma. \tag{4.4}$$

Suponha agora que γ seja finito. Considere $\mathcal{B}$ como sendo o espaço compacto (e separável) de todas as matrizes de ordem $d \times d$ cuja norma é igual a um.

$$\mathcal{B} = \{u \in M(d,\mathbb{R}); \|u\| = 1\}.$$

Seja g em $GL(d,\mathbb{R})$ e u em $\mathcal{B}$. Defina agora a ação do grupo topológico $GL(d,\mathbb{R})$ sobre o espaço topológico $\mathcal{B}$ como sendo

$$\begin{aligned}
\theta\colon GL(d,\mathbb{R}) \times \mathcal{B} &\longrightarrow \mathcal{B}\\
(g,u) &\longmapsto g\cdot u = \frac{gu}{\|gu\|}.
\end{aligned}$$

Assim $\mathcal{B}$ se torna um $GL(d,\mathbb{R})$-espaço. Seja δ a medida de Dirac na matriz identidade Id em $\mathcal{B}$, isto é, para $A \subset \mathcal{B}$, $\delta(A) = 1$ se $Id \in A$ e $\delta(A) = 0$ se $Id \notin A$. Considere μ como sendo a distribuição de g_1. Para todo inteiro n defina a distribuição ν_n em $\mathcal{B}$ como sendo

$$\nu_n = \frac{1}{n}\sum_{i=0}^{n-1} \mu^i * \delta. \tag{4.5}$$

Se (ν_{n_i}) é uma subsequência de (ν_n), então seu limite ν é uma medida μ-invariante pela Proposição 3.

Defina o cociclo aditivo $\sigma : GL(d,\mathbb{R}) \times \mathcal{B} \longrightarrow \mathbb{R}$, como sendo $\sigma(g,u) = \log\|gu\|$. Observe que $\sigma(g,u) \leqslant \log\|g\| + \log\|u\| = \log\|g\|$, então $\sigma^+(g,u) \leqslant \log^+\|g\|$. Como $\mathbb{E}(\log^+\|g_1\|) < \infty$, então

$$\begin{aligned}
\left|\int \sigma^+(g,u)\ d\mu(g)\right| &\leqslant \int |\sigma^+(g,u)|\ d\mu(g)\\
&\leqslant \int \log^+\|g\|\ d\mu(g)\\
&= \mathbb{E}(\log^+\|g\|) < \infty.
\end{aligned}$$

Portanto $\int \sigma^+(g,u)\ d\mu(g)$ é uma função contínua e limitada em $\mathcal{B}$, assim temos que

$$\lim_i \iint \sigma^+(g,u)\ d\mu(g)\ d\nu_{n_i}(u) = \iint \sigma^+(g,u)\ d\mu(g)\ d\nu(u).$$

Analogamente para $\sigma^- = \sup\{-\sigma, 0\}$ e para qualquer $k > 0$ temos que $\int \inf\{k, \sigma^-(g,u)\}\ d\mu(g)$ é uma função contínua e limitada em $\mathcal{B}$, daí

$$\lim_i \iint \inf\{k, \sigma^-(g,u)\}\ d\mu(g) d\nu_{n_i}(u) = \iint \inf\{k, \sigma^-(g,u)\}\ d\mu(g) d\nu(u). \tag{4.6}$$

Além disso, temos também que para qualquer $k > 0$

$$\liminf_i \iint \sigma^-(g,u) d\mu(g)\ d\nu_{n_i}(u) \geqslant \lim_i \iint \inf\{k, \sigma^-(g,u)\}\ d\mu(g) d\nu_{n_i}(u). \tag{4.7}$$

Então de (4.6) e (4.7) obtemos

$$\liminf_i \iint \sigma^-(g,u) d\mu(g)\ d\nu_{n_i}(u) \geqslant \iint \inf\{k, \sigma^-(g,u)\}\ d\mu(g) d\nu(u). \tag{4.8}$$

Seja a sequência de funções (f_k) onde $f_k : \mathcal{B} \to \mathbb{R}$ é dada por

$$f_k(u) = \int \inf\{k, \sigma^-(g,u)\}\ d\mu(g).$$

Logo (f_k) é uma sequência de funções mensuráveis não negativas, pela propriedade do truncamento horizontal

$$\lim_k \iint \inf\{k, \sigma^-(g,u)\}\ d\mu(g) d\nu(u) = \iint \sigma^-(g,u)\ d\mu(g) d\nu(u).$$

Então, para nossa sequência de funções, temos

$$\lim_k \int f_k(u) d\nu(u) = \iint \sigma^-(g,u) d\ \mu(g) d\ \nu(u). \tag{4.9}$$

Logo pelo lema de Fatou e por (4.9) segue que

$$\begin{aligned}
\liminf_k \iint \inf\{k, \sigma^-(g,u)\}\ d\mu(g) d\nu(u) &\geqslant \int \liminf_k f_k(u)\ d\nu(u) && (4.10)\\
&= \iint \sigma^-(g,u)\ d\mu(g) d\nu(u). && (4.11)
\end{aligned}$$

Agora, ao tomarmos o limite em (4.8) quando $k \to \infty$ obtemos, por (4.10),

$$\begin{aligned}\liminf_i \iint \sigma^-(g,u)\ d\mu(g)\ d\nu_{n_i}(u) &\geqslant \liminf_k \iint \inf\{k, \sigma^-(g,u)\}\ d\mu(g)d\nu(u)\\ &\geqslant \iint \sigma^-(g,u)\ d\mu(g)d\nu(u).\end{aligned}$$

Portanto

$$\liminf_i \iint \sigma^-(g,u)\ d\mu(g)d\nu_{n_i}(u) \geqslant \iint \sigma^-(g,u)\ d\mu(g)d\nu(u). \tag{4.12}$$

Por outro lado, pela propriedade do $\limsup$ e por (4.12) segue que

$$\begin{aligned}\limsup_i \iint \sigma(g,u)\ d\mu(g)d\nu_{n_i}(u) &= \limsup_i \iint (\sigma^+(g,u) - \sigma^-(g,u))\ d\mu(g)d\nu_{n_i}(u)\\ &\leqslant \limsup_i \iint \sigma^+(g,u)\ d\mu(g)d\nu_{n_i}(u)\\ &\quad - \liminf_i \iint \sigma^-(g,u)\ d\mu(g)d\nu_{n_i}(u)\\ &\leqslant \iint \sigma^+(g,u)d\mu(g)d\nu(u) - \iint \sigma^-(g,u)\ d\mu(g)d\nu(u)\\ &= \iint \sigma(g,u)d\mu(g)d\nu(u).\end{aligned}$$

Portanto

$$\limsup_i \iint \sigma(g,u)\ d\mu(g)d\nu_{n_i}(u) \leqslant \iint \sigma(g,u)\ d\mu(g)d\nu(u). \tag{4.13}$$

Por outro lado, de (4.5) segue que

$$\begin{aligned}\iint \sigma(g,u)\ d\mu(g)d\nu_n(u) &= \iint \sigma(g,u)\ d\mu(g)d\left(\frac{1}{n}\sum_{i=0}^{n-1}\mu^i * \delta\right)(u)\\ &= \frac{1}{n}\sum_{i=0}^{n-1}\iint \sigma(g,u)\ d\mu(g)d(\mu^i * \delta)(u). \qquad (4.14)\end{aligned}$$

Fazendo uso da definição da convolução de medidas temos que para g_i em $GL(d,\mathbb{R})$

$$\begin{aligned}\iint \sigma(g,u)\ d\mu(g)d(\mu^i * \delta)(u) &= \iiint \sigma(g, g_1\cdot u)\ d\mu(g)d\mu(g_1)d(\mu^{i-1} * \delta)(u)\\ &= \vdots\\ &= \underbrace{\int\cdots\int}_{i+2} \sigma(g, g_i\cdot(\ldots\cdot(g_1\cdot u)))\ d\mu(g)d\mu(g_i)\ldots d\mu(g_1)d\delta(u)\\ &= \underbrace{\int\cdots\int}_{i+2} \sigma(g, (g_i\ldots g_1)\cdot u)\ d\mu(g)d\mu(g_i)\ldots d\mu(g_1)d\delta(u).\end{aligned}$$

Substituindo está última igualdade em (4.14), obtemos:

$$\begin{aligned}
\iint \sigma(g,u)\ d\mu(g)d\nu_n(u) &= \frac{1}{n}\sum_{i=0}^{n-1}\underbrace{\int}_{i+2} \sigma(g_{i+1}, g_i \dots g_1 \cdot u)\ d\mu(g_{i+1})d\mu(g_i)\dots d\mu(g_1)d\delta(u) \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\underbrace{\int\cdots\int}_{i+2} \sigma(g_{i+1}, (g_i \dots g_1)\cdot \mathrm{Id})\ d\mu(g_{i+1})d\mu(g_i)\dots d\mu(g_1) \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\int \sigma(g_{i+1}, (g_i g_{i-1} \dots g_1)\cdot \mathrm{Id})\ d\mathbb{P} \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}(\sigma(g_{i+1}, (g_i g_{i-1} \dots g_1)\cdot \mathrm{Id})) \qquad (4.15)\\
&= \frac{1}{n}\mathbb{E}(\sigma(g_n g_{n-1}\dots g_1, \mathrm{Id})) \\
&= \frac{1}{n}\mathbb{E}(\log \|g_n g_{n-1}\dots g_1\|).
\end{aligned}$$

Observação 13. Observe que pela definição de cociclo aditivo, para $g_i, g_{i-1}, \dots, g_1$ em $GL(d,\mathbb{R})$ e u em $\mathcal{B}$ obtemos o seguinte:

$$\begin{aligned}
\sigma(g_n g_{n-1}\dots g_1, u) &= \sigma(g_n, g_{n-1}\dots g_1\cdot u) + \sigma(g_{n-1}\dots g_1, u) \\
&= \sigma(g_n, g_{n-1}\dots g_1\cdot u) + \sigma(g_{n-1}, g_{n-2}\dots g_1\cdot u) + \sigma(g_{n-2}\dots g_1, u) \\
&= \sigma(g_n, g_{n-1}\dots g_1\cdot u) + \sigma(g_{n-1}, g_{n-2}\dots g_1\cdot u) + \dots \sigma(g_1, x) \\
&= \sum_{i=0}^{n-1}\sigma(g_{i+1}, g_i\dots g_1\cdot u).
\end{aligned}$$

Está observação foi considerada em (4.15).

Portanto, aplicando o limite obtemos

$$\lim_n \iint \sigma(g,u)\ d\mu(g)d\nu_n(u) = \lim_n \frac{1}{n}\mathbb{E}(\log\|g_n g_{n-1}\dots g_1\|) = \gamma. \qquad (4.16)$$

De (4.13) e (4.16) concluímos que

$$\gamma \leqslant \iint \sigma(g,u)\ d\mu(g)d\nu(u). \qquad (4.17)$$

Está última desigualdade juntamente com alguns arranjos na identidade $\sigma = \sigma^+ - \sigma^-$ garante a integrabilidade do cociclo σ. De fato, como $\sigma = \sigma^+ - \sigma^-$ e $|\sigma| = \sigma^+ + \sigma^-$ então

$|\sigma| = 2\sigma^+ - \sigma$, logo temos que

$$\begin{aligned}
\iint |\sigma(g,u)|\ d\mu(g)d\nu(u) &= \iint (2\sigma^+ - \sigma)\ d\mu(g)d\nu(u) \\
&= 2\iint \sigma^+(g,u)\ d\mu(g)d\nu(u) - \iint \sigma(g,u)\ d\mu(g)d\nu(u) \\
&\leqslant 2\mathbb{E}(\log^+\|g\|) - \gamma \\
&< \infty.
\end{aligned}$$

Assim, pela Proposição 4 existe $f : \Omega \times \mathcal{B} \to \mathbb{R}$ de modo que com probabilidade um

$$\begin{aligned}
\lim_n \frac{1}{n}\sigma(g_n(\omega)\dots g_1(\omega),u) &= \lim_n \frac{1}{n}\log\|(g_n(\omega)\dots g_1(\omega)\cdot u\| \\
&= f(\omega,u).
\end{aligned}$$

Lembre-se que ν é uma medida μ-invariante, então

$$\begin{aligned}
\iint f(\omega,u)\ d\mathbb{P}(\omega)d\nu(u) &= \lim_n \iint \frac{1}{n}\sigma(g_n(\omega)\dots g_1(\omega),u)\ d\mathbb{P}(\omega)d\nu(u) \\
&= \lim_n \iint \frac{1}{n}\sum_{i=1}^{n}\sigma(g_i(\omega),(g_{i-1}(\omega)\dots g_1(\omega))\cdot u)\ d\mathbb{P}(\omega)d\nu(u) \\
&= \lim_n \iint \frac{1}{n}\sum_{i=1}^{n}\sigma(g_i(\omega),u)\ d\mathbb{P}(\omega)d\nu(u) \\
&= \lim_n \iint \frac{1}{n}n\sigma(g,u)\ d\mu(g)d\nu(u) \\
&= \iint \sigma(g,u)\ d\mu(g)d\nu(u).
\end{aligned}$$

Da desigualdade em (4.17) temos o seguinte

$$\gamma \leqslant \iint f(\omega,u)\ d\mathbb{P}(\omega)d\nu(u) \tag{4.18}$$

Além disso,

$$\frac{1}{n}\log\|g_n(\omega)\dots g_1(\omega)\cdot u\| \leqslant \frac{1}{n}\log\|g_n(\omega)\dots g_1(\omega)\| + \frac{1}{n}\log\|u\|.$$

Tomando o limite quando $n \to \infty$

$$\begin{aligned}
f(\omega,u) &= \liminf_n \frac{1}{n}\log\|(g_n(\omega)\dots g_1(\omega)\| && (4.19)\\
&\leqslant \limsup_n \frac{1}{n}\log\|(g_n(\omega)\dots g_1(\omega)\| && (4.20)\\
&= \frac{1}{m}\mathbb{E}(\log\|(g_m(\omega)\dots g_1(\omega)\| && (4.21)\\
&= \gamma. && (4.22)
\end{aligned}$$

Portanto, segue que

$$f(\omega,u) \leqslant \gamma. \tag{4.23}$$

De (4.18) e (4.23) temos que

$$f(\omega,u) = \gamma \quad \text{quase sempre.} \tag{4.24}$$

Finalmente de (4.19), (4.20) e (4.24) temos

$$\lim_{n} \frac{1}{n} \log \|g_n \dots g_1\| = \gamma \quad \text{quase sempre.}$$

$\square$

CAPÍTULO

5

TEOREMA ERGÓDICO SUBADITIVO DE KINGMAN

Neste capítulo, demonstraremos o Teorema ergódico subaditivo de Kingman. Começamos por enunciar o Teorema e logo dar uma breve explicação do esquema da prova para facilitar a leitura. Em seguida apresentamos e demostramos alguns lemas fundamentais usados durante a demonstração do teorema principal do capítulo. Finalmente, fazemos a demonstração do Teorema dividindo-o em duas partes, a primeira como desigualdade inferior e a segunda como superior.

A parte positiva e a parte negativa de uma função mensurável $\varphi : M \to [-\infty, +\infty)$ são as funções não negativas definidas por

$$\varphi^{+}(x) = \max\{\varphi(x), 0\} \quad \text{e} \quad \varphi^{-}(x) = \max\{-\varphi(x), 0\}.$$

Considere um espaço de probabilidade $(M, \mathcal{A}, \mu)$ e $f : M \to M$. A função φ é chamada de *invariante* se $(\varphi \circ f)(x) = \varphi(x)$ para μ-quase todo ponto $x \in M$. Uma sequência $\varphi_n : M \to [-\infty, +\infty)$, $n \geqslant 1$, de funções mensuráveis é *subaditiva*, relativa à aplicação f, se

$$\varphi_{m+n} \leqslant \varphi_m + \varphi_n \circ f^m$$

para todo $n, m \geqslant 1$.

Exemplo 9. Dada uma função mensurável $A : M \to GL(d, \mathbb{R})$, considere a sequência de funções

$$\varphi_n(x) = \log \|A^n(x)\|$$

onde $A^n(x) = A(f^{n-1}(x)) \ldots A(f(x))A(x)$. Então $(\varphi_n)_n$ é subaditiva: de fato, dados $n, m \geqslant 1$

temos

$$\begin{aligned}
\varphi_{m+n}(x) &= \log\|A^{m+n}(x)\| \\
&= \log\|A^m(f^n(x))A^n(x)\| \\
&\leqslant \log(\|A^m(f^n(x))\|\|A^n(x)\|) \\
&\leqslant \log\|A^m(f^n(x))\| + \log\|A^n(x)\| \\
&= (\varphi_m \circ f^n)(x) + \varphi_n(x).
\end{aligned}$$

Teorema 10. *[Kingman (VIANA, 2014)] Seja $(M, \mathcal{A}, \mu)$ um espaço de probabilidade e $f : M \to M$ uma transformação mensurável que preserva a medida μ. Seja $(\varphi_n)_n$ uma sucessão subaditiva de funções mensuráveis tais que $\varphi_1^+ \in L^1(\mu)$. Então $(\frac{\varphi_n}{n})_n$ converge μ-quase todo ponto para uma função invariante $\varphi : M \to [-\infty, +\infty)$. Além disso, a parte positiva φ^+ é integrável e*

$$\int \varphi \; d\mu = \lim_n \frac{1}{n}\int \varphi_n \; d\mu = \inf_n \frac{1}{n}\int \varphi_n \; d\mu \in [-\infty, +\infty). \tag{5.1}$$

A ideia da prova é a seguinte. Considere $(\varphi_n)_n$ uma sequência subaditiva de funções mensuráveis de modo que $\varphi_1^+ \in L^1(\mu)$, então como $(\varphi_n)_n$ é subaditiva, temos que

$$\varphi_{m+n} \leqslant \varphi_m + \varphi_n \circ f^m.$$

Além disso

$$\varphi_n \leqslant \varphi_1 + \varphi_1 \circ f + \ldots + \varphi_1 \circ f^{n-1}.$$

Esta desigualdade permanece verdadeira se substituímos φ_n e φ_1 por φ_n^+ e φ_1^+. De uma forma resumida, temos

$$\varphi_n \leqslant \sum_{j=0}^{n-1} \varphi_1 \circ f^j \quad \text{e} \quad \varphi_n^+ \leqslant \sum_{j=0}^{n-1} \varphi_1^+ \circ f^j. \tag{5.2}$$

A integrabilidade de φ_1^+ implica que φ_n^+ é também integrável para todo $n \in \mathbb{N}$. De fato

$$\int |\varphi_n^+| \; d\mu \leqslant \sum_{j=0}^{n-1} \int |\varphi_1^+ \circ f^j| \; d\mu = \sum_{j=0}^{n-1} \int |\varphi_1^+| \; d\mu = n \int |\varphi_1^+| \; d\mu < \infty.$$

Por outro lado, a subaditividade de $(\varphi_n)_n$ implica também que

$$h_n = \int \varphi_n \; d\mu \quad \text{para todo } \; n \geqslant 1 \tag{5.3}$$

é uma sequência subaditiva em $[-\infty, +\infty)$. Logo pelo Lema de Fekete o seguinte limite existe, onde $L \in [-\infty, +\infty)$:

$$\lim_n \frac{h_n}{n} = \inf \frac{h_n}{n} = L.$$

Defina $\varphi_+ : M \to [-\infty, +\infty]$ e $\varphi_- : M \to [-\infty, +\infty]$ como sendo

$$\varphi_+(x) = \limsup_n \frac{\varphi_n(x)}{n} \quad \text{e} \quad \varphi_-(x) = \liminf_n \frac{\varphi_n(x)}{n}. \tag{5.4}$$

É claro que a seguinte desigualdade é valida para todo $x \in M$:

$$\varphi_-(x) \leqslant \varphi_+(x). \tag{5.5}$$

Nosso objetivo é mostrar que

$$\int \varphi_+(x)\ d\mu(x) \leqslant L \leqslant \int \varphi_-(x)\ d\mu(x), \tag{5.6}$$

isto implica que

$$\int \varphi_+(x)\ d\mu(x) = L = \int \varphi_-(x)\ d\mu(x),$$

e portanto $\varphi_+ = \varphi_-$ μ-quase sempre. Desde que toda função φ_n seja limitada a partir de $-\infty$. Consequentemente, as duas funções φ_+ e φ_- coincidem em μ-quase todos os pontos e suas integrais são iguais a L. Isso prova o teorema neste caso. No final, usamos um método de truncamento para remover a condição de uniformemente limitado inferiormente, para assim, provar o teorema em completa generalidade.

5.1 Lemas Fundamentais

Nesta seção, vamos considerar o que chamaremos de hipótese de limitação. Vamos assumir inicialmente que a sequência $(\frac{\varphi_n}{n})_n$ está uniformemente limitada inferiormente, isto é, existe $c \in \mathbb{R}$, $c > 0$ de modo que para todo $n \in \mathbb{N}$

$$\frac{\varphi_n}{n} \geqslant c.$$

Além disso vamos ter as seguintes considerações iniciais. Dado $\epsilon > 0$ fixo, considere a sucessão de conjuntos $(E_k)_k$ definida por

$$E_k = \{x \in M;\ \varphi_j(x) \leqslant j(\varphi_-(x) + \epsilon)\ \text{ para algum }\ j \in \{1, \dots, k\}\}. \tag{5.7}$$

Observação 14. Lembre do seguinte resultado: $\liminf_n a_n = -\infty$ se e somente se (a_n) não tem limite inferior.

Note que pela definição de φ_- e o fato de ser limitada inferiormente (não toma o valor de $-\infty$), então cada $x \in M$ pertence a E_k para todo k suficientemente grande, assim

$$M = \bigcup_{k=1}^{\infty} E_k.$$

É claro que para todo $k \in \mathbb{N}$, $E_k \subset E_{k+1}$, então

$$\lim_k \mu(E_k) = \mu(\bigcup_{k=1}^{\infty} E_k) = \mu(M) = 1. \tag{5.8}$$

Agora, defina a sucessão de funções $\psi_k : M \to \mathbb{R} \cup \{\infty\}$ como sendo

$$\psi_k(x) = \begin{cases} \varphi_-(x) + \epsilon, & \text{se } x \in E_k \\ \varphi_1(x), & \text{se } x \notin E_k. \end{cases} \tag{5.9}$$

Afirmamos que $\psi_k(x) \geqslant \varphi_-(x) + \epsilon$ para cada $x \in M$. De fato, considere o caso em que $x \notin E_k$, então $\varphi_j(x) > j(\varphi_-(x) + \epsilon)$ para cada $j = 1, \ldots, k$. Em particular, para $j = 1$ e pela definição de ψ_k temos que $\psi_k(x) = \varphi_1(x) > \varphi_-(x) + \epsilon$. E além disso, se $x \in E_k$ então $\psi_k(x) = \varphi_-(x) + \epsilon$.

Então para todo $k \in \mathbb{N}$ temos que $\psi_k(x) \geqslant \varphi_-(x) + \epsilon$. Assim, pelo Teorema da convergência monótona

$$\lim_k \int \psi_k \, d\mu = \int (\varphi_-(x) + \epsilon) \, d\mu. \tag{5.10}$$

O seguinte lema diz que a sequência (φ_n) é limitada por uma soma orbital de ψ_k e um termo que chamaremos de "termo de erro". Como as somas orbitais são sequências aditivas, isso praticamente reduz nosso ajuste subaditivo ao caso aditivo. Dado que a última soma no lado direito da desigualdade (o termo de erro) é a soma de um número fixo de funções integráveis, seu valor será desconsiderado.

Lema 7. *Para* μ*-quase todo ponto* $x \in M$ *e para todo* $n > k \geqslant 1$,

$$\varphi_n(x) \leqslant \sum_{i=0}^{n-k-1} \psi_k(f^i(x)) + \sum_{i=n-k}^{n-1} \max\{\psi_k, \varphi_1\}(f^i(x)).$$

Demonstração. Seja $x \in M$ tal que $\varphi_-(x) = \varphi_-(f^j(x))$ (φ_- é invariante) para qualquer $j \geqslant 1$. Introduzimos duas sequências de inteiros $(m_j)_j$ e $(n_j)_j$ tais que

$$m_0 \leqslant n_1 < m_1 \leqslant n_2 < m_2 \leqslant \ldots \tag{5.11}$$

definidos indutivamente da seguinte forma: considere $m_0 = 0$ e seja n_1 como sendo o menor inteiro maior ou igual a m_0 tal que $f^{n_1}(x)$ pertence a E_k, então pela definição de E_k existe $j_1 \in \{1, \ldots, k\}$ tal que $\varphi_{j_1}(f^{n_1}(x)) \leqslant j_1(\varphi_-(f^{n_1}(x)) + \epsilon) = j_1(\varphi_-(x) + \epsilon)$ (a última desigualdade é verdade já que φ_- é invariante). Agora considere $m_1 = n_1 + j_1$ e seja n_2 como sendo o menor inteiro maior ou igual a m_1 de modo que $f^{n_2}(x)$ pertence a E_k, então existe $j_2 \in \{1, \ldots, k\}$ tal que $\varphi_{j_2}(f^{n_2}(x)) \leqslant j_2(\varphi_-(f^{n_2}(x)) + \epsilon) = j_2(\varphi_-(x) + \epsilon)$. Nós podemos aplicar o mesmo argumento acima para obter em geral que dado $j \geqslant 1$, seja n_j como sendo o menor inteiro maior ou igual a m_{j-1} de modo que $f^{n_j}(x)$ pertence a E_k, então pela definição de E_k, existe m_j tal que $1 \leqslant m_j - n_j \leqslant k$ e

$$\varphi_{m_j - n_j}(f^{n_j}(x)) \leqslant (m_j - n_j)(\varphi_-(f^{n_j}(x)) + \epsilon).$$

Assim, obtemos a definição da sequência em (5.11). Pela subaditividade temos que

$$\begin{aligned} \varphi_n(x) \leqslant\ & \varphi_{n_1}(x) + \varphi_{m_1 - n_1}(f^{n_1}(x)) + \varphi_{n_2 - m_1}(f^{m_1}(x)) \\ & + \varphi_{m_2 - n_2}(f^{n_2}(x)) + \ldots + \varphi_{m_s - n_s}(f^{n_s}(x)) + \varphi_{n - m_s}(f^{m_s}(x)). \end{aligned} \tag{5.12}$$

Vamos primeiro analisar os elementos da forma $\varphi_{m_j - n_j}$.
Da definição dos conjuntos (E_k) temos que

$$
\begin{aligned}
\varphi_{m_1-n_1}(f^{n_1}(x)) &\leqslant (m_1-n_1)(\varphi_-(x)+\epsilon) \leqslant \sum_{i=n_1}^{m_1-1} \psi_k(f^i(x)) \\
\varphi_{m_2-n_2}(f^{n_2}(x)) &\leqslant (m_2-n_2)(\varphi_-(x)+\epsilon) \leqslant \sum_{i=n_2}^{m_2-1} \psi_k(f^i(x)) \\
\vdots \quad & \quad \vdots \\
\varphi_{m_s-n_s}(f^{n_s}(x)) &\leqslant (m_s-n_s)(\varphi_-(x)+\epsilon) \leqslant \sum_{i=n_s}^{m_s-1} \psi_k(f^i(x)).
\end{aligned}
$$

Observação 15. As últimas desigualdades à direita em cada linha se justificam da seguinte maneira. Para qualquer $l = \{1, \ldots, s\}$ e pela invariância de φ_- temos que

$$
\begin{aligned}
\varphi_{m_l-n_l}(f^{n_l}(x)) &\leqslant (m_l-n_l)(\varphi_-(x)+\epsilon) \\
&= (\varphi_-(f^{n_l})+\epsilon)+(\varphi_-(f^{n_l+1})+\epsilon)+\ldots+(\varphi_-(f^{m_l-1})+\epsilon) \\
&\leqslant \sum_{i=n_l}^{m_l-1} \psi_k(f^i(x)).
\end{aligned}
$$

Por outro lado, se agora analisamos os elementos da forma $\varphi_{n_{i+1}-m_i}$, lembre que neste caso $x \notin E_k$ então pela definição dos (ψ_k) temos

$$
\begin{aligned}
\varphi_{n_1}(x) &\leqslant \varphi_1(x)+\ldots+\varphi_1(f^{n_1-1}(x)) = \sum_{i=0}^{n_1-1} \psi_k(f^i(x)) \\
\varphi_{n_2-m_1}(f^{m_1}(x)) &\leqslant \varphi_1(f^{m_1}(x))+\ldots+\varphi_1(f^{n_2-1}(x)) = \sum_{i=m_1}^{n_2-1} \psi_k(f^i(x)) \\
\vdots \quad & \quad \vdots \\
\varphi_{n_s-m_{s-1}}(f^{m_{j-1}}(x)) &\leqslant \varphi_1(f^{m_{s-1}}(x))+\ldots+\varphi_1(f^{n_s-1}(x)) = \sum_{i=m_{s-1}}^{n_s-1} \psi_k(f^i(x)) \\
\varphi_{n-m_s}(f^{m_s}(x)) &\leqslant \varphi_1(f^{m_s}(x))+\ldots+\varphi_1(f^{n-1}(x)) = \sum_{i=m_s}^{n-1} \psi_k(f^i(x)).
\end{aligned}
$$

Se juntamos as duas análises feitas anteriormente em (5.12), obtemos o resultado

$$
\varphi_n(x) \leqslant \sum_{i=0}^{n-1} \psi_k(f^i(x)) \leqslant \sum_{i=0}^{n-k-1} \psi_k(f^i(x)) + \sum_{i=n-k}^{n-1} \max\{\psi_k, \varphi_1\}(f^i(x)). \tag{5.13}
$$

$\square$

Lema 8. *Se $\varphi : M \to \mathbb{R}$ é integravel com respeito à medida μ e $f : M \to M$ é uma transformação mensurável que preserva μ então para μ-quase todo ponto, temos*

$$
\lim_n \frac{1}{n} \varphi(f^n(x)) = 0.
$$

Demonstração. Basta mostrar que, para cada $\epsilon > 0$, o conjunto de $x \in M$ tal que $|\varphi(f^n(x))| \geqslant \epsilon n$ para infinitos $n \in \mathbb{N}$ tem medida nula. Isto é $\mu(\mathcal{C}_\epsilon) = 0$, onde

$$\mathcal{C}_\epsilon = \{x \in M : |\varphi(f^n(x))| \geqslant \epsilon n \text{ para infinitos valores de } n \in \mathbb{N}\}.$$

Assim, $\mathcal{C} = \cup_{i \geqslant 1} \mathcal{C}_{\frac{1}{i}}$ também tem medida nula. Primeiramente observemos que

$$\begin{aligned}
\sum_{n=1}^{\infty} \mu(\{x \in M \,;\, |\varphi(f^n(x))| \geqslant n\epsilon\}) &= \sum_{n=1}^{\infty} \mu(\{x \in M ; |\varphi(x)| \geqslant n\epsilon\}) \\
&= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \mu(\{x \in M \,;\, k \leqslant \frac{|\varphi(x)|}{\epsilon} < k+1\}) \\
&= \sum_{k=1}^{\infty} k\mu(\{x \in M \,;\, k \leqslant \frac{|\varphi(x)|}{\epsilon} < k+1\}) \\
&\leqslant \int_{\{|\varphi| > \epsilon\}} \frac{|\varphi(x)|}{\epsilon}\, d\mu < \infty.
\end{aligned}$$

Portanto, pelo Lema de Borel-Cantelli temos que $\mu(\mathcal{C}_\epsilon) = 0$. Pela definição de $\mathcal{C}_\epsilon$, para todo $x \notin \mathcal{C}_\epsilon$ existe $p \geqslant 1$ de modo que $|\varphi(f^n(x))| < \epsilon n$ para todo $n \geqslant p$, além disso $\mu(\mathcal{C}) = 0$. Portanto para todo $x \notin \mathcal{C}$ temos que

$$\lim_n \frac{1}{n} \varphi(f^n(x)) = 0.$$

□

Agora também podemos mostrar a seguinte versão forte do Lema 8.

Corolário 1. *Seja* $f : M \to M$ *uma aplicação mensurável,* μ *uma medida invariante por* f *e* $\varphi : M \to \mathbb{R}$ *uma função mensurável tal que a função* $\varphi \circ f - \varphi$ *é integravel com respeito à medida* μ*. Então para* μ*-quase todo ponto, temos que*

$$\lim_n \frac{1}{n} \varphi(f^n(x)) = 0.$$

Demonstração. Aplicando o teorema ergódico de Birkhoff à função integrável $\varphi \circ f - \varphi$ temos que a sequência $\frac{1}{n} \varphi \circ f^n$ converge em quase todo ponto para alguma função mensurável ψ. De fato:

$$\begin{aligned}
\psi(x) &= \lim_n \sum_{j=0}^{n-1} (\varphi \circ f - \varphi)(f^j(x)) = \lim_n \sum_{j=0}^{n-1} ((\varphi \circ f^{j+1})(x) - (\varphi \circ f^j)(x)) \\
&= \lim_n \frac{1}{n} \left(\sum_{j=0}^{n-2} (\varphi \circ f^{j+1})(x) + (\varphi \circ f^n)(x) - \sum_{j=1}^{n-1} (\varphi \circ f^j)(x) - \varphi(x) \right) \\
&= \lim_n \frac{1}{n} (\varphi \circ f^n(x) - \varphi(x)) \\
&= \lim_n \frac{1}{n} \varphi \circ f^n(x).
\end{aligned}$$

Por outro lado como a medida μ é f-invariante,

$$\mu\left(\left\{x \in M \ : \ \left|\frac{1}{n}(\varphi\circ f^n)(x)\right| \geqslant c\right\}\right) = \mu(\{y \in M \ : \ |\varphi(y)| > nc\}) \to 0 \quad \text{quando} \ \ n \to \infty.$$

Portanto a sequência $\frac{1}{n}\varphi\circ f^n$ converge para zero em medida, logo $\frac{1}{n_k}\varphi\circ f^{n_k}$ converge para zero quase todo ponto para alguma subsequência. Daí que $\psi(x) = 0$ em μ-quase todo ponto. $\square$

5.2 Desigualdade Inferior

Lema 9. *Sob as hipóteses do Teorema de Kingman, temos que*

$$\int \varphi_- \, d\mu = L. \tag{5.14}$$

Demonstração. Assuma primeiramente que $(\frac{\varphi_n}{n})$ é uniformemente limitada inferiormente, logo considere a sequência de funções $\left(\frac{\varphi_n}{n}+k\right)_n$ mensuráveis e não negativas, então pelo Lema de Fatou

$$\int \liminf_n \left(\frac{\varphi_n}{n}+k\right) \, d\mu \leqslant \liminf_n \int \left(\frac{\varphi_n}{n}+k\right) \, d\mu. \tag{5.15}$$

Pelas propriedades do limite inferior

$$\int \left(\liminf_n \frac{\varphi_n}{n} + \liminf_n k\right) \, d\mu \leqslant \int \liminf_n \left(\frac{\varphi_n}{n}+k\right) \, d\mu \leqslant \liminf_n \int \frac{\varphi_n}{n} \, d\mu + \liminf_n \int k \, d\mu$$

e como μ é uma medida de probabilidade

$$\begin{aligned} \int \liminf_n \frac{\varphi_n}{n} \, d\mu + \int k \, d\mu &\leqslant \liminf_n \int \frac{\varphi_n}{n} \, d\mu + \int k \, d\mu \\ \int \liminf_n \frac{\varphi_n}{n} \, d\mu &\leqslant \liminf_n \int \frac{\varphi_n}{n} \, d\mu \\ \int \varphi_- \, d\mu &\leqslant L < \infty. \end{aligned}$$

Portanto

$$\int \varphi_- \, d\mu \leqslant L. \tag{5.16}$$

Para a desigualdade contrária, lembre do resultado obtido no Lema 7. Usaremos tal resultado da seguinte forma. Primeiramente integramos a desigualdade do Lema 7 e depois dividimos por n, assim

$$\frac{1}{n}\int \varphi_n \, d\mu \leqslant \frac{1}{n}\sum_{i=0}^{n-k-1} \int \psi_k \circ f^i \, d\mu + \frac{1}{n}\sum_{i=n-k}^{n-1} \int \max\{\psi_k, \varphi_1\} \circ f^i \, d\mu.$$

Como f é μ-invariante

$$\frac{1}{n}\int \varphi_n \, d\mu \leqslant \frac{1}{n}\sum_{i=0}^{n-k-1} \int \psi_k \, d\mu + \frac{1}{n}\sum_{i=n-k}^{n-1} \int \max\{\psi_k, \varphi_1\} \, d\mu.$$

Pela não dependência do índice i, segue que

$$\frac{1}{n}\int \varphi_n \, d\mu \leqslant \frac{n-k}{n}\int \psi_k \, d\mu + \frac{k}{n}\int \max\{\psi_k, \varphi_1\} \, d\mu.$$

Assim, fixado k e tomando o limite quando $n \to \infty$

$$L \leqslant \int \psi_k \, d\mu. \tag{5.17}$$

Note que pela definição de ψ_k em (5.9), temos que

$$\begin{aligned}\int \psi_k \, d\mu &= \int_{E_k} \psi_k \, d\mu + \int_{E_k^c} \psi_k \, d\mu \\ &= \int_{E_k} (\varphi_- + \epsilon) \, d\mu + \int_{E_k^c} \varphi_1 \, d\mu. \end{aligned} \tag{5.18}$$

Logo, fazendo $k \to \infty$, de (5.17) e (5.18), e do fato de que $\epsilon > 0$ é arbitrario, concluímos finalmente que

$$L \leqslant \int \varphi_- \, d\mu. \tag{5.19}$$

Isto prova o resultado, mas com a hipótese de $(\frac{\varphi_n}{n})$ ser uniformemente limitado inferiormente. No caso geral, utilizaremos o método de truncagem, considerando para cada $c \in \mathbb{R}$, $c > 0$, as funções

$$\varphi_n^c = \max\{\varphi_n, -cn\} \quad \text{e} \quad \varphi_-^c = \max\{\varphi_-, -c\}.$$

A sequência de funções $(\varphi_n^c)_n$ é subaditiva. De fato:

$$\begin{aligned}\varphi_{m+n}^c &= \max\{\varphi_{m+n}, -c(m+n)\} \\ &= \frac{1}{2}(\varphi_{m+n} - c(m+n) + |\varphi_{m+n} + c(m+n)|) \\ &\leqslant \frac{1}{2}(\varphi_m + \varphi_n \circ f^m - cm - cn + |\varphi_m + \varphi_n \circ f^m + cm + cn|) \\ &\leqslant \frac{1}{2}(\varphi_m - cm + |\varphi_m + cm|) + \frac{1}{2}(\varphi_n \circ f^m - cn + |\varphi_n \circ f^m + cn|) \\ &= \max\{\varphi_m, -cm\} + \max\{\varphi_n \circ f^m, -cn\} \\ &= \varphi_m^c + \varphi_n^c \circ f^m.\end{aligned}$$

Pela definição, para cada $n \in \mathbb{N}$ temos que $\varphi_n^c/n \geqslant -c$ e

$$\liminf_n \frac{\varphi_n^c}{n} = \max\{\varphi_-, -c\} = \varphi_-^c.$$

Pelo provado anteriormente (com a hipótese de limitação uniforme), temos que

$$\int \varphi_-^c \, d\mu = L_c = \inf_n \frac{1}{n}\int \varphi_n^c \, d\mu. \tag{5.20}$$

Observação 16. Note que, dados $c, d \in \mathbb{R}$ tais que $c \leqslant d$ então

$$\varphi_n^d = \max\{\varphi_n, -dn\} \leqslant \max\{\varphi_n, -cn\} = \varphi_n^c.$$

Analogamente temos que $\varphi_-^d \leqslant \varphi_-^c$. Assim, para cada $n \in \mathbb{N}$ fixo temos uma sequência monótona de funções $(\varphi_n^k)_k$ e $(\varphi_-^k)_k$ de forma que $\lim_k \varphi_n^k = \varphi_n$ e $\lim_k \varphi_-^k = \varphi_-$.

Da observação acima e do Teorema da Convergência Monótona aplicado à sequência $(\varphi_n^k)_k$ obtemos

$$\int \varphi_n \, d\mu = \lim_k \int \varphi_n^k \, d\mu = \inf_k \int \varphi_n^k \, d\mu. \tag{5.21}$$

Da mesma forma para a sequência $(\varphi_-^k)_k$, temos que

$$\int \varphi_- \, d\mu = \lim_k \int \varphi_-^k \, d\mu = \inf_k \int \varphi_-^k \, d\mu. \tag{5.22}$$

Finalmente, de (5.21) e (5.22) juntamente com (5.20) obtemos

$$\int \varphi_- \, d\mu = \inf_c \int \varphi_-^c \, d\mu = \inf_c \inf_n \frac{1}{n} \int \varphi_n^c \, d\mu = \inf_n \inf_c \frac{1}{n} \int \varphi_n^c \, d\mu = \inf_n \frac{1}{n} \int \varphi_n \, d\mu = L.$$

Portanto

$$\int \varphi_- \, d\mu = L.$$

□

5.3 Desigualdade Superior

Lema 10. *Para qualquer* $k \in \mathbb{N}$ *fixo*

$$\limsup_n \frac{\varphi_{kn}}{n} = k \limsup_n \frac{\varphi_n}{n}. \tag{5.23}$$

Demonstração. Para mostrar uma das desigualdades utilizaremos uma propriedade do limite superior. Como $\left(\frac{\varphi_{kn}}{kn}\right)$ é uma subsequência de $\left(\frac{\varphi_n}{n}\right)$, então

$$\limsup_n \frac{\varphi_{kn}}{n} \leqslant k \limsup_n \frac{\varphi_n}{n}. \tag{5.24}$$

Para a desigualdade contrária, seja $n \in \mathbb{N}$, então

$$n = q_n k + r_n \quad \text{tal que} \quad 0 \leqslant r_n < k. \tag{5.25}$$

Pela subaditividade de $(\varphi_n)_n$ segue que

$$\varphi_n \leqslant \varphi_{q_n k} + \varphi_{r_n} \circ f^{q_n k} \leqslant \varphi_{q_n k} + \psi \circ f^{q_n k} \tag{5.26}$$

onde $\psi = \max\{\varphi_1^+, \ldots \varphi_k^+\} \geqslant \max\{\varphi_1, \ldots \varphi_k\}$. Uma vez que r_n está limitado por k, isto é $r_n < k$ e já que k está fixado, logo temos que $q_n \to \infty$ quando $n \to \infty$. De (5.25) obtemos

$$\frac{n}{q_n} = k + \frac{r_n}{q_n}.$$

Daí, $n/q_n \to k$ quando $n \to \infty$. Pela definição de ψ temos que $\psi \in L^1(\mu)$, então pelo Lema 8 segue que

$$\lim_n \frac{1}{n} \psi \circ f^{q_n k} = 0 \quad \mu\text{-quase sempre}. \tag{5.27}$$

Se consideramos o conjunto de medida total onde a igualdade (5.27) vale, e em (5.26) dividimos por n e tomamos o limite superior, obtemos

$$\limsup_n \frac{\varphi_n}{n} \leqslant \limsup_n \frac{\varphi_{q_n k}}{n} + \limsup_n \frac{\psi \circ f^{q_n k}}{n} = \limsup_n \frac{\varphi_{q_n k}}{n}. \tag{5.28}$$

Fazendo uso das propriedades do limite superior e tendo em conta que $\lim_n \frac{q_n}{n} = \frac{1}{k} > 0$, obtemos o seguinte

$$\limsup_n \frac{\varphi_{q_n k}}{n} = \limsup_n \frac{q_n}{n} \frac{\varphi_{q_n k}}{q_n} = \frac{1}{k} \limsup_q \frac{\varphi_{kq}}{q}. \tag{5.29}$$

De (5.29) e (5.28) segue que

$$\limsup_n \frac{\varphi_n}{n} \leqslant \limsup_n \frac{\varphi_{kn}}{kn} \tag{5.30}$$

Finalmente de (5.24) e (5.30) obtemos o resultado. □

Lema 11. *Considere as hipóteses do Teorema de Kingman, e suponha também que* $\inf \varphi_n > -\infty$ *para todo* $n \in \mathbb{N}$. *Então*

$$\int \varphi_+ \, d\mu \leqslant L.$$

Demonstração. Para cada k fixo e $n \geqslant 1$, considere as somas temporais de Birkhoff de φ_k com respeito a f^k, isto é, as funções

$$\theta_n(x) = \sum_{j=0}^{n-1} -\varphi_k(f^{jk}(x)). \tag{5.31}$$

Note que

$$\int \theta_n \, d\mu = -n \int \varphi_k \, d\mu \quad \text{para cada } n. \tag{5.32}$$

De fato, como f^k é uma transformação que preserva a medida μ, então

$$\begin{aligned} \int \theta_n \, d\mu &= \int \sum_{j=0}^{n-1} -\varphi_k(f^{jk}(x)) \, d\mu \\ &= \sum_{j=0}^{n-1} \int -\varphi_k \, d\mu \\ &= -n \int \varphi_k \, d\mu. \end{aligned}$$

Por outro lado, pela subaditividade da sequência $(\varphi_n)_n$

$$\begin{aligned} \varphi_{kn} &\leqslant \varphi_k + \varphi_k \circ f^k + \varphi_k \circ f^{2k} + \ldots + \varphi_k \circ f^{k(n-1)} \\ &= \sum_{j=0}^{n-1} \varphi_k \circ f^{jk} \\ &= -\theta_n. \end{aligned}$$

Portanto
$$\varphi_{kn} \leqslant -\theta_n \quad \text{para todo} \quad n. \tag{5.33}$$
Agora, defina $\theta_-(x) = \liminf_n \frac{1}{n}\theta_n(x)$, então fazendo uso de (5.33) e do Lema 10 afirmamos que $\theta_- \leqslant -k\varphi_+$. De fato
$$\begin{aligned}
\theta_- &= \liminf_n \frac{1}{n}\theta_n \\
&\leqslant \liminf_n \frac{-\varphi_{kn}}{n} \\
&= -\limsup_n \frac{\varphi_{kn}}{n} \\
&= -k\limsup_n \frac{\varphi_n}{n} \\
&= -k\varphi_+.
\end{aligned}$$
Assim
$$\int \theta_- \, d\mu \leqslant \int -k\varphi_+ \, d\mu. \tag{5.34}$$
Observe também que a sequência $(\theta_n)_n$ é aditiva, de fato, sejam $m, n \geqslant 1$, então
$$\begin{aligned}
\theta_{n+m} &= \sum_{j=0}^{m+n-1} -\varphi_k \circ f^{jk} \\
&= \sum_{j=0}^{m-1} -\varphi_k \circ f^{jk} + \sum_{j=m}^{m+n-1} -\varphi_k \circ f^{jk} \\
&= \sum_{j=0}^{m-1} -\varphi_k \circ f^{jk} + \sum_{i=0}^{n-1} -\varphi_k \circ f^{(i+m)k} \\
&= \sum_{j=0}^{m-1} -\varphi_k \circ f^{jk} + \sum_{i=0}^{n-1} -\varphi_k \circ f^{ik} \circ f^{mk} \\
&= \theta_m + \theta_n \circ f^{mk}.
\end{aligned}$$
Agora, de (5.31) temos que $\theta_1 = -\varphi_k$, logo θ_1 é limitada já que
$$\theta_1(x) = -\varphi_k(x) \leqslant \sup_x -\varphi_k(x) = -\inf_x \varphi_k(x) < \infty.$$
Assim, a função θ_1^+ e limitada é limitada e consequentemente integrável. Pelo Lema 9 e a igualdade em (5.32), temos que
$$\begin{aligned}
\int \theta_- \, d\mu &= \lim_n \frac{1}{n}\int \theta_n \, d\mu \\
&= -\int \varphi_k \, d\mu.
\end{aligned} \tag{5.35}$$
Então de (5.34) e (5.35), obtemos
$$\int \varphi_+ \, d\mu \leqslant -\frac{1}{k}\int \theta_- \, d\mu = \frac{1}{k}\int \varphi_k \, d\mu.$$

Finalmente, tomando o ínfimo sobre k obtemos que

$$\int \varphi_+ \, d\mu \leqslant L. \tag{5.36}$$

□

Demonstração. do Teorema de Kingman
Primeiramente mostramos que φ é invariante. Pela subaditividade da sequência $(\varphi_n)_n$ temos que

$$\varphi_- = \liminf_n \frac{\varphi_n}{n} \leqslant \liminf_n \frac{\varphi_1 + \varphi_{n-1} \circ f}{n} = \varphi_- \circ f.$$

Pelo Lema 3, φ_- é invariante, isto é $\varphi_- = \varphi_- \circ f$ em μ-quase todo ponto, analogamente $\varphi_+ = \varphi_+ \circ f$ em μ-quase todo ponto. Daí que $\varphi = \varphi \circ f$ em μ-quase todo ponto.

Finalizamos a demonstração do Teorema Ergódico Subaditivo. Dos Lemas 9 e 11 se verifica a desigualdade (5.6) e assim o Teorema 10 no caso de ser limitada inferiormente é valido. Lembremos agora das seguintes definições, para qualquer $c \in \mathbb{R}$ onde $c > 0$:

$$\varphi_n^c = \max\{\varphi_n, -cn\}, \qquad \varphi_+^c = \max\{\varphi_+, -c\} \quad \text{e} \quad \varphi_-^c = \max\{\varphi_-, -c\}.$$

Os argumentos anteriores podem ser aplicados à sequência $(\varphi_n^c)_n$ para qualquer $c > 0$ fixo, assim

$$\int \varphi_+^c \, d\mu = \int \varphi_-^c \, d\mu.$$

Daí que, para μ-quase todo ponto e qualquer $c > 0$ temos que $\varphi_+^c = \varphi_-^c$. Como $\lim_c \varphi_+^c = \varphi_+$ e $\lim_c \varphi_-^c = \varphi_-$ então segue que $\varphi_+ = \varphi_-$ μ-quase todo ponto. □

Vamos agora dar uma prova, utilizando o teorema ergódico subaditivo de Kingman, do teorema de Furstenberg-Kesten, que estende o teorema de Furstenberg provado no capítulo anterior.

Teorema 11 (Furstenberg- Kesten). *Seja* $(M, \mathcal{A}, \mu)$ *um espaço de probabilidade,* $f : M \to M$ *uma transformação mensurável que preserva a medida* μ, *e* $A : M \to GL(d, \mathbb{R})$ *uma aplicação mensurável tal que* $\log^+ \|A^{\pm 1}\| \in L^1(\mu)$, *então*

$$\lambda_+(x) = \lim_n \frac{1}{n} \log \|A^n(x)\| \quad e \quad \lambda_-(x) = \lim_n \frac{1}{n} \log \|(A^n(x))^{-1}\|^{-1}, \tag{5.37}$$

existem para μ*-quase todo ponto* $x \in M$. *Além disso, as funções* λ_+ *e* λ_- *são invariantes e* μ*-integráveis com*

$$\int \lambda_+(x) \, d\mu(x) = \lim_n \frac{1}{n} \int \log \|A^n(x)\| \, d\mu(x)$$

e

$$\int \lambda_-(x) \, d\mu(x) = \lim_n \frac{1}{n} \int \log \|(A^n(x))^{-1}\|^{-1} \, d\mu(x).$$

Demonstração. Defina as seguintes funções

$$\varphi_n(x) = \log \|A^n(x)\| \quad \text{e} \quad \psi_n(x) = \log \|(A^n(x))^{-1}\|. \tag{5.38}$$

Assim

$$\varphi_1(x) = \log \|A(x)\| \quad \text{e} \quad \psi_1(x) = \log \|(A(x))^{-1}\|. \tag{5.39}$$

Portanto, a hipótese implica que φ_1^+ e ψ_1^+ são integráveis em relação à medida μ, por conseguinte $\varphi_1(x)$, $\psi_1(x) \in [-\infty, +\infty)$ para μ-quase todo ponto $x \in M$. Finalmente pela subaditividade das sequências φ_n e ψ_n e pelo Teorema 10 temos que

$$\lim_n \frac{1}{n} \log \|A^n(x)\| = \lambda_+(x) \quad \text{e} \quad \lim_n \frac{1}{n} \log \|(A^n(x))^{-1}\|^{-1} = \lambda_-(x). \tag{5.40}$$

□

CAPÍTULO

6

TEOREMA ERGÓDICO MULTIPLICATIVO DE OSELEDETS

Para motivar o resultado que provaremos neste capítulo, considere uma única matriz A agindo em $\mathbb{R}^d$, com valores próprios positivos $e^{\lambda_1} > e^{\lambda_2} > \ldots > e^{\lambda_k}$ (talvez com multiplicidade). Dado um vetor $v \in \mathbb{R}^d$, consideremos o comportamento de $\|A^n \cdot v\|$ quando n é muito grande. Para um vetor qualquer, $\|A^n \cdot v\|$ irá crescer a uma taxa de $e^{n\lambda_1}$. Mas para v contido no espaço gerado pelos vetores próprios com valores próprios e^{λ_2} ou menor, a taxa de crescimento de $\|A^n \cdot v\|$ será diferente. Repetindo a análise isto dá uma decomposição de $\mathbb{R}^d$ por subespaços com diferentes ordens de crescimento sob iteração da matriz A.
Este resultado é conhecido como o Teorema Ergódico Multiplicativo de Oseledets. Neste capítulo demostraremos tal Teorema para cociclos em dimensão 2. A razão para incluir essa discussão em um caso especial é que ela pode ser manipulada por métodos muito mais simples e, assim, fornece um visão útil das principais características do teorema.

Considere $F : M \times \mathbb{R}^2 \to M \times \mathbb{R}^2$ como sendo $F(x,v) = (f(x), A(x) \cdot v)$, para alguma função mensurável $A : M \to SL(2,\mathbb{R})$ satisfazendo a condição de integrabilidade.

Teorema 12. *[Teorema de Oseledets (AVILA; BOCHI, 2008)] Seja $f : M \to M$ uma transformação que preserva a medida μ, e $A : M \to SL(2,\mathbb{R})$ um cociclo que satisfaz a condição de integrabilidade*

$$\int \log \|A(x)\| \, d\mu(x) < \infty.$$

Para μ-quase todo ponto $x \in M$, uma das condições abaixo vale:

1. $\lambda_-(x) = \lambda_+(x)$ *e*

$$\lim_{n\to\infty} \frac{1}{n} \log \|A^n(x)v\| = \lambda_\pm(x) \quad \textit{para todo} \quad v \in \mathbb{R}^2 \setminus \{0\}. \tag{6.1}$$

2. *Ou* $\lambda_-(x) < \lambda_+(x)$ *e existe um espaço vetorial unidimensional* $E^s_x \subset \mathbb{R}^2$ *tal que*

$$\lim_n \frac{1}{n} \log \|A^n(x)v\| = \begin{cases} \lambda_-(x), & \textit{se } v \in E^s_x \setminus \{0\} \\ \lambda_+(x), & \textit{se } v \in \mathbb{R}^2 \setminus E^s_x. \end{cases} \tag{6.2}$$

Além disso, o espaço E^s_x *depende mensuravelmente do ponto* x *e é invariante pelo cociclo, isto é* $A(x)E^s_x = E^s_{f(x)}$ *para todo* $x \in M$.

Demonstração. Vamos considerar $x \in M$ como na conclusão do Teorema 11 e seja $\lambda(x) = \lambda_+(x) = -\lambda_-(x)$. Primeiro, seja $x \in M$ tal que $\lambda(x) = 0$. Logo, dado $n \in \mathbb{N}$ e qualquer $v \in \mathbb{R}^2 \setminus \{0\}$ sabemos que

$$\|A^n(x)\|^{-1}\|v\| \leqslant \|A^n(x)v\| \leqslant \|A^n(x)\|\|v\|. \tag{6.3}$$

Tomando o logaritmo na desigualdade e dividindo por n obtemos o seguinte

$$\lim_n \frac{1}{n} \log \|A^n(x)\|^{-1}\|v\| \leqslant \lim_n \frac{1}{n} \log \|A^n(x)v\| \leqslant \lim_n \frac{1}{n} \log \|A^n(x)\|\|v\|,$$

e assim

$$\lim_n \frac{1}{n} \log \|A^n(x)\|^{-1} \leqslant \lim_n \frac{1}{n} \log \|A^n(x)v\| \leqslant \lim_n \frac{1}{n} \log \|A^n(x)\|.$$

Logo

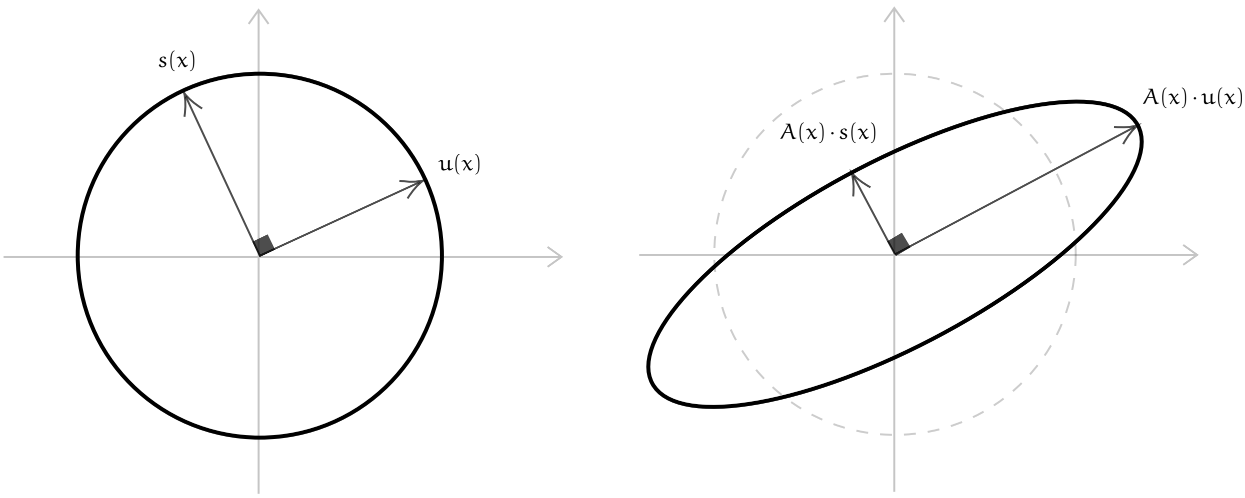
$$-\lambda_-(x) \leqslant \lim_n \frac{1}{n} \log \|A^n(x)v\| \leqslant \lambda_+(x).$$

Portanto, quando $\lambda(x) = 0$ obtemos o resultado.

Na segunda parte do teorema usamos a decomposição de Cartan. Para cada $A \in SL(2,\mathbb{R})$ existem duas rotações que denotaremos por R_1 e R_2 e uma matriz diagonal que vamos denotar por D tais que $A = R_1DR_2$, além disso $\|A\| = \|D\|$. Como estamos no caso em que $\lambda(x) > 0$, existe $n_0 \in \mathbb{N}$ tal que $\|A^n(x)\| \approx e^{n\lambda}$ é maior que um para todo $n \geqslant n_0$, caso contrário temos que $\lambda(x) = 0$. Então existem vetores ortonormais $s(x)$ e $u(x)$ tais que

$$\|A(x) \cdot s(x)\| = \|A(x)\|^{-1} \quad \text{e} \quad \|A(x) \cdot u(x)\| = \|A(x)\|.$$

Figura 1 – Vetores ortonormais $s(x)$ e $u(x)$



Fonte: Elaborada pelo autor.

Assim, para cada n definimos

$$\|A^n(x)\cdot s_n(x)\| = \|A^n(x)\|^{-1} \quad \text{e} \quad \|A^n(x)\cdot u_n(x)\| = \|A^n(x)\|$$

onde $u_n(x)$ e $s_n(x)$ são respectivamente o vetor mais expandido e o vetor mais contraído sob $A^n(x)$. A prova do teorema consiste em provar que o subespaço gerado por $s_n(x)$ é justamente o espaço procurado e no caso invertível o subespaço gerado por $u_n(x)$ convergirá também ao espaço procurado.

Lema 12. *O ângulo* $\alpha_n = \measuredangle(s_n(x), s_{n+1}(x))$ *decresce exponencialmente, com*

$$\limsup_n \frac{1}{n}\log|\sin\alpha_n| \leqslant -2\lambda(x).$$

Demonstração. Seja $\alpha_n = \measuredangle(s_n(x), s_{n+1}(x))$, isto é, o ângulo entre os vetores s_n e s_{n+1}. Logo, s_n se escreve como $s_n(x) = \cos\alpha_n \cdot s_{n+1}(x) + \sin\alpha_n \cdot u_{n+1}(x)$ na base ortonormal $\{s_{n+1}, u_{n+1}\}$. A imagem por $A^{n+1}(x)$ desta base, por definição, também é ortonormal, logo segue que

$$\begin{aligned} \|A^{n+1}(x)\cdot s_n(x)\| &= \|\cos\alpha_n(A^{n+1}(x)\cdot s_{n+1}) + \sin\alpha_n(A^{n+1}(x)\cdot u_{n+1})\| \\ &\geqslant \|\sin\alpha_n(A^{n+1}(x)\cdot u_{n+1})\| \\ &= |\sin\alpha_n|\|A^{n+1}(x)\|. \end{aligned} \tag{6.4}$$

Por outro lado, pela identidade de cociclos temos que

$$\begin{aligned} \|A^{n+1}(x)\cdot s_n(x)\| &= \|A(f^n(x))\cdot A^n(x)\cdot s_n(x)\| \\ &\leqslant \|A(f^n(x))\|\|A^n(x)\cdot s_n(x)\| \\ &= \|A(f^n(x))\|\|A^n(x)\|^{-1}. \end{aligned} \tag{6.5}$$

Das desigualdades em (6.4) e (6.5) obtemos que

$$|\sin\alpha_n| \leqslant \frac{\|A(f^n(x))\|}{\|A^n(x)\|\|A^{n+1}(x)\|}. \tag{6.6}$$

Tomando o logaritmo e dividindo tudo entre n temos

$$\limsup_n \frac{1}{n}|\sin\alpha_n| \leqslant \limsup_n \frac{\log\|A(f^n(x))\|}{n} - \limsup_n \frac{\log\|A^n(x)\|}{n} - \limsup_n \frac{\log\|A^{n+1}(x)\|}{n}. \tag{6.7}$$

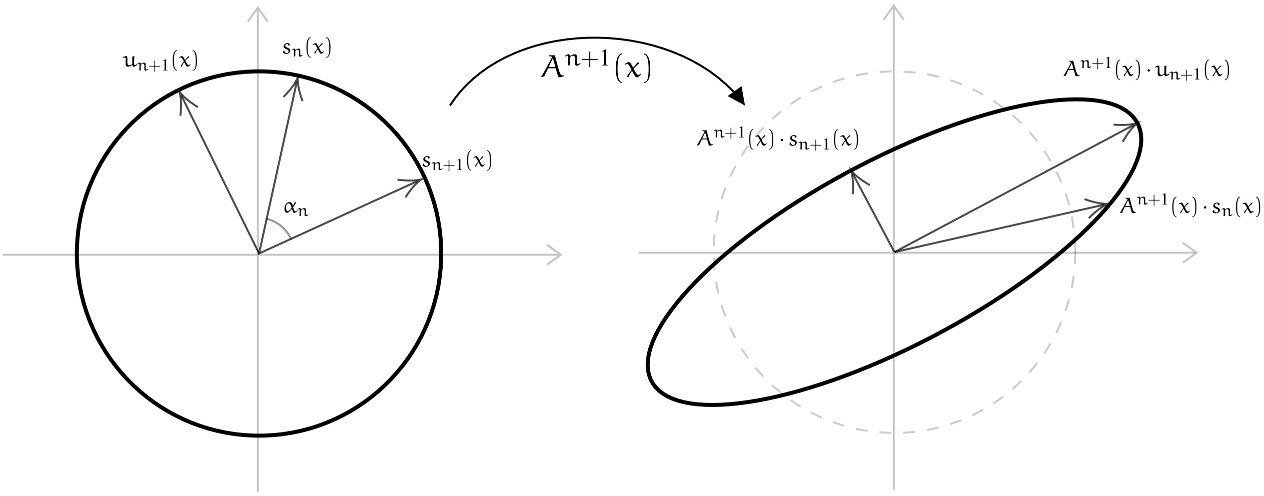
Pelo Corolário 1 podemos garantir que para μ-quase todo ponto $x \in M$ temos

$$\lim_n \frac{1}{n}\log\|A(f^n(x))\| = 0.$$

Portanto de (6.7) temos que

$$\limsup_n \frac{1}{n}|\sin\alpha_n| \leqslant -2\lambda(x).$$

Figura 2 – Imagem por $A^{n+1}(x)$ da base ortonormal $\{s_{n+1}, u_{n+1}\}$.



Fonte: Elaborada pelo autor.

□

Lema 13. *A sequência $(s_n(x))_n$ é de Cauchy no espaço projetivo.*

Demonstração. Seja $\epsilon > 0$ tal que $-2\lambda(x) + \epsilon < 0$. Pelo Lema 12 e da definição do limite superior, existe $n_0 \in \mathbb{N}$ tal que para todo $n \geqslant n_0$

$$|\sin\alpha_n| \leqslant e^{n(-2\lambda(x)+\epsilon)}. \tag{6.8}$$

Considere o intervalo $K = [0, 2\pi]$. Se o ângulo entre $s_n(x)$ e $s_{n+1}(x)$ está em $[0, \frac{2\pi}{3}]$, sabemos que $\sin\frac{\alpha_n}{2} \leqslant \sin\alpha_n$, logo

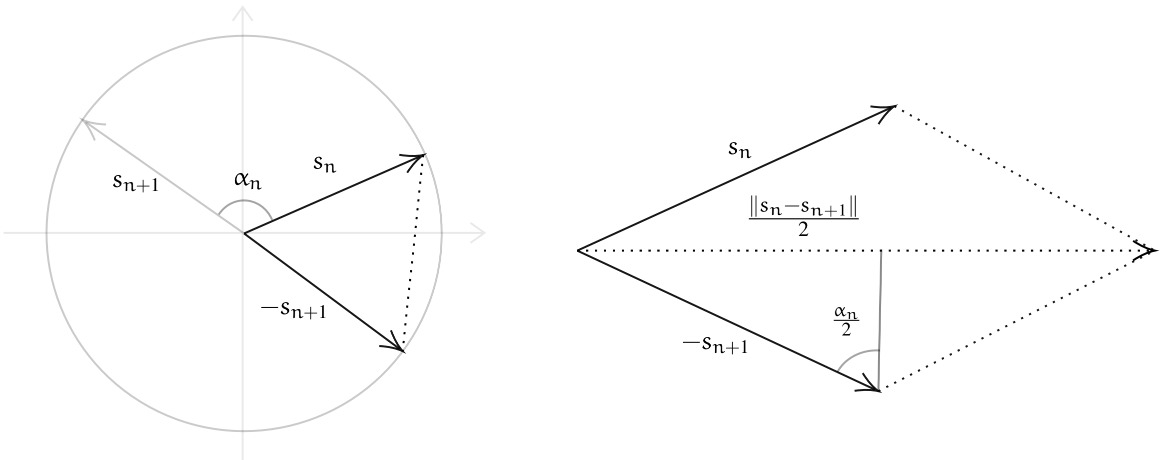
$$\frac{1}{2}\|s_n(x) - s_{n+1}(x)\| = \sin\frac{\alpha_n}{2} \leqslant |\sin\alpha_n|. \tag{6.9}$$

Pela desigualdade em (6.8) temos que

$$\|s_n(x) - s_{n+1}(x)\| \leqslant 2e^{n(-2\lambda(x)+\epsilon)}. \tag{6.10}$$

Se o ângulo entre $s_n(x)$ e $s_{n+1}(x)$ está em $K \setminus [0, \frac{2\pi}{3}]$ então, substituindo alguns $s_j(x)$ por $-s_j(x)$ como se mostra na Figura 3 obtemos o mesmo resultado que em (6.11).

Figura 3 – Vetores unitários $s_n(x)$ e $s_{n+1}(x)$.



Fonte: Elaborada pelo autor.

Consequentemente, existe $C \in \mathbb{R}$, $C > 0$, tal que para cada $k \geqslant 1$ e n suficientemente grande

$$\|s_{n+k}(x) - s_n(x)\| \leqslant Ce^{n(-2\lambda(x)+\epsilon)}. \tag{6.11}$$

Em particular, a sequência é de Cauchy. $\square$

Defina $s(x) = \lim_n s_n(x)$ sempre que o limite existe. Seja $\beta_n(x)$ o ângulo entre $s(x)$ e $s_n(x)$. Como α_n diminui exponencialmente, o mesmo acontece com $\beta_n(x)$, com a mesma taxa exponencial. Logo

$$\limsup_n \frac{1}{n} \log|\sin \beta_n| \leqslant -2\lambda(x). \tag{6.12}$$

Lema 14. *O vetor $s(x)$ é contraído à taxa de $-\lambda(x)$, isto é:*

$$\lim_n \frac{1}{n} \log \|A^n(x)s(x)\| = -\lambda(x).$$

Demonstração. Seja $\beta_n = \measuredangle(s(x), s_n(x))$. Logo, $s(x)$ se escreve como $s(x) = \cos \beta_n \cdot s_n(x) + \sin \beta_n \cdot u_n(x)$ na base ortonormal $\{s_n, u_n\}$. A imagem por $A^{n+1}(x)$ desta base, por definição, também é ortonormal, logo segue que

$$\begin{aligned}
\|A^n(x) \cdot s(x)\| &= \|\cos \beta_n (A^n(x) \cdot s_n) + \sin \beta_n (A^n(x) \cdot u_n)\| \\
&\geqslant \|\cos \beta_n (A^n(x) \cdot s_n(x))\| \\
&= |\cos \beta_n| \|A^n(x)\|^{-1}.
\end{aligned}$$

Então

$$\begin{aligned}\liminf_n \frac{1}{n}\log\|A^n(x)\cdot s(x)\| &\geqslant \liminf_n \frac{1}{n}\log|\cos\beta_n|\|A^n(x)\|^{-1}\\ &= \liminf_n \frac{1}{n}\log|\cos\beta_n| - \lim_n \frac{1}{n}\log\|A^n(x)\|\\ &= 0-\lambda(x).\end{aligned}$$

Portanto

$$\liminf_n \frac{1}{n}\log\|A^n(x)\cdot s(x)\| \geqslant -\lambda(x). \tag{6.13}$$

Por outro lado como $\|A^n(x)\cdot s(x)\| \leqslant |\cos\beta_n|\|A^n(x)\cdot s_n\| + |\sin\beta_n|\|A^n(x)\cdot u_n\|$, temos que

$$\begin{aligned}&\limsup_n \frac{1}{n}\log\|A^n(x)\cdot s(x)\|\\ &\leqslant \limsup_n \frac{1}{n}\log(|\cos\beta_n|\|A^n(x)s_n(x)\| + |\sin\beta_n|\|A^n(x)u_n(x)\|)\\ &= \max\left\{\limsup_n \frac{1}{n}\log|\cos\beta_n|\|A^n(x)\cdot s_n(x)\|\ ,\ \limsup_n \frac{1}{n}\log|\sin\beta_n|\|A^n(x)\cdot u_n(x)\|\right\}\\ &\leqslant \max\left\{\limsup_n \frac{1}{n}\log\|A^n(x)\|^{-1}\ ,\ \limsup_n \frac{1}{n}\log|\sin\beta_n| + \limsup_n \frac{1}{n}\log\|A^n(x)\|\right\}\\ &= \max\{-\lambda(x)\ ,\ -2\lambda(x)+\lambda(x)\}\\ &= -\lambda(x).\end{aligned}$$

Assim

$$\lim_n \frac{1}{n}\log\|A^n(x)s(x)\| = -\lambda(x).$$

□

O lema anterior nos dá uma idéia de quem poderia ser o espaço E^s_x no Teorema de Oseledets, mas é necessário provar uma condição que será mostrada no seguinte lema. Considere um vetor $v\in\mathbb{R}^2$ que não seja colinear com $s(x)$ e nós chamamos γ_n ao ângulo entre os vetores $s_n(x)$ e v.

Lema 15. *Seja $v\in\mathbb{R}^2$ que não seja colinear com $s(x)$, então*

$$\lim_n \frac{1}{n}\log\|A^n(x)\cdot v\| = \lambda(x). \tag{6.14}$$

Demonstração. Denote $\gamma_n = \measuredangle(v, s_n(x))$, então podemos escrever $v = \cos\gamma_n\cdot s_n(x) + \sin\gamma_n\cdot u_n(x)$, e assim

$$\begin{aligned}\|A^n(x)\cdot v\| &= \|\cos\gamma_n(A^n(x)\cdot s_n(x)) + \sin\gamma_n(A^n(x)\cdot u_n(x))\|\\ &\geqslant |\sin\gamma_n|\|A^n(x)\cdot u_n(x)\|\\ &= |\sin\gamma_n|\|A^n(x)\|.\end{aligned}$$

Observe que para $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ de modo que para todo $n \geqslant n_0$ temos que $\epsilon < |\sin \gamma_n| \leqslant 1$ já que v não é colinear com $s(x)$, logo

$$\begin{aligned}
\liminf_n \frac{1}{n} \log \|A^n(x) \cdot v\| &\geqslant \liminf_n \frac{1}{n} \log |\sin \gamma_n| \|A^n(x)\| \\
&= \liminf_n \frac{1}{n} \log \epsilon + \liminf_n \frac{1}{n} \log \|A^n(x)\| \\
&= 0 + \lambda(x).
\end{aligned}$$

Portanto

$$\liminf_n \frac{1}{n} \log \|A^n(x) \cdot v\| \geqslant \lambda(x). \tag{6.15}$$

Por outro lado, como $\|A^n(x) \cdot v\| \leqslant \|A^n(x)\| \|v\|$, então temos a desigualdade contrária, isto é

$$\limsup_n \frac{1}{n} \log \|A^n(x) \cdot v\| \leqslant \lim_n \frac{1}{n} \log \|A^n(x)\| = \lambda(x). \tag{6.16}$$

Assim, temos que

$$\lim_n \frac{1}{n} \log \|A^n(x) \cdot v\| = \lambda(x). \tag{6.17}$$

□

Lema 16. $A(x)s(x)$ *é colinear a* $s(f(x))$.

Demonstração. De fato, pelo Lema 14 temos que

$$\begin{aligned}
-\lambda(x) &= \lim_n \frac{1}{n+1} \log \|A^{n+1}(x) \cdot s(x)\| \\
&= \lim_n \frac{n}{n+1} \frac{1}{n} \log \|A^n(f(x)) \cdot A(x)s(x)\| \\
&= \lim_n \frac{1}{n} \log \|A^n(f(x)) \cdot A(x)s(x)\|.
\end{aligned}$$

Pelo Lema 15 para v não colinear a $s(f(x))$ temos

$$\begin{aligned}
\lim_n \frac{1}{n} \log \|A^n(f(x)) \cdot v\| &= \lambda(f(x)) \\
&= \lambda(x).
\end{aligned}$$

Como $\lambda(x) \neq 0$, segue que $A(x)s(x)$ é colinear a $s(f(x))$. □

Parte final da prova do Teorema de Oseledets. Tome $E^s_x \subset \mathbb{R}^2$ como sendo a reta $\mathbb{R}s(x) = \{as(x) : a \in \mathbb{R}\}$ gerada por $s(x)$. O Lema 14 prova a primeira afirmação em (6.2), o Lema 15 prova a segunda afirmação em (6.2). Além disso temos que $A(x)E^s_x = E^s_{f(x)}$ já que dado $s(x)$ em E^s_x então o Lema 16 diz que $A(x)s(x) \in E^s_{f(x)}$ portanto E^s_x é invariante pelo cociclo. □

Teorema 13. *[Caso invertível] Seja* $f : M \to M$ *uma transformação bijetiva que preserva a medida* μ *e* $A : M \to SL(2, \mathbb{R})$ *um cociclo que satisfaz a condição de integrabilidade*

$$\int \log \|A(x)\| \, d\mu(x) < \infty.$$

Para μ*-quase todo ponto* $x \in M$*, uma das condições abaixo vale:*

1. $\lambda_-(x) = \lambda_+(x)$ *e*

$$\lim_n \frac{1}{n} \log \|A^n(x)v\| = \lambda_\pm(x) \quad \textit{para todo} \quad v \in \mathbb{R}^2. \tag{6.18}$$

2. *Ou* $\lambda_-(x) < \lambda_+(x)$ *e existe uma decomposição de soma direta* $\mathbb{R}^2 = E^s_x \oplus E^u_x$ *tal que*

$$\lim_{n\to+\infty} \frac{1}{n} \log \|A^n(x)\cdot v\| = \begin{cases} \lambda_-(x), & \textit{se } v \in E^s_x \setminus \{0\} \\ \lambda_+(x), & \textit{se } v \in \mathbb{R}^2 \setminus E^s_x \end{cases} \tag{6.19}$$

$$\lim_{n\to-\infty} \frac{1}{n} \log \|A^n(x)\cdot v\| = \begin{cases} \lambda_+(x), & \textit{se } v \in E^u_x \setminus \{0\} \\ \lambda_-(x), & \textit{se } v \in \mathbb{R}^2 \setminus E^u_x. \end{cases} \tag{6.20}$$

Além disso, neste último caso, os espaços E^s_x *e* E^u_x *dependem mensuravelmente do ponto* x*, são invariantes pelo cociclo, isto é* $A(x)E^s_x = E^s_{f(x)}$ *e* $A(x)E^u_x = E^u_{f(x)}$ *para todo* $x \in M$*, e o ângulo entre eles diminui subexponencialmente ao longo de órbitas, isto é:*

$$\lim_n \frac{1}{n} \log |\sin \measuredangle(E^s_{f^n(x)}, E^u_{f^n(x)})| = 0. \tag{6.21}$$

Vamos precisar dos resultados abaixo para poder dar a prova do teorema de Oseledets (caso invertível).

Definição 26. Seja $\mathbb{RP}^d$ o espaço projetivo de $\mathbb{R}^d$. Para cada vetor não nulo de $\mathbb{R}^d$, denotamos por $\overline{x}$ sua projeção em $\mathbb{RP}^d$. Então para $\overline{x}$ e $\overline{y}$ em $\mathbb{RP}^d$, definimos a *métrica no espaço projetivo* como sendo

$$d(\overline{x},\overline{y}) = |\sin \measuredangle(x,y)|.$$

Proposição 5. *Sejam* $p, q \in \mathbb{R}^n \setminus \{0\}$*, então*

$$\left\| \frac{p}{\|p\|} - \frac{q}{\|q\|} \right\| \leqslant \max\left\{ \frac{1}{\|p\|}, \frac{1}{\|q\|} \right\} \|p-q\|.$$

Demonstração. Dado os vetores u, v em $\mathbb{R}^n$ com $\|u\| \geqslant \|v\| = 1$, então

$$\left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\| \leqslant \|u-v\|. \tag{6.22}$$

Suponha que $\|p\| \leqslant \|q\|$, logo

$$\max\left\{ \frac{1}{\|p\|}, \frac{1}{\|q\|} \right\} = \frac{1}{\|p\|}.$$

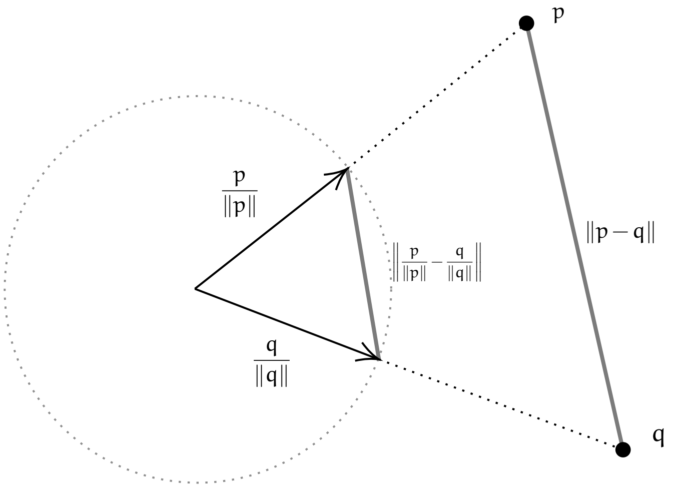
Consideremos $u = \frac{q}{\|p\|}$ e $v = \frac{p}{\|p\|}$, substituindo em (6.22) nós conseguimos obter

$$\begin{aligned} \left\| \frac{q}{\|q\|} - \frac{p}{\|p\|} \right\| = \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\| &\leqslant \|u-v\| \\ &= \left\| \frac{q}{\|p\|} - \frac{p}{\|p\|} \right\| \\ &= \frac{1}{\|p\|}\|q-p\| \\ &= \max\left\{ \frac{1}{\|p\|}, \frac{1}{\|q\|} \right\} \|q-p\|. \end{aligned}$$

Portanto

$$\left\| \frac{p}{\|p\|} - \frac{q}{\|q\|} \right\| \leqslant \max\left\{ \frac{1}{\|p\|} \,,\, \frac{1}{\|q\|} \right\} \|p-q\|.$$

Figura 4 – Interpretação geométrica da desigualdade



Fonte: Elaborada pelo autor.

□

Proposição 6. *Seja* $g \in GL(d,\mathbb{R})$ *e considere a ação projetiva de* $g \in GL(d,\mathbb{R})$*,* $\varphi_g : \mathbb{RP}^d \to \mathbb{RP}^d$ *como sendo* $\varphi_g(x) = \frac{g(x)}{\|g(x)\|}$*, então*

$$\frac{1}{\|g\|\|g^{-1}\|} \leqslant \frac{|\sin\measuredangle(g(x),g(y))|}{|\sin\measuredangle(x,y)|} \leqslant \|g\|\|g^{-1}\|. \tag{6.23}$$

Demonstração. Para $x \in \mathbb{RP}^d$ e $v \in T_v\mathbb{RP}^d$ temos que

$$\begin{aligned} D\varphi_g(x))v &= \lim_{\lambda\to 0}\frac{1}{\lambda}(\varphi_g(x+\lambda v)-\varphi_g(x)) \\ &= \lim_{\lambda\to 0}\frac{1}{\lambda}\left(\frac{g(x+\lambda v)}{\|g(x+\lambda v)\|}-\frac{g(x)}{\|g(x)\|}\right). \end{aligned}$$

Se φ_g é Lipschitz, então o resultado segue. De fato pelo Lema 5 temos que

$$\begin{aligned} \|D\varphi_g(x))v\| &= \|\lim_{\lambda\to 0}\frac{1}{\lambda}(\varphi_g(x+\lambda v)-\varphi_g(x))\| \\ &\leqslant \lim_{\lambda\to 0}\frac{1}{|\lambda|}\left\|\frac{g(x+\lambda v)}{\|g(x+\lambda v)\|}-\frac{g(x)}{\|g(x)\|}\right\| \\ &\leqslant \lim_{\lambda\to 0}\frac{1}{|\lambda|}\max\left\{\frac{1}{\|g(x+\lambda v)\|},\frac{1}{\|g(x)\|}\right\}|\lambda|\|g(v)\| \\ &= \lim_{\lambda\to 0}\max\left\{\frac{1}{\|g(x+\lambda v)\|},\frac{1}{\|g(x)\|}\right\}\|g(v)\| \\ &= \frac{\|g(v)\|}{\|g(x)\|} \\ &\leqslant \frac{\|g\|\|v\|}{m(g)}. \end{aligned}$$

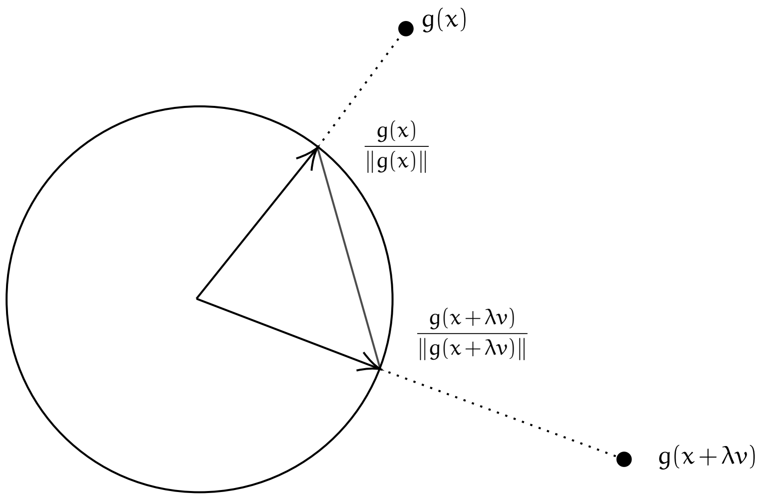
onde $m(g) = \|g^{-1}\|^{-1}$. Daí que

$$\frac{m(g)}{\|g\|} \leqslant \|D\varphi_g(x)\| \leqslant \frac{\|g\|}{m(g)}.$$

Portanto, como $d(\varphi_g(x), \varphi_g(y)) = |\sin\measuredangle(g(x), g(y))|$ então

$$\frac{m(g)}{\|g\|} \leqslant \frac{|\sin\measuredangle(g(x), g(y))|}{|\sin\measuredangle(x,y)|} \leqslant \frac{\|g\|}{m(g)}. \tag{6.24}$$

Figura 5 – Ação projetiva.



Fonte: Elaborada pelo autor.

□

Lema 17. *Seja* $\theta(y) = \measuredangle(E^s_y, E^u_y)$. *Para* μ*-quase todo* x *em* M *com* $\lambda(x) > 0$, *temos*

$$\lim_{n\to\pm\infty} \frac{1}{n} \log|\sin\theta(f^n(x))| = 0. \tag{6.25}$$

Demonstração. Para a prova deste Lema fazemos uso da desigualdade obtida na Proposição 6, então

$$\frac{m(A(x))}{\|A(x)\|} \leqslant \frac{|\sin\measuredangle(A(x)E^s_x, A(x)E^u_x)|}{|\sin\measuredangle(E^s_x, E^u_x)|} \leqslant \frac{\|A(x)\|}{m(A(x))}.$$

Como os espaços E^s_x e E^u_x são invariantes pelo cociclo, além disso como $A(x) \in SL(2,\mathbb{R})$ implica que $\|A(x)\| = \|A(x)^{-1}\|$. Dito de outro modo $\|A(x)\|^{-1} = m(A(x))$, logo

$$m(A(x))^2 \leqslant \frac{|\sin\theta(f(x))|}{|\sin\theta(x)|} \leqslant \|A(x)\|^2.$$

Tomando o logaritmo na desigualdade, obtemos

$$|\log|\sin\theta(f(x))| - \log|\sin\theta(x)|| \leqslant 2\log\|A(x)\|.$$

Se integramos ambos lados da desigualdade e lembramos da condição de integrabilidade temos que

$$\int |\log|\sin\theta(f(x))| - \log|\sin\theta(x)|| \; d\mu(x) \leqslant 2\int \log\|A(x)\| \; d\mu(x) < \infty.$$

Portanto $\log|\sin\theta(f(x))| - \log|\sin\theta(x)| \in L^1(\mu)$. Se denotamos por $\varphi(x) = \log|\sin(x)|$ temos pelo Corolário 1 que

$$\lim_{n\to\pm} \frac{1}{n}\log|\sin\measuredangle(E^s_{f^n(x)}, E^u_{f^n(x)})| = 0.$$

□

Lema 18. *Os vetores* $u(x)$ *e* $s(x)$ *são não colineares para* μ*-quase todo ponto em* $\{\, x \,:\lambda(x) > 0\}$.

Demonstração. Este lema diz que $E^s_x \neq E^u_x$, então é suficiente provar que

$$\lim_{n\to-\infty} \frac{1}{n}\log\|A^n(x)|E^s_x\| = -\lambda(x) \tag{6.26}$$

para μ-quase todo ponto em $\{\, x \,: \lambda(x) > 0\}$. Seja

$$\psi_n(x) = \frac{1}{n}\log\|A^{-n}(x)|E^s_x\|$$

e

$$\phi_n(x) = \frac{1}{n}\log\|(A^n(x)|E^s_x)^{-1}\|.$$

Lembre-se também da seguinte propriedade dos cociclos

$$A^{m+n}(x) = A^m(f^n(x))A^n(x),$$

para todo $x \in M$ e para todo $m, n \in \mathbb{Z}$. Então em particular temos que

$$(A^n(f^{-n}(x))|E^s_{f^{-n}(x)})^{-1} = A^{-n}(x)|E^s_x.$$

Daí que

$$\begin{aligned}
\phi_n \circ f^{-n}(x) &= \phi_n(f^{-n}(x)) \\
&= \frac{1}{n}\log\|(A^n(f^{-n}(x))|E^s_{f^{-n}(x)})^{-1}\| \\
&= \frac{1}{n}\log\|A^{-n}(x)|E^s_x\| \\
&= \psi_n(x).
\end{aligned}$$

Assim, $\phi_n \circ f^{-n} = \psi_n$ e por (6.2) temos que μ-quase todo ponto $x \in M$. O espaço gerado por $s(x)$ é contraído à taxa de $-\lambda(x)$, isto é ϕ_n converge para $-\lambda$. Em particular a sequência ϕ_n converge para $-\lambda$ em medida, isto é, dado $\epsilon > 0$

$$\lim_n \mu(\{x \in M \;:\; |\phi_n(x) + \lambda(x)| > \epsilon\}) = 0. \tag{6.27}$$

Como λ é f-invariante, segue que para todo $k \in \mathbb{Z}$

$$\begin{aligned}\{x \in M : |\phi_n(x) + \lambda(x)| > \epsilon\} &= \{x \in M : |\phi_n(f^k(x)) + \lambda(f^k(x))| > \epsilon\} \\ &= \{x \in M : |\phi_n(f^k(x)) + \lambda(x)| > \epsilon\}.\end{aligned}$$

Então, em particular temos que

$$\begin{aligned}\lim_n \mu(\{x \in M : |\phi_n(x) + \lambda(x)| > \epsilon\}) &= \lim_n \mu(\{x \in M : |\phi_n(f^{-n}(x)) + \lambda(x)| > \epsilon\}) \\ &= \lim_n \mu(\{x \in M : |\psi_n(x) + \lambda(x)| > \epsilon\}).\end{aligned}$$

Além disso, de (6.27) temos que

$$\lim_n \mu(\{x \in M : |\psi_n(x) + \lambda(x)| > \epsilon\}) = 0. \tag{6.28}$$

De (6.28) concluímos que ψ_n converge para $-\lambda$ em medida e como ψ_n converge para ψ então $\psi = -\lambda$ μ-quase todo ponto. Assim temos que

$$\lim_n \frac{1}{n} \log \|A^{-n}(x)|E_x^s\| = -\lambda(x) \tag{6.29}$$

e por outro lado

$$\lim_n \frac{1}{n} \log \|A^{-n}(x)|E_x^u\| = \lambda(x). \tag{6.30}$$

Portanto $E_x^s \neq E_x^u$. □

Demonstração. Teorema de Oseledets caso invertível
Vamos considerar $x \in M$ como na conclusão do Teorema 11, e $\lambda(x) = \lambda_+(x) = -\lambda_-(x)$. O caso (6.18) é obtido diretamente do Teorema 12 aplicado a F e à inversa de F. Para o caso em que $\lambda(x) > 0$ considere E_x^s e E_x^u os subespaços obtidos ao aplicar o teorema 12 a F e a inversa F^{-1} respectivamente. O Lema 18 garante que esses dois subespaços são transversais e o Lema 17 garante que o ângulo diminui subexponencialmente. □

CAPÍTULO

7

TEOREMA PRINCIPAL

Teorema 14 (Teorema de Furstenberg (AVILA; BOCHI, 2008))**.** *Seja* μ *uma medida de probabilidade em* $\mathrm{SL}(2,\mathbb{R})$ *satisfazendo a condição de integrabilidade*

$$\int_{\mathrm{SL}(2,\mathbb{R})} \log \|M\| \; d\mu(M) < \infty.$$

Denote por G_μ *o menor subgrupo fechado de* $\mathrm{SL}(2,\mathbb{R})$ *que contém* $\mathrm{supp}(\mu)$. *Assuma*

(i) G_μ *não é compacto,*

e que uma das seguintes condições é satisfeita:

(ii) Não há um conjunto finito, não vazio $L \subset \mathbb{RP}^2$ *tal que* $M(L) = L$ *para* μ*-quase todo* M.

(ii') Não há conjunto $L \subset \mathbb{RP}^2$ *com 1 ou 2 elementos tal que* $M(L) = L$ *para* μ*-quase todo* M.

Então o expoente de Lyapunov λ *é positivo.*

Proposição 7. *Se* μ *satisfaz* (i) *do Teorema de Furstenberg, então* (ii) *e* (ii') *são equivalentes.*

Demonstração. É imediato que (ii) implica (ii'). Para provar o contrário vamos mostrar que (i) e (ii') implica (ii). De fato, sejam $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ elementos distintos de $\mathbb{RP}^2$ de modo que para todo $M \in G_\mu$

$$M(\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k\}) = \{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k\}.$$

Note que, cada $M \in G_\mu$ induz uma permutação $\pi(M)$ de $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k\}$ e

$$\pi : G_\mu \to S_k$$

é um homomorfismo de grupos. O kernel de π é dado por

$$\begin{aligned} \ker(\pi) &= \{\, M \in G_\mu \;:\; \pi(M) = \mathrm{Id} \,\} \\ &= \{\, M \in G_\mu \;:\; M \cdot \bar{v}_i = \bar{v}_i \;\text{ para todo }\; i = 1, 2, \ldots, k. \,\}. \end{aligned}$$

O kernel de π é um subgrupo normal e fechado de G_μ. Pelo primeiro teorema dos isomorfismos temos que

$$\frac{G_\mu}{\ker(\pi)} \simeq F \subset \pi(G_\mu)$$

assim, $\frac{G_\mu}{\ker(\pi)}$ é finito já que a cardinalidade de S_k é menor ou igual a $k!$. Pelo item (i) temos que $\ker(\pi)$ e um subgrupo de G_μ que não é compacto, então $\ker(\pi)$ não é finito.

Agora, se $k \geqslant 3$ considere representantes das classes $\overline{v}_1$,$\overline{v}_2$ e $\overline{v}_3$ denotados por v_1, v_2 e v_3 respectivamente. Podemos escrever a seguinte igualdade para α, $\beta \in \mathbb{R} \setminus \{0\}$

$$v_3 = \alpha v_1 + \beta v_2.$$

Como $M \in \ker(\pi)$, existem ρ_1, ρ_2 e ρ_3 em $\mathbb{R}$ tal que $Mv_1 = \rho_1 v_1$, $Mv_2 = \rho_2 v_2$ e $Mv_3 = \rho_3 v_3$, logo temos

$$\begin{aligned}
\rho_3(\alpha v_1 + \beta v_2) &= \rho_3 v_3 \\
&= M v_3 \\
&= M(\alpha v_1 + \beta v_2) \\
&= M\alpha v_1 + M\beta v_2 \\
&= \alpha M v_1 + \beta M v_2 \\
&= \alpha \rho_1 v_1 + \beta \rho_1 v_2
\end{aligned}$$

então

$$(\rho_3 \alpha - \rho_1 \alpha) v_1 + (\rho_3 \beta - \rho_2 \beta) v_2 = 0.$$

Daí segue $\rho_3 = \rho_2 = \rho_1 = \rho$ portanto $M = \rho Id$. Tomando o determinante em ambos lados, temos $\det(M) = \det(\rho Id) = \rho^2 \det(Id) = \rho^2$, logo $\rho^2 = 1$ já que $M \in SL(2, \mathbb{R})$. Assim $M = \pm Id$. Isso prova que $\ker(\pi) \subseteq \{\pm Id\}$. Segue que $\ker(\pi)$ é finito, o que é uma contradição. Portanto $k \leqslant 2$. $\square$

7.1 Medidas não atômicas

Seja $\mathcal{M}(\mathbb{RP}^2)$ o espaço de probabilidade das medidas de Borel em $\mathbb{RP}^2$. Uma medida $\nu \in \mathcal{M}(\mathbb{RP}^2)$ é chamada de *não atômica* se $\nu(\{\overline{x}\}) = 0$ para todo $\overline{x} \in \mathbb{RP}^2$.

Lema 19. *Seja A uma matriz em $M(2, \mathbb{R})$ diferente da matriz nula (não necessariamente invertível) e ν uma medida não atômica em $\mathcal{M}(\mathbb{RP}^2)$ então a equação*

$$\int f(\overline{x})\, d(A_*\nu)(\overline{x}) = \int f(A \cdot \overline{x})\, d\nu(\overline{x})$$

é valida para toda função Borel limitada em $\mathbb{RP}^2$ e define uma medida de probabilidade $A_*\nu$ em $\mathbb{RP}^2$.

Demonstração. Para todo x em $\mathbb{R}^2$ tal que $Ax \neq 0$, então $A \cdot \overline{x} = \overline{Ax}$ é bem definida. Como $A \neq 0$, existe no máximo uma direção $\overline{u}$ em $\mathbb{RP}^2$ tal que para todo vetor nessa direção $Au = 0$. Logo $A \cdot \overline{u}$ não está bem definido em $\mathbb{RP}^2$. Assim $A \cdot \overline{x}$ está definido para todo $\overline{x}$ em $\mathbb{RP}^2$ exceto $\overline{u}$. Mas a medida ν é não atômica, isto é, $\nu(\{\overline{u}\}) = 0$, portanto, $A_*\nu$ pode ser definido. $\square$

Lema 20. *Se ν é uma medida de probabilidade em $\mathcal{M}(\mathbb{RP}^2)$ não atômica e A_n uma sequência de matrizes não nulas que convergem para $A \neq 0$, então $(A_n)_*\nu$ converge na topologia fraca* para $A_*\nu$.*

Demonstração. Pelo Lema 19 temos que $A_n \cdot \overline{x}$ está bem definido para todo $\overline{x}$ em $\mathbb{RP}^2$ exceto a direção $\overline{u}_n$ com $\nu(\{\overline{u}_n\}) = 0$ para todo $n \geqslant 1$. Então $A_n \cdot \overline{x}$ está bem definido para todo $\overline{x}$ em $\mathbb{RP}^2$ menos um conjunto enumerável $D = \{\overline{u}_n \in \mathbb{RP}^2, n \geqslant 1\}$ tal que $\nu(D) = 0$, assim temos que

$$\begin{aligned}
\lim_n \int f(\overline{x})\, d((A_n)_*\nu)(\overline{x}) &= \lim_n \int f(A_n \cdot \overline{x})\, d\nu(\overline{x}) && (7.1)\\
&= \int f(A \cdot \overline{x})\, d\nu(\overline{x}) && (7.2)\\
&= \int f(\overline{x})\, d(A_*\nu)(\overline{x}). && (7.3)
\end{aligned}$$

$(A_n)_*\nu$ converge na topologia fraca* para $A_*\nu$. $\square$

Lema 21. *Se ν é uma medida de probabilidade em $\mathcal{M}(\mathbb{RP}^2)$ não atômica, então $H_\nu = \{M \in SL(2,\mathbb{R}); M_*\nu = \nu\}$ é um subgrupo compacto de $SL(2,\mathbb{R})$.*

Demonstração. Sejam A e B em H_ν e E subconjunto mensurável de $\mathbb{RP}^2$, logo

$$(AB)_*\nu(E) = \nu(B^{-1}(A^{-1}(E))) = B_*\nu(A^{-1}(E)) = \nu(A^{-1}(E)) = A_*\nu(E) = \nu(E)$$

além disso

$$A_*^{-1}\nu(E) = \nu(A(E)) = \nu(E)$$

isto é, dado A em H_ν então A^{-1} está em H_ν. Portanto H_ν é subgrupo de $SL(2,\mathbb{R})$. Agora suponha que H_ν não é compacto, neste caso existe uma sequência de matrizes A_n em H_ν tal que $\|A_n\| \to \infty$. Vamos examinar um ponto de acumulação A da sequência renormalizada $\|A_n\|^{-1}A_n$. A matriz A não pode ser invertível já que

$$\det(A) = \lim_n \det\left(\frac{A_n}{\|A_n\|}\right) = 0.$$

No entanto, A tem norma 1, então não é a matriz nula, logo pelo Lema 20 temos que $A_*\nu = \nu$. Assim, A tem posto um e então a medida $\nu = A_*\nu$ deve ser uma medida de Dirac, o que é uma contradição pois a medida de Dirac é atômica. $\square$

Lema 22. *Se μ satisfaz as hipóteses do teorema de Furstenberg, então toda ν em $\mathcal{M}(\mathbb{RP}^2)$ μ-invariante é não atômica.*

Demonstração. Considere ν em $\mathcal{M}(\mathbb{RP}^2)$ μ-invariante. Para todo $\kappa > 0$ temos que o conjunto $\{\overline{x} \in \mathbb{RP}^2\,;\, \nu(\{\overline{x}\}) > \kappa\}$ é finito. Suponha que ν é atômica então existe $\beta > 0$ tal que $\nu(\{\overline{x}\}) \leqslant \beta$ para todo $\overline{x}$ em $\mathbb{RP}^2$. Se consideramos o conjunto

$$\mathcal{L} = \{\overline{x} \in \mathbb{RP}^2\,;\, \nu(\{\overline{x}\}) = \beta\}$$

temos que para $\overline{x}_0$ em $\mathcal{L}$

$$\begin{aligned}
\beta &= \nu(\{\overline{x}_0\}) \\
&= (\mu * \nu)(\{\overline{x}_0\}) \\
&= \iint \chi_{\{\overline{x}_0\}}(M \cdot \overline{x})\, d\mu(M)\, d\nu(\overline{x}) \\
&= \int \nu(\{M^{-1} \cdot \overline{x}_0\})\, d\mu(M).
\end{aligned}$$

Mas $\nu(\{M^{-1} \cdot \overline{x}_0\}) \leqslant \beta$ para todo M, assim $\nu(\{M^{-1} \cdot \overline{x}_0\}) = \beta$ então $M^{-1} \cdot \overline{x}_0 \in \mathcal{L}$ para μ-quase todo M. Isto é $M^{-1}(\mathcal{L}) \subset \mathcal{L}$ para μ-quase todo M. Por outro lado, seja $\overline{x}_0$ em $\mathcal{L}$, então $\overline{x}_0 \in M^{-1}(\mathcal{L})$. De fato

$$\begin{aligned}
\nu(\{M \cdot \overline{x}_0\}) &= (\mu * \nu)(\{M \cdot \overline{x}_0\}) \\
&= \iint \chi_{\{M \cdot \overline{x}_0\}}(M \cdot \overline{x})\, d\mu(M)\, d\nu(\overline{x}) \\
&= \int \nu(\{M^{-1} M \cdot \overline{x}\})\, d\mu(M) \\
&= \int \nu(\{\overline{x}_0\})\, d\mu(M) \\
&= \beta
\end{aligned}$$

Portanto temos que $M(\mathcal{L}) = \mathcal{L}$ para μ-quase todo M $\square$

A partir de agora vamos supor que μ satisfaz as hipóteses do Teorema de Furstenberg, e que ν é uma medida não atômica μ-invariante em $\mathbb{RP}^2$. Nosso primeiro objetivo é expressar o expoente de Lyapunov em termos das duas medidas mencionadas.

Lema 23.

$$\lambda = \iint \log \frac{\|Mx\|}{\|x\|}\, d\mu(M)\, d\nu(x). \tag{7.4}$$

Demonstração. Considere o deslocamento à esquerda

$$\sigma\colon SL(2,\mathbb{R})^{\mathbb{N}} \longrightarrow SL(2,\mathbb{R})^{\mathbb{N}}, \qquad (Y_n(\omega))_{n\in\mathbb{N}} \mapsto (Y_{n+1}(\omega))_{n\in\mathbb{N}}$$

cconhecido como o deslocamento de Bernoulli, no espaço de sequências tem a medida invariante ergódica $\mu^{\mathbb{N}}$. Considere a aplicação

$$\begin{aligned}
T\colon SL(2,\mathbb{R})^{\mathbb{N}} \times \mathbb{RP}^2 &\longrightarrow SL(2,\mathbb{R})^{\mathbb{N}} \times \mathbb{RP}^2 \\
(\omega, x) &\longmapsto (\sigma(\omega), Y_1(\omega)x)
\end{aligned}$$

que deixa invariante a medida $\mu^{\mathbb{N}} \times \nu$.
Considere $f : SL(2,\mathbb{R})^{\mathbb{N}} \times \mathbb{RP}^2 \to \mathbb{R}$ como sendo

$$f(\omega, x) = \log \frac{\|Y_1(\omega)x\|}{\|x\|}. \tag{7.5}$$

Então

$$\frac{1}{n}\sum_{j=0}^{n-1}(f\circ T^j)(\omega,x) = \frac{1}{n}\log\frac{\|Y_n(\omega)Y_{n-1}(\omega)\dots Y_1(\omega)x\|}{\|x\|}.$$

O lado direito converge para λ para quase todos os ω e quase todos os x pelo teorema de Oseledets, isto é, para $\mu^{\mathbb{N}}$-quase todo ω

$$\lim_n \frac{1}{n}\log\frac{\|Y_n(\omega)Y_{n-1}(\omega)\dots Y_1(\omega)x\|}{\|x\|} = \lambda \quad \text{para todo } x \in \mathbb{RP}^2 \text{ a menos de } E^s_\omega. \tag{7.6}$$

Por outro lado, pelo Teorema Ergódico de Birkhoff, temos que para $\mu^{\mathbb{N}} \times \nu$-quase todo ponto

$$\iint \lim_n \frac{1}{n}\log\frac{\|Y_n(\omega)Y_{n-1}(\omega)\dots Y_1(\omega)x\|}{\|x\|}\, d(\mu^{\mathbb{N}}\times\nu) = \iint f\, d(\mu^{\mathbb{N}}\times\nu)$$

logo

$$\begin{aligned}
\iint f(\omega,x)\, d\mu^{\mathbb{N}}(\omega)\, d\nu(x) &= \iint \lim_n \frac{1}{n}\log\frac{\|Y_n(\omega)\dots Y_1(\omega)x\|}{\|x\|}\, d\mu^{\mathbb{N}}(\omega)\, d\nu(x)\\
&= \int_{\{(\omega,x):x\in E^s_\omega\}} \lim_n \frac{1}{n}\log\frac{\|Y_n(\omega)\dots Y_1(\omega)x\|}{\|x\|}\, d\mu^{\mathbb{N}}(\omega)\, d\nu(x)\\
&\quad + \int_{\{(\omega,x):x\notin E^s_\omega\}} \lim_n \frac{1}{n}\log\frac{\|Y_n(\omega)\dots Y_1(\omega)x\|}{\|x\|}\, d\mu^{\mathbb{N}}(\omega)\, d\nu(x).
\end{aligned}$$

Na última igualdade, lembre-se que ν não dá massa para $x \in E^s_\omega$, isto é

$$\iint f(\omega,x)\, d\mu^{\mathbb{N}}(\omega)\, d\nu(x) = \int_{\{(\omega,x):x\notin E^s_\omega\}} \lim_n \frac{1}{n}\log\frac{\|Y_n(\omega)\dots Y_1(\omega)x\|}{\|x\|}\, d\mu^{\mathbb{N}}(\omega)\, d\nu(x). \tag{7.7}$$

Portanto de (7.6) e (7.5) temos que

$$\lambda = \iint \log\frac{\|Mx\|}{\|x\|}\, d\mu(M)d\nu(x).$$

□

7.2 Convergencia do Push-forward de medidas

Nesta seção vamos considerar $S_n(\omega) = Y_1(\omega)Y_2(\omega)\cdots Y_n(\omega)$, e também utilizaremos a seguinte notação $A_*\nu = A\nu$.

Lema 24. *Para $\mu^{\mathbb{N}}$-quase todo ω, existe ν_ω em $\mathcal{M}(\mathbb{RP}^2)$ tal que*

$$S_n(\omega)\nu \to \nu_\omega.$$

Demonstração. Fixe $f : \mathbb{RP}^2 \to \mathbb{R}$ mensurável e limitado. Associe a f a função $F : SL(2,\mathbb{R}) \to \mathbb{R}$ definida como

$$F(M) = \int f(M \cdot \overline{x}) d\nu(\overline{x})$$

Seja $\mathcal{F}_n$ a σ-álgebra gerada por $Y_1, Y_2, \cdots, Y_n$ então $S_n(\cdot)$ é $\mathcal{F}_n$-mensurável e como ν é uma distribuição μ-invariante temos que

$$\begin{aligned} \mathbb{E}(F(S_{n+1}|\mathcal{F}_n) &= \int F(S_n M) d\mu(M) \\ &= \iint f(S_n M \cdot \overline{x}) d\mu(M) d\nu(\overline{x}) \\ &= \int f(S_n \cdot \overline{y}) d\nu(\overline{y}) \\ &= F(S_n). \end{aligned}$$

Além disso, como $\mathbb{E}(|F(S_n|)$ é limitado pois f é limitado, então $\{F(S_n); n \geqslant 1\}$ é um martingale, assim pelo teorema da convergência de martingale, $F(Sn)$ converge quase sempre para algum Γf, isto é

$$\Gamma f(\omega) = \lim_{n\to\infty} F(S_n(\omega)) \quad \text{Para quase todo } \omega.$$

Seja $\{f_k; k \in \mathbb{N}\}$ um subconjunto de funções em $\mathcal{C}(\mathbb{RP}^2, \mathbb{R})$ enumerável e denso. Então, qualquer ponto limite ν_ω da sequência de medidas $S_n(\omega)\nu$ satisfaz que para todo k.

$$\begin{aligned} \int f_k d\nu_\omega &= \lim_{i\to\infty} \int f_k d(S_{n_i}\nu) \\ &= \lim_{i\to\infty} \int f_k \circ S_{n_i} d\nu \\ &= \Gamma f_k(\omega). \end{aligned}$$

Assim, pela unicidade do ponto limite ν_ω a convergência segue. □

Lema 25. *Para quase todo ponto* ω *as medidas* ν_ω *do Lema anterior satisfaz que* $S_n(\omega)M\nu$ *converge a* ν_ω *para* μ*-quase todo* M.

Demonstração. Fixe $f : \mathbb{RP}^2 \to \mathbb{R}$ mensurável e limitado então demonstraremos que

$$\lim_{n\to\infty} \mathbb{E}(F(S_n(\omega))) = \mathbb{E}(\Gamma f(\omega)) = \lim_{n\to\infty} \mathbb{E}(F(S_n(\omega))) \tag{7.8}$$

para μ-quase todo M em $SL(2,\mathbb{R})$.
Para isso, é suficiente provar que

$$\lim_{n\to\infty} \mathbb{E}((F(S_{n+1}) - F(S_n))^2) = 0. \tag{7.9}$$

Assim, então teríamos o seguinte

$$\begin{aligned}\mathbb{E}((F(S_{n+1})-F(S_n))^2) &= \mathbb{E}\left(\left(\iint f(S_nM\cdot\overline{x})d\nu(\overline{x})d\mu(M)-\int f(S_n\cdot\overline{x})d\nu(\overline{x})\right)^2\right)\\ &= \mathbb{E}\left(\left(\iint f(S_nM\cdot\overline{x})d\nu(\overline{x})d\mu(M)-\iint f(S_n\cdot\overline{x})d\nu(\overline{x})d\mu(M)\right)^2\right)\\ &= \mathbb{E}\left(\left(\iint (f(S_nM\cdot\overline{x})-f(S_n\cdot\overline{x}))d\nu(\overline{x})d\mu(M)\right)^2\right). \qquad (7.10)\end{aligned}$$

Se tomamos o limite em ambos lados de (7.10) temos o seguinte

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}((F(S_{n+1})-F(S_n))^2) &= \lim_{n\to\infty}\mathbb{E}\left(\left(\iint (f(S_nM\cdot\overline{x})-f(S_n\cdot\overline{x}))d\nu(\overline{x})d\mu(M)\right)^2\right)\\ 0 &= \lim_{n\to\infty}\mathbb{E}\left(\left(\iint (f(S_nM\cdot\overline{x})-f(S_n\cdot\overline{x}))d\nu(\overline{x})d\mu(M)\right)^2\right).\end{aligned}$$

Assim temos que para quase todo ω

$$\lim_{n\to\infty}\iint (f(S_nM\cdot\overline{x})-f(S_n\cdot\overline{x}))d\nu(\overline{x})d\mu(M)=0.$$

Então a igualdade em 7.8 segue. Resta apenas mostrar que 7.9 é valido. De fato, observe que

$$\begin{aligned}\mathbb{E}(F(S_{n+1})F(S_n)) &= \mathbb{E}\left(\int f(S_{n+1})d\nu\cdot\int f(S_n)d\nu\right)\\ &= \mathbb{E}\left(\iint f(S_nM\cdot\overline{x})d\nu(\overline{x})d\mu(M)\cdot\int f(S_n\cdot\overline{x})d\nu(\overline{x})\right)\\ &= \mathbb{E}\left(\int f(S_n\cdot\overline{x})d\nu(\overline{x})\cdot\int f(S_n\cdot\overline{x})d\nu(\overline{x})\right)\\ &= \mathbb{E}\left(\left(\int f(S_n\cdot\overline{x})d\nu(\overline{x})\right)^2\right)\\ &= \mathbb{E}(F(S_n)^2).\end{aligned}$$

Assim

$$\begin{aligned}\mathbb{E}((F(S_{n+1})-F(S_n))^2) &= \mathbb{E}(F(S_{n+1})^2)-2\mathbb{E}(F(S_{n+1})F(S_n))+\mathbb{E}(F(S_n)^2)\\ &= \mathbb{E}(F(S_{n+1})^2)-\mathbb{E}(F(S_n)^2).\end{aligned}$$

Se aplicamos o limite em ambos lados, temos que

$$\lim_{n\to\infty}\mathbb{E}((F(S_{n+1})-F(S_n))^2)=\mathbb{E}(\Gamma f)^2)-\mathbb{E}(\Gamma f)^2)=0.$$

□

Lema 26. *Para $\mu^{\mathbb{N}}$-quase todo ω existe $Z(\omega)$ em $\mathbb{RP}^2$ tal que $\nu_\omega=\delta_{Z(\omega)}$.*

Demonstração. Considere ω fixo, então para μ-quase todo M temos que

$$\lim_n S_n \nu = \lim_n S_n M \nu.$$

Seja A um ponto limite da sequência normalizada de matrizes $\|S_n\|^{-1} S_n$, logo para uma subsequência temos que

$$\lim_k \|S_{n_k}\|^{-1} S_{n_k} = A.$$

Como A é diferente da matriz nula, já que $\|A\| = 1$, pelo Lema 20 temos que

$$\lim_k \frac{S_{n_k}}{\|S_{n_k}\|} \nu = A\nu \quad \text{e} \quad \lim_k \frac{S_{n_k}}{\|S_{n_k}\|} M\nu = AM\nu.$$

Afirmamos que

$$A\nu = AM\nu \quad \mu\text{-quase todo } M.$$

De fato, seja f em $\mathcal{C}(\mathbb{RP}^2, \mathbb{R})$, então

$$\begin{aligned}
\int f(\overline{x}) d(A\nu) &= \lim \int f(\overline{x}) d(\|S_n\|^{-1} S_n \nu) \\
&= \lim \int f(\|S_n\|^{-1} S_n \cdot \overline{x}) d\nu \\
&= \lim \int f(S_n \cdot \overline{x}) d\nu \\
&= \lim \int f(\overline{x}) d(S_n \nu) \\
&= \lim \int f(\overline{x}) d(S_n M \nu) \\
&= \lim \int f(S_n M \cdot \overline{x}) d\nu \\
&= \lim \int f(\|S_n\|^{-1} S_n M \cdot \overline{x}) d\nu \\
&= \lim \int f(\overline{x}) d(\|S_n\|^{-1} S_n M \nu) \\
&= \int f(\overline{x}) d(AM\nu).
\end{aligned}$$

Agora se A é invertível, então $\nu = M\nu$ μ-quase todo M, logo M pertence ao grupo compacto H_ν, o que é uma contradição à hipótese (i) do Teorema de Furstenberg. Assim A não é invertível, isto é, o posto de A é unidimensional. Seja $Z(\omega)$ a direção do posto então $A\nu = \delta_{Z(\omega)}$ e como $A\nu = \nu_\omega$ segue que $\nu_\omega = \delta_{Z(\omega)}$. □

O próximo passo é mostrar que a convergência a uma medida de Dirac implica crescimento da norma.

Lema 27. *Seja η uma medida de probabilidade em $\mathcal{M}(\mathbb{RP}^2)$ não atômica e (A_n) uma sequência em $\mathrm{SL}(2,\mathbb{R})$ de modo que $A_n\eta$ converge fraco para $\delta_{\overline{z}}$, para algum $\overline{z}$ em $\mathbb{RP}^2$. Então*

$$\lim_n \|A_n\| = \lim_n \|A_n^*\| = \infty.$$

Além disso, para cada w em $\mathbb{R}^2$ e se z é um vetor unitário na direção de $\overline{z}$, então

$$\lim_n \frac{\|A_n^* w\|}{\|A_n^*\|} = |\langle w, z\rangle|.$$

Demonstração. Suponha que $\|A_n\|^{-1}A_n$ converge para alguma matriz A. Como A é diferente da matriz nula, já que $\|A\| = 1$, assim, pelo Lema 20 temos que $\|A_n\|^{-1}A_n\eta$ converge fraco a $A\eta$. Pela hipótese temos que $A\eta = \delta_z$, então temos que $\det(A) = 0$, caso contrário, temos que $\eta = A^{-1}\delta_z = \delta_{A^{-1}z}$, assim η seria atômica. Portanto

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

Para a segunda afirmação, lembre que $A\eta = \delta_z$, então o posto de A é uma linha na direção z. Se z é um vetor unitário na direção de $\overline{z}$ e se considerarmos $\{e_1, e_2\}$ a base canônica de $\mathbb{R}^2$ então podemos escrever $Ae_1 = \pm\|Ae_1\|z$ e $Ae_2 = \pm\|Ae_2\|z$. Para qualquer w em $\mathbb{R}^2$. Temos que

$$\begin{aligned}\|A^*w\|^2 &= |\langle A^*w, e_1\rangle|^2 + |\langle A^*w, e_2\rangle|^2 \\ &= |\langle w, Ae_1\rangle|^2 + |\langle w, Ae_2\rangle|^2 \\ &= (\|Ae_1\|^2 + \|Ae_2\|^2)|\langle w, z\rangle|^2\end{aligned}$$

Observe que, em $w = z$ e como $\|A^*\| = \|A\| = 1$ obtemos $\|Ae_1\|^2 + \|Ae_2\|^2 = 1$ assim temos que para todo w em $\mathbb{R}^2$.

$$\|A^*w\| = |\langle w, z\rangle|.$$

Por outro lado, temos $\lim_n \|A_n\|^{-1}A_n = A$ implica $\lim_n \|A_n^*\|^{-1}A_n^* w = A^*w$ e assim

$$\lim_n \frac{\|A_n^* w\|}{\|A_n^*\|} = \|A^*w\| = |\langle w, z\rangle|.$$

□

Lema 28. *Seja $(X, \mathscr{A}, \lambda)$ um espaço de probabilidade e $T : X \to X$ uma transformação mensurável que preserva a medida λ. Se $f \in L^1(\lambda)$ e também*

$$\lim_n \sum_{j=0}^{n-1} f(T^j(x)) = \infty \quad \lambda\text{-quase todo } x \text{ em} X.$$

Então

$$\int f \, d\lambda > 0.$$

Demonstração. Pelo Teorema Ergódico de Birkhoff, existe para λ-quase todo x em X

$$\tilde{f}(x) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)).$$

Além disso $\tilde{f}$ é invariante pela transformação mensurável T e

$$\int f \, d\lambda = \int \tilde{f} \, d\lambda.$$

Suponha que

$$\int f \, d\lambda = 0,$$

logo

$$0 = \int f \, d\lambda = \int \tilde{f} \, d\lambda.$$

Então $\tilde{f}(x) = 0$ para x em Ω_0 onde $\lambda(\Omega_0) = 1$. Escreva

$$s_n(x) = \sum_{j=0}^{n-1} f(T^j(x)).$$

Para $\epsilon > 0$ considere os seguintes conjuntos

$$U_\epsilon = \{x \in \Omega_0; s_n(x) \geqslant \epsilon \ \text{ para todo } \ n \geqslant 1\},$$

e também

$$V_\epsilon = \bigcup_{k \geqslant 0} T^{-k}(U_\epsilon)$$

Agora, considere x em V_ϵ e denote $k(x) \geqslant 0$ como sendo o menor inteiro de modo que $T^{k(x)}(x)$ esteja em U_ϵ. Denote $k = k(x)$, assim para $n \geqslant k$

$$\begin{aligned} s_n(x) &= s_k(x) + s_{n-k}(T^k(x)) \\ &\geqslant s_k(x) + \sum_{j=k}^{n-1} \epsilon \chi_{U_\epsilon}(T^j(x)) \end{aligned}$$

Dividindo por n e no limite temos que

$$\lim_n \frac{1}{n} s_n(x) \geqslant \lim_n \frac{1}{n} s_k(x) + \lim_n \frac{1}{n} \sum_{j=k}^{n-1} \epsilon \chi_{U_\epsilon}(T^j(x))$$

daí

$$\tilde{f}(x) \geqslant \epsilon \tilde{\chi}_{U_\epsilon}(x)$$

Mas como temos que $\tilde{f}(x) = 0$ para x em Ω_0 então $0 \geqslant \epsilon \tilde{\chi}_{U_\epsilon}(x)$, isso mostra que o conjunto U_ϵ tem medida nula.

$$\begin{aligned} \lambda(U_\epsilon) &= \int_X \tilde{\chi}_{U_\epsilon}(x) d\lambda(x) \\ &= \int_{V_\epsilon} \tilde{\chi}_{U_\epsilon}(x) d\lambda(x) \\ &= 0. \end{aligned}$$

Como T é uma transformação mensurável que preserva a medida λ então para todo $\epsilon > 0$ temos que $\lambda(V_\epsilon) = 0$. De fato, para todo $\epsilon > 0$.

$$\lambda(V_\epsilon) = \lambda\left(\bigcup_{k\geqslant 0} T^{-k}(U_\epsilon)\right) = \sum_{k\geqslant 0} \lambda(T^{-k}(U_\epsilon)) = \sum_{k\geqslant 0} \lambda(U_\epsilon) = 0.$$

Como V_ϵ tem medida nula, daí que

$$\lambda\left(\bigcup_{\epsilon>0}(V_\epsilon)\right) = \sum_{\epsilon>0} \lambda(V_\epsilon) = 0.$$

O que é uma contradição, já que pela hipótese, se $\lim_n s_n(x) = \infty$ então x está em V_ϵ para algum $\epsilon > 0$. Este fato é facilmente visto já que se x não está em V_{1/m^2} para qualquer $m \geqslant 1$, pela definição dos V_{1/m^2} temos que para todo $m \geqslant 1$ e para todo $k \geqslant 0$, $T^k(x)$ não está em U_{1/m^2}. Assim temos que para todo $m \geqslant 1$ e para todo $k \geqslant 0$ existe $n \geqslant 1$ tal que

$$s_n(T^k(x)) \leqslant \frac{1}{m^2}.$$

A partir daí é possível construir uma sequência $(n_m)_m$, onde $n_0 = 0$ e para todo $m \geqslant 1$

$$s_{n_m}(T^{n_1+n_2+\ldots+n_{m-1}}(x)) \leqslant \frac{1}{m^2}.$$

Pela construção e pela aditividade das somas s_n temos que

$$\begin{aligned}
s_{n_1+n_2+\ldots+n_m}(x) &= \sum_{j=0}^{n_1+\ldots+n_m-1} f(T^j(x)) \\
&= s_{n_1}(x) + s_{n_2}(T^{n_1}(x) + s_{n_3}(T^{n_2+n_1}(x)) + \ldots + s_{n_m}(T^{n_{m-1}+n_{m-2}+\ldots+n_1}(x)) \\
&\leqslant 1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{m^2}.
\end{aligned}$$

Da última desigualdade, vemos que $s_{n_1+\ldots+n_m}(x)$ não converge a ∞. $\square$

7.3 Demostração do Teorema de Furstenberg

Substituir em todos os lugares Y_i por Y_i^*. Defina μ^* como sendo a distribuição de Y_i sob a aplicação adjunta, isto e $\mu^*(Y_i) = \mu(Y_i^*)$. Note que μ^* também satisfaz a hipótese do teorema se μ o faz. De fato, suponha que G_{μ^*} satisfaz (ii) no teorema de Furstenberg. Suponha que exista $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_k$ elementos distintos de $\mathbb{RP}^2$ e considere os representantes deles, isto é $v_1, v_2, \ldots, v_k$ de modo que para cada $1 \leqslant j \leqslant k$ e $M \in G_{\mu^*}$ temos que

$$M(v_i) = v_j \quad \text{para algum } \; 1 \leqslant j \leqslant k.$$

Escreva $v_i^\perp$ como sendo ortogonal de v_i, assim temos que

$$M^*(v_j^\perp) = v_i^\perp.$$

Então
$$\{M^*(v_1^\perp), M^*(v_2^\perp), \ldots, M^*(v_k^\perp)\} = \{v_1^\perp, v_2^\perp, \ldots, v_k^\perp\}.$$
Portanto para cada $M \in G_\mu = \{M^* \ : \ M \in G_{\mu^*}\}$, temos que
$$M\left(\bigcup_{i=1}^{k} v_i^\perp\right) = \bigcup_{i=1}^{k} v_i^\perp$$
e assim (ii) não vale para G_μ.

Por outro lado pelo Lema 27 temos que
$$\lim_n \|S_n^*(\omega)x\| = |\langle x, Z(\omega)\rangle| \lim_n \|S_n(\omega)\| = \infty$$
se $x \notin \{Z^\perp(\omega)\}$, assim $\langle x, Z(\omega)\rangle \neq 0$. Daí que
$$\lim_n \log \frac{\|S_n^*(\omega)x\|}{\|x\|} = \infty$$
para $\mu^{\mathbb{N}}$-quase todo ω e para todo $x \in \mathbb{RP}^2 \setminus \{Z^\perp(\omega)\}$. Em particular temos que $\mu^{\mathbb{N}} \times \nu$-quase todo (ω, x), então
$$\lim_n \sum_{j=0}^{n-1} (f \circ T^j)(\omega, x) = \lim_n \log \frac{\|S_n^*(\omega)x\|}{\|x\|} = \infty.$$
Portanto pelo Lema 28 concluímos que $\lambda > 0$. $\square$

REFERÊNCIAS

AVILA, A.; BOCHI, J. Trieste lecture notes on lyapunov exponent. 2008. Citado nas páginas 67 e 79.

BOUGEROL, P. *et al.* **Products of random matrices with applications to Schrödinger operators**. Basileia, Suíça: Springer Science & Business Media, 2012. v. 8. Citado nas páginas 19 e 45.

CHUNG, K. L. **A course in probability theory**. Cambridge, Massachusetts, EUA: Academic press, 2001. Citado na página 32.

FOLLAND, G. B. **Real analysis: modern techniques and their applications**. EUA: John Wiley & Sons, 2013. Citado na página 27.

FURSTENBERG, H. Noncommuting random products. **Transactions of the American Mathematical Society**, JSTOR, v. 108, n. 3, p. 377–428, 1963. Citado nas páginas 7 e 9.

FURSTENBERG, H.; KESTEN, H. Products of random matrices. **The Annals of Mathematical Statistics**, JSTOR, v. 31, n. 2, p. 457–469, 1960. Citado nas páginas 7 e 9.

RUFFINO, P. R. C. **Uma iniciaçao aos sistemas dinâmicos estocásticos**. Rio de Janeiro: SBM, 2012. Citado na página 30.

VIANA, M. **Lectures on Lyapunov Exponents**. Reino Unido: Cambridge University Press, 2014. (Cambridge Studies in Advanced Mathematics). Citado nas páginas 21 e 54.

VIANA, M.; OLIVEIRA, K. **Fundamentos da teoria ergódica**. Rio de Janeiro: SBM, 2014. v. 90. Citado na página 32.

ICMC USP
SÃO CARLOS