UNIVERSIDADE FEDERAL DE ALAGOAS

INSTITUTO DE MATEMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

PEDRO HENRIQUE GOMES DE CARVALHO

FÓRMULA DE ROKHLIN

Maceió-AL

2017

PEDRO HENRIQUE GOMES DE CARVALHO

FÓRMULA DE ROKHLIN

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal de Alagoas, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador: Prof. Dr. Krerley Irraciel Martins de Oliveira

Maceió-AL
2017

Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecária Responsável: Helena Cristina Pimentel do Vale

C331f Carvalho, Pedro Henrique Gomes de.
Fórmula de Rokhlin / Pedro Henrique Gomes de Carvalho. – 2017.
41 f.

Orientador: Krerley Irraciel Martins de Oliveira.
Dissertação (mestrado em Matemática) – Universidade Federal de Alagoas. Instituto de Matemática. Maceió, 2017.

Bibliografia: f. 41.

1. Matemática – Estudo e ensino. 2. Entropia. 3. Jacobiano. 4. Rokhlin, Fórmula de. 5. Sistemas dinâmicos. I. Título.

CDU: 519.218.84

UNIVERSIDADE FEDERAL DE ALAGOAS

PEDRO HENRIQUE GOMES DE CARVALHO

Fórmula de Rokhlin

Dissertação de mestrado na área de concentração em Sistemas Dinâmicos submetida em 21/11/17 à banca examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Matemática da Universidade Federal de Alagoas, como parte dos requisitos necessários à obtenção do grau de mestre em Matemática.

Banca Examinadora

Krerley Oliveira

Dr. Krerley Irraciel Martins Oliveira - Orientador (UFAL)

José Ferreira Alves

Dr. José Ferreira Alves (Universidade do Porto)

Davi dos Santos Lima

Dr. Davi dos Santos Lima (UFAL)

Resumo

Neste trabalho será apresentado a demonstração de um teorema importante de Teoria Ergódica. A Fórmula de Rokhlin afirma que dado uma transformação localmente invertível $f : M \to M$ que admita Jacobiano, se existir uma partição finita ou enumerável que gere a σ-álgebra de M, então pode-se calcular a entropia do sistema dinâmico através do Jacobiano; a entropia é dada por $h_\mu(f) = \int \log J_\mu f \ d\mu$. Depois deste resultado, será dado a definição de Sistema Dinâmico Aleatório e como a Fórmula de Rokhlin pode ser estendida para esta dinâmica.

Palavras-chave: Entropia, Jacobiano, Fórmula de Rokhlin

Abstract

In this work a demonstration of an important theorem of Ergodic Theory is presented. The Rokhlin's formula states that given a locally invertible transformation $f : M \to M$ which Jacobian admits, if there is a finite or enumerable partition that generates the σ-algebra of M, then one can compute an entropy of the dynamic system through the Jacobian; the entropy is given by $h_\mu(f) = \int \log J_\mu f \; d\mu$. After this result, is given a definition of Random Dynamic System and how a Rokhlin's Formula can be extended to this dynamic.

Keywords: Entropy, Jacobian, Rokhlin's Formula

Sumário

Introdução

Começaremos no primeiro capítulo definindo conceitos básicos da Teoria Ergódica. Discutiremos alguns conceitos importantes, como: Esperança Condicional, Desintegração de Rokhlin.

No segundo capítulo definiremos Entropia, será dada uma noção intuitiva do surgimento do conceito de Entropia e também serão apresentados alguns teoremas que nos ajudarão no cálculo da entropia.

O capítulo 03 é destinado ao tema principal deste trabalho. A noção de Jacobiano é claramente uma generalização da ideia de Jacobiano vista em cálculo, o Jacobiano de um difeomorfismo com respeito ao volume de $\mathbb{R}^n$. Neste capítulo iremos definir o que vem a ser o Jacobiano de uma função localmente invertível relativamente a uma medida η. Mais adiante, mostraremos o quão importante é a existência de Jacobiano para facilitar o cálculo da entropia. Por fim, será demonstrado a Fórmula de Rokhlin, um teorema que relaciona o jacobiano de uma função com a sua entropia.

No capítulo 04 faremos uma aplicação da fórmula de Rokhlin. Apresentaremos a fórmula de Rokhlin para o caso de Sistemas Dinâmicos Aleatórios.

1 Esperança condicional

Sabemos no caso de subconjuntos convexos de espaços vetoriais com dimensão finita tem-se que todo elemento do convexo pode ser escrito como combinação convexa dos elementos extremais. Por exemplo, todo ponto de um triângulo pode ser escrito como combinação convexa dos vértices do triângulo. Tratando do conjunto das medidas invariantes por uma transformação, as medidas ergódicas são pontos extremais. Agora, uma pergunta interessante em Teoria Ergódica é saber se toda medida invariante é uma combinação linear de medidas ergódicas. O teorema da decomposição ergódica responde essa pergunta. No entanto, estamos interessados nos conceitos que surgem na preparação desse teorema. Um conceito importante é o de esperança condicional; este conceito será fundamental para a demonstração da Fórmula de Rokhlin, bem como o conceito da desintegração de uma medida. Abordaremos esses conceitos neste capítulo.

1.1 Algumas definições essenciais

Nesta seção $(M, \mathcal{B}, \mu)$ será um espaço de probabilidade.

Definição 1.1.1 *Dizemos que uma coleção $\mathcal{P} = (P_n)_n$ $n \in \mathbb{N}$, de conjuntos mensuráveis de M é uma **partição** se $\mu(P_i \cap P_j) = 0$ para $i \neq j$ e $\mu(X \setminus \cup_i P_i) = 0$.*

Exemplo 1.1.1 *Seja $M = \mathbb{R}^2$. Considere $\mathcal{P} = (P_j)_j$ onde $P_j = \{j\} \times \mathbb{R}$, com $j \in \mathbb{R}$. Então $\mathcal{P}$ é uma partição de $\mathbb{R}^2$.*

Exemplo 1.1.2 *Seja $(\mathbb{R}, \mathcal{B}, m)$ um espaço de medida, onde $\mathcal{B}$ é a σ-álgebra de Borel e m é a medida de Lebesgue. Então $\mathcal{P} = (P_\lambda)_\lambda$ onde $P_\lambda = \{\lambda\} \in \mathbb{R}$ é uma partição de $\mathbb{R}$ em conjuntos mensuráveis. A chamada **partição em pontos**.*

Definição 1.1.2 *Sejam $\mathcal{P}$ e $\mathcal{Q}$ duas partições, dizemos que $\mathcal{P}$ é **menos fina** que $\mathcal{Q}$, se todo elemento de $\mathcal{Q}$ está contido em algum elemento de $\mathcal{P}$, a menos de medida nula, denotamos por $\mathcal{P} \prec \mathcal{Q}$.*

Denotamos por $\mathcal{P}(x)$ o elemento da partição que contém um ponto x. A **soma** $\mathcal{P} \vee \mathcal{Q}$ de duas partições $\mathcal{P}$ e $\mathcal{Q}$ é a partição cujos elementos são as interseções $P \cap Q$ com $P \in \mathcal{P}$ e $Q \in \mathcal{Q}$. Mais geralmente, dada qualquer família enumerável de partições $\mathcal{P}_n$, definimos

$$\bigvee_n \mathcal{P}_n = \{\bigcap_n P_n : P_n \in \mathcal{P}_n \text{ para cada } n\}.$$

A soma $\mathcal{P} \vee \mathcal{Q}$ é precisamente a menos fina de todas as partições $\mathcal{R}$ tais que $\mathcal{P} \prec \mathcal{R}$ e $\mathcal{Q} \prec \mathcal{R}$.

Neste trabalho, $\mathcal{P}$ será uma partição de M em conjuntos mensuráveis. Denotaremos por $\pi : M \to \mathcal{P}$ a projeção natural que associa a cada $x \in M$ o elemento $\mathcal{P}(x)$ da partição que contém x.

Este mapa π possibilita munir a partição $\mathcal{P}$ com uma estrutura de espaço de probabilidade, da seguinte forma: Um subconjunto $\mathcal{Q}$ de $\mathcal{P}$ é mensurável se, e somente se, a pré imagem

$$\pi^{-1}(\mathcal{Q}) = \text{união dos elementos } P \in \mathcal{P} \text{ que pertencem a } \mathcal{Q}$$

é um subconjunto mensurável de M.

Proposição 1.1.1 *A família $\hat{\mathcal{B}}$ dos subconjuntos mensuráveis é uma σ-álgebra em $\mathcal{P}$.*

Demonstração.

- $\{\emptyset\}$ e $\{\mathcal{P}\} \in \hat{\mathcal{B}}$. De fato, $\emptyset = \pi^{-1}(\emptyset)$ e $M = \pi^{-1}(\mathcal{P})$ que é mensurável;

- Se $P \in \hat{\mathcal{B}}$, então $\pi^{-1}(P)$ é um conjunto mensurável em M, logo $(\pi^{-1}(P))^c \in M$ é mensurável. Como $(\pi^{-1}(P))^c = \pi^{-1}(P^c)$, concluímos que $P^c \in \hat{\mathcal{B}}$;

- Se $P_j \in \hat{\mathcal{B}}$ é uma sequência enumerável de subconjuntos mensuráveis, então $\pi^{-1}(P_j) \in M$ é mensurável para todo j. Logo $\cup_j \pi^{-1}(P_j)$ é um subconjunto mensurável de M, portanto $\pi^{-1}(\cup_j(P_j))$ é mensurável, o que nos mostra que $\cup_j(P_j) \in \hat{\mathcal{B}}$. Logo, $\hat{\mathcal{B}}$ é uma σ-álgebra em $\mathcal{P}$.

■

Com isso, definimos a **medida quociente** $\hat{\mu}$ por

$$\hat{\mu}(\mathcal{Q}) = \mu(\pi^{-1}(\mathcal{Q})) \text{ para cada } \mathcal{Q} \in \hat{\mathcal{B}}.$$

Definição 1.1.3 *Uma* ***desintegração*** *de μ relativamente a uma partição $\mathcal{P}$ é uma família $\{\mu_P : P \in \mathcal{P}\}$ de probabilidades em M tal que, para todo conjunto mensurável $E \subset M$:*

i. $\mu_P(P) = 1$ *para $\hat{\mu}$-quase todo $P \in \mathcal{P}$;*

ii. *a aplicação $\mathcal{P} \to \mathbb{R}$, definida por $P \to \mu_P(E)$ é mensurável;*

iii. $\mu(E) = \int \mu_P(E) d\hat{\mu}(P)$.

As μ_P são chamadas ***probabilidades condicionais*** *de μ relativamente a $\mathcal{P}$.*

Exemplo 1.1.3 *Seja $\mathcal{P} = \{P_1, P_2, P_3, ..., P_n\}$ uma partição finita de M em subconjuntos mensuráveis com $\mu(P_i) > 0$ para todo i. A medida quociente $\hat{\mu}$ é dada por $\hat{\mu}(\{P_i\}) = \mu(P_i)$. Considere a restrição normalizada μ_i de μ a cada P_i*

$$\mu_i(E) = \frac{\mu(E \cap P_i)}{\mu(P_i)} \text{ para todo } E \subset M \text{ mensurável.}$$

Então $\{\mu_1, \mu_2, ..., \mu_n\}$ é uma desintegração de μ relativa a $\mathcal{P}$, já que

$$\mu(E) = \sum_{i=1}^{n} \mu(\{P_i\})\mu(E).$$

Definição 1.1.4 *Dizemos que $\mathcal{P}$ é uma* ***partição mensurável*** *se, restrita a algum subconjunto de M com medida total, ela é o limite de uma sequência crescente de partições enumeráveis. Mais precisamente, a partição é mensurável se existe algum conjunto mensurável $M_0 \subset M$ com medida total tal que, restrito a M_0*

$$\mathcal{P} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$$

para alguma sequência crescente $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \mathcal{P}_n \prec \cdots$ de partições enumeráveis. Lembre que $\mathcal{P}_i \prec \mathcal{P}_{i+1}$ significa que todo elemento de $\mathcal{P}_{i+1}$ está contido em algum elemento de $\mathcal{P}_i$.

A definição acima é equivalente a seguinte definição:

Definição 1.1.5 *Dizemos que $\mathcal{P}$ é uma* ***partição mensurável*** *se existir um conjunto mensurável $M_0 \subseteq M$, com medida total, e uma família enumerável $\{A_n; n \in \mathbb{N}\}$ de conjuntos mensuráveis tais que, para cada $P \in \mathcal{P}$, existe uma sequência $B_n \in \{A_n, M\backslash A_n\}$ com*

$$P \cap M_0 = \bigcap_{n\in\mathbb{N}} (B_n \cap M_0).$$

Exemplo 1.1.4 *Seja $M = (0,1) \times (0,1)$, considere a partição $\mathcal{P} = (0,1) \times \{c\}$; com $c \in (0,1)\}$ de M. $\mathcal{P}$ é mensurável, de fato, seja $\{a_i; i \in \mathbb{N}\}$ uma enumeração do conjunto $\mathbb{Q}\cap(0,1)$ e defina a família de conjuntos mensuráveis*

$$A^j_{a_i} = (0,1) \times [(a_i - \frac{1}{2^j}, a_i + \frac{1}{2^j}) \cap (0,1)]$$

Assim, temos que $\{A^j_{a_i}; i,j \in \mathbb{N}\}$ é uma família enumerável de conjuntos mensuráveis de M tal que, para todo $P = (0,1) \times \{c\} \in \mathcal{P}$,

$$P = \bigcap_{i,j\in\mathbb{N}} B_{i,j} \text{ com } B_{i,j} \in \{A^j_{a_i}, (A^j_{a_i})^c\}.$$

Depois de definirmos o que vem a ser uma medida ergódica, daremos um exemplo de uma partição que não é mensurável.

Teorema 1.1.1 *(Desintegração de Rokhlin) Suponha que o espaço métrico M é completo separável e que $\mathcal{P}$ é partição mensurável. Então a probabilidade μ admite alguma desintegração relativamente a $\mathcal{P}$.*

Não demonstraremos o teorema acima, mas ele será bastante útil. A demonstração pode ser encontrada na referência [2], Capítulo 5.

1.2 Esperanças condicionais

Para o restante do capítulo, fixemos uma sequência $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \mathcal{P}_n \prec \cdots$ de partições enumeráveis tais que $\mathcal{P} = \vee_{n=1}^{\infty}\mathcal{P}_n$ restrito a algum conjunto $M_0 \subset M$ com medida total.

Seja $\psi : M \to \mathbb{R}$ uma função mensurável limitada qualquer. Para cada $n \geq 1$, defina $e_n(\psi) : M \to \mathbb{R}$ da seguinte forma:

$$e_n(\psi, x) = \begin{cases} \dfrac{1}{\mu(\mathcal{P}_n(x))} \displaystyle\int_{\mathcal{P}_n(x)} \psi \, d\mu, & \text{se } \mu(\mathcal{P}_n(x)) > 0 \\ 0, & \text{caso contrário} \end{cases}$$

Como as partições $\mathcal{P}_n$ são enumeráveis, o segundo caso da definição se aplica apenas num conjunto de medida zero. Observe que $e_n(\psi)$ é constante em cada $P_n \in \mathcal{P}_n$; denotamos por $E_n(\psi, P_n)$ o valor desta constante. Então

$$\int \psi d\mu = \sum_{P_n} \int_{P_n} \psi d\mu = \sum_{P_n} \mu(P_n) E_n(\psi, P_n) = \int e_n(\psi) d\mu \tag{1.1}$$

para todo $n \in \mathbb{N}$.

Lema 1.2.1 *Seja $\psi : M \to \mathbb{R}$ função mensurável qualquer, então existe um subconjunto M_ψ de M com $\mu(M_\psi) = 1$ tal que*

1. *$e(\psi, x) = \lim_n e_n(\psi, x)$ existe para todo $x \in M_\psi$;*
2. *$e(\psi) : M \to \mathbb{R}$ é mensurável e é constante em cada $P \in \mathcal{P}$;*
3. *$\int \psi d\mu = \int e(\psi) d\mu$.*

Demonstração: Vamos supor que $\psi \geq 0$. Para cada $\alpha < \beta$, seja $S(\alpha, \beta)$ o conjunto dos pontos $x \in M$ tais que

$$\liminf_n e_n(\psi, x) < \alpha < \beta < \limsup_n e_n(\psi, x)$$

É claro que a sequência $e_n(\psi, x)$ diverge se, e somente se, $x \in S(\alpha, \beta)$ (óbvio, pois se $\liminf = \limsup$ a sequência converge) para algum $\alpha < \beta$. Em outras palavras, o limite $e_n(\psi, x)$ existe se, e somente se, x pertence a interseção M_ψ de todos os $S(\alpha, \beta)^c$ com $\alpha < \beta$ racionais. Como se trata de uma interseção enumerável, para provar que $\mu(M_\psi) = 1$ basta mostrar que $\mu(S(\alpha, \beta)) = 0$ para todo $\alpha < \beta$.

Como α, β estão fixos, podemos escrever $S = S(\alpha, \beta)$. Dado $x \in S$, fixe uma sequência de inteiros $1 \leq a_1^x < b_1^x < ... < a_i^x < b_i^x < ...$ tais que

$$e_{a_i^x}(\psi, x) < \alpha \quad \text{e} \quad e_{b_i^x}(\psi, x) > \beta \text{ para todo } i \geq 1$$

Defina A_i como $A_i(x) = \mathcal{P}_{a_i^x}(x)$ e B_i como sendo a união dos elementos $B_i(x) = \mathcal{P}_{b_i^x}(x)$ obtidos deste modo, para todos os pontos $x \in S$. Por construção, $S \subset A_{i+1} \subset B_i \subset A_i$ para todo $i \geq 1$. Em particular, S está contido no conjunto

$$\hat{S} = \bigcap_{i=1}^{\infty} B_i = \bigcap_{i=1}^{\infty} A_i.$$

Como a sequência $\mathcal{P}_n$ é crescente, $n \geq 1$, dados dois quaisquer dos conjuntos $A_i(x) = \mathcal{P}_{a_i^x}(x)$ que formam A_i, ou eles são disjuntos ou um deles está contido no outro. Então os conjuntos $A_i(x)$ maximais são disjuntos dois-a-dois e, portanto, constituem uma partição de A_i. Logo, somando apenas sobre estes conjuntos maximais com medida positiva, temos

$$\int_{A_i} \psi \, d\mu = \sum_{A_i(x)} \int_{A_i(x)} \psi \, d\mu \leq \sum_{A_i(x)} \alpha\mu(A_i(x)) = \alpha\mu(A_i),$$

para qualquer $i \geq 1$. Analogamente,

$$\int_{B_i} \psi \, d\mu = \sum_{B_i(x)} \int_{B_i(x)} \psi \, d\mu \geq \sum_{B_i(x)} \beta\mu(B_i(x)) = \beta\mu(B_i).$$

Como $A_i \supset B_i$ e nós estamos supondo que $\psi \geq 0$, segue que

$$\alpha\mu(A_i) \geq \int_{A_i} \psi \, d\mu \geq \int_{B_i} \psi \, d\mu \geq \beta\mu(B_i),$$

para todo $i \geq 1$. Tomando o limite quando $i \to \infty$, obtemos que $\alpha\mu(\hat{S}) \geq \beta\mu(\hat{S})$. Isto implica que $\mu(\hat{S}) = 0$ e, portanto, $\mu(S) = 0$. Isto prova a afirmação quando ψ é não-negativa. O caso geral segue imediatamente, sabendo que sempre podemos escrever $\psi = \psi^+ - \psi^-$, onde $\psi^\pm$ são mensuráveis, não-negativas e limitadas. Note que $e_n(\psi) = e_n(\psi^+) - e_n(\psi^-)$ para todo $n \geq 1$ e, portanto, a conclusão do lema é verdadeira para ψ se ela vale para ψ^+ e ψ^-. Isto conclui a prova da afirmação (1).

A mensurabilidade de $e(\psi)$ segue do fato de que o limite de funções mensuráveis é mensurável. Dado que $\mathcal{P}_n$ é menos fina que $\mathcal{P}$, é claro que $e_n(\psi)$ é constante em cada $P \in \mathcal{P}$, restrito a um subconjunto de M com medida total. Logo o mesmo vale para $e(\psi)$. Isto demonstra (2). Observe também que $\mid e_n(\psi) \mid \leq \sup \mid \psi \mid$ para todo $n \geq 1$. Assim, usando o teorema da convergência dominada, podemos passar o limite em (1.1). Desta forma obtemos a afirmação (3).

Estaremos interessados no caso que ψ é uma função característica, $\psi = \mathcal{X}_A$ para algum conjunto mensurável $A \subset M$. Ou seja,

$$e(\psi, x) = \lim_n \frac{\mu(\mathcal{P}_n(x) \cap A)}{\mu(\mathcal{P}_n(x))}$$

2 Entropia

2.1 Definições e alguns resultados

Neste capítulo $(M, \mathcal{B}, \mu)$ indicará um espaço de probabilidade e M é um espaço métrico compacto. Por *partição*, sempre entenderemos uma família finita ou enumerável $\mathcal{P}$ de subconjuntos mensuráveis de M disjuntos dois-a-dois e cuja união tem medida total.

Definição 2.1.1 *Seja $f : M \to M$ uma transformação mensurável, dizemos que a medida μ é **invariante** por f, ou que f **preserva** μ se*

$$\mu(E) = \mu(f^{-1}(E)) \text{ para todo conjunto mensurável } E \subset M.$$

Exemplo 2.1.1 *Consideremos a transformação $f : [0, 1] \to [0, 1]$ definida por*

$$f(x) = \begin{cases} 2x, & se\ 0 \leq\ x < \dfrac{1}{2} \\ 2x - 1, & se\ \dfrac{1}{2} \leq x \leq 1. \end{cases}$$

A medida de Lebesgue em [0,1] invariante por f. A demonstração pode ser encontrada na referência [2]. Uma ideia é demonstrar que a medida é preservada em intervalos, depois demonstrar que é preservada numa união de intervalos e daí usar um teorema de extensão para obter que a medida é preservada em toda a σ-álgebra.

Um teorema bastante importante é dado a seguir, ele assegura que dada uma medida invariante finita, quase todo ponto de qualquer conjunto mensurável E regressa a E um número infinito de vezes:

Teorema 2.1.1 *(Recorrência de Poincaré) Seja $f : M \to M$ uma transformação mensurável e seja μ uma medida finita invariante por f. Seja $E \subset M$ qualquer conjunto mensurável com $\mu(E) > 0$. Então, para μ-quase todo ponto $x \in E$ existem infinitos valores de n para os quais $f^n(x)$ também está em E*

Demonstração. Seja E_0 o conjunto dos pontos $x \in E$ que nunca regressam a E. Observe que as pré-imagens $f^{-n}(E_0)$ são disjuntas duas-a-duas. De fato, suponhamos que existem $m > n \geq 1$ tais que $f^{-m}(E_0)$ intersecta $f^{-n}(E_0)$. Seja x um ponto na intersecção e seja $y = f^n(x)$. Então $y \in E_0$ e $f^{m-n}(y) = f^m(x) \in E_0$, que está contido em E. Isto quer dizer que y volta pelo menos uma vez a E, o que contradiz a definição de E. Assim, as pré-imagens são duas-a-duas disjuntas.

Como a medida é invariante por f, $\mu(f^{-n}(E_0)) = \mu(E_0)$ para todo $n \leq 1$, concluímos que

$$\mu(\bigcup_{n=1}^{\infty} f^{-n}(E_0)) = \sum_{n=1}^{\infty} \mu(f^{-n}(E_0)) = \sum_{n=1}^{\infty} \mu(E_0).$$

Como a medida é finita, a expressão do lado esquerda é finita. Por outro lado, do lado direito temos uma soma de infinitos termos, todos iguais, assim o único jeito dessa série não explodir (da soma ser finita) é que as parcelas sejam nulas, isto é, $\mu(E_0) = 0$.

Agora, denotemos por F o conjunto dos pontos $x \in E$ que regressam apenas um número finito de vezes. Por definição de E, temos que todo ponto $x \in F$ tem algum iterado $f^k(x) \in E_0$. Ou seja,

$$F \subset \bigcup_{k=0}^{\infty} f^{-k}(E_0).$$

Agora, como $\mu(E_0) = 0$ e μ é invariante, temos:

$$\mu(F) \leq \mu(\bigcup_{k=0}^{\infty} f^{-k}(E_0)) \leq \sum_{k=0}^{\infty} \mu(f^{-k}(E_0)) = \sum_{k=0}^{\infty} \mu(E_0) = 0.$$

Portanto, $\mu(F) = 0$ e concluímos assim o teorema. ■

Agora, introduziremos um dos conceitos fundamentais para o estudo da Teoria Ergódica. Para motivar o enunciado, considere um conjunto mensurável $E \subset M$ com medida positiva e um ponto $x \in M$ qualquer. Queremos estudar o conjunto dos iterados de x que retornam para E, isto é,

$$\{j \geq 0 : f^j(x) \in E\}.$$

Por exemplo, o teorema anterior (Recorrência de Poincaré) afirma que, para quase todo ponto $x \in E$, este conjunto é infinito. Porém, gostaríamos de uma informação mais precisa, de uma natureza quantitativa. Chamamos de **tempo médio de visita** de x a E o valor de

$$\tau(E,x) = \lim_{n\to\infty} \frac{1}{n}\#\{0 \leq j < n : f^j(x) \in E\}.$$

Seria interessante saber, por exemplo, em que condições este tempo médio de visita é positivo. Antes de abordar este problema, é necessário responder a uma questão ainda mais básica: esse limite existe?

O próximo resultado responde essa pergunta e generaliza.

Teorema 2.1.2 *Seja $f : M \to M$ uma transformação mensurável e seja μ uma probabilidade invariante por f. Dada qualquer função integrável $\varphi : M \to \mathbb{R}$, o limite*

$$\hat{\varphi}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x))$$

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

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

Demonstração: Omitiremos a demonstração deste resultado por não ser o foco principal do trabalho, porém a demonstração pode ser obtida no livro "Fundamentos da Teoria Ergódica"Cap5.

O limite $\hat{\varphi}$ é chamado **média temporal**, ou **média orbital** de φ. Uma outra observação é que esse teorema resolve a pergunta feita anteriormente, basta tomarmos $\varphi = \mathcal{X}_E$. Podemos definir então o que vem a ser uma medida ergódica.

Definição 2.1.2 *Seja μ uma medida de probabilidade invariante por uma transformação mensurável $f : M \to M$. Diremos que o sistema (f, μ) é* ***ergódico*** *se, dado qualquer conjunto mensurável E, temos que $\tau(E, x) = \mu(E)$ para μ-quase todo ponto $x \in M$.*

Dizemos que uma função $\phi : M \to \mathbb{R}$ é **invariante** se $\phi = \phi \circ f$ em μ-quase todo ponto. Ou seja, a menos de um conjunto com medida nula, a função é constante em toda trajetória de f. Além disso, dizemos que um conjunto mensurável $B \subset M$ é **invariante** se a sua função característica $\mathcal{X}_B$ é uma função invariante. Em outras palavras, B é invariante se ele difere da sua pré-imagem $f^{-1}(B)$ por um conjunto de medida nula:

$$\mu(B\Delta f^{-1}(B)) = 0$$

Proposição 2.1.1 *Seja μ uma probabilidade invariante de uma transformação mensurável $f : M \to M$. São equivalentes:*

a) (f, μ) é ergódico.

b) Para todo subconjunto invariante A tem-se $\mu(A) = 0$ ou $\mu(A) = 1$

Agora, daremos exemplo de uma partição que não é mensurável.

Exemplo 2.1.2 *Seja $f : M \to M$ uma transformação invertível e μ a medida de Lebesgue invariante, tal que $\mu(M) = 1$. Considere $\mathcal{P}$ a partição de M em órbitas, onde $\mathcal{O}_x = \{f^n(x) | n \in \mathbb{Z}\}$. Primeiramente, é fácil perceber que as órbitas são subconjuntos invariantes por f. Agora, suponha por absurdo que exista uma sequência crescente de partições enumeráveis $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \mathcal{P}_n \prec \cdots$ tais que*

$$\bigvee_{n=1}^{\infty} \mathcal{P}_n = \mathcal{P}.$$

Seja $P_n \in \mathcal{P}_n$, perceba que P_n é a união de elementos da partição de $\mathcal{P}$, que são órbitas. Logo P_n é também uma conjunto invariante de f. Portanto, $\mu(P_n) = 0$ ou $\mu(P_n) = 1$. Assim, existe um único elemento da partição $\mathcal{P}_n$ com medida total, vamos denotar este elemento por $\hat{P}_n$. Agora, seja $P \in \vee_{n=1}^{\infty} \mathcal{P}_n$, então $P = \cap_{n=1}^{\infty} P_n$, onde $P_n \in \mathcal{P}_n$. Com isso, temos duas possibilidades para P:

$$P = \bigcap_{n=1}^{\infty} P_n, \text{ onde cada } P_n = \hat{P}_n \text{ e portanto } \mu(P) = 1$$

ou então

$$P = \bigcap_{n=1}^{\infty} P_n, \text{ onde existe ao menos um } j \in \mathbb{N} \text{ tal que } P_j \neq \hat{P}_j. \text{ Assim, } \mu(P) = 0.$$

Concluímos então que existe um único $\hat{P} \in \mathcal{P}$ com medida total, absurdo! Pois, como $\mathcal{P}$ é uma partição em órbitas, cada elemento $P \in \mathcal{P}$ é um conjunto enumerável, e portanto $\mu(P) = 0$, já que todo conjunto enumerável tem medida de Lebesgue igual a zero.

2.2 Entropia

Comecemos por discutir, simplesmente por motivação, o que o conceito de ganho de "informação"poderia significar. Um experimento nesse espaço significa que um ponto $x \in M$ é selecionado aleatoriamente. Trataremos os conjuntos da partição $A \in \mathcal{B}$ como os possíveis resultados desse experimento, regido pela probabilidade μ. Isso significa que A ocorre com uma probabilidade $\mu(A)$. Aqui podemos fazer uma pergunta informal acerca de um $A \in \mathcal{B}$ fixo:

Suponha que x seja um ponto escolhido aleatoriamente em M. O quanto de informação ganhamos ao saber que $x \in A$?

Por exemplo, se $A = M$, não ganhamos informação alguma sobre x. Podemos dizer que a quantidade de ganho de informação é zero (de fato, não é surpresa saber que $x \in M$). Se A é apenas um ponto, esse ponto deve ser o próprio x e nós aprendemos a localização exata de x. Então podemos dizer que a quantidade de ganho de informação é máxima. A pergunta anterior torna-se interessante quando consideramos uma partição $\mathcal{B}$ e fazemos a seguinte pergunta:

Suponhamos que x seja selecionado aleatoriamente em M e seremos informados sobre qual conjunto da partição ele se encontra. Quanta informação sobre a localização precisa de x esperamos ganhar?

Por um lado, quando menor for $\mu(A)$, maior será o ganho de informação ao dizer que $x \in A$. Vamos abordar a primeira pergunta. Estamos a procura de uma função $I : \mathcal{B} \to [0, \infty]$ tal que:

$$I(A) := \text{o quanto de informação ganhamos ao saber que } x \in A.$$

Além do mais, tenha as seguintes propriedades:

i. $\mu(A) = \mu(B) \Rightarrow I(A) = I(B)$ (a informação depende apenas da medida do conjunto).

ii. $\mu(A) = 1 \Rightarrow I(A) = 0$ (nenhuma informação é obtida sobre a localização de x)

iii. $\mu(A) = 0 \Rightarrow I(A) = M$ onde $M := \sup I \in [0, \infty]$ (temos o maior ganho de informação sobre x)

iv. $\mu(A) < \mu(B) \Rightarrow I(B) < I(A)$

v. $\mu(A \cap B) = \mu(A)\mu(B)$ (isto é, A e B são independentes) $\Rightarrow I(A \cap B) = I(A) + I(B)$.

Através da propriedade i., podemos imaginar que estamos a procura de uma função $\phi : [0, 1] \to [0, \infty]$ tal que

$$I(A) = \phi(\mu(A))$$

para todo $A \in \mathcal{B}$. Vamos tentar definir uma ϕ universal, no sentido de que não depende do espaço de probabilidade estabelecido. Procuramos então uma $\phi : [0, 1] \to [0, 1]$ que seja estritamente decrescente, com $\phi(1) = 0$ e que assuma o máximo no ponto zero. Além do mais, pela propriedade v., queremos que

$$\phi(ab) = \phi(a) + \phi(b)$$

para todo $a, b \in [0, 1]$. Segue que ϕ é dada por

$$\phi(x) = -\log x.$$

Podemos entar definir formalmente a função informação.

Definição 2.2.1 *Seja $(M, \mathcal{B}, \mu)$ um espaço de probabilidade. Nós dizemos que $I : \mathcal{B} \to [0, \infty]$ dada por*

$$I(A) = -\log(\mu(A))$$

*é a **função informação** em $(M, \mathcal{B}, \mu)$. Nós podemos pensar em $I(A)$ como o valor da informação que se obtém sobre a localização precisa de um ponto x escolhido aleatoriamente em M, se for dito que $x \in A$.*

Definição 2.2.2 *Seja $f : M \to M$ uma transformação preservando a medida de probabilidade μ. Dada uma partição $\mathcal{P}$ de M, a **entropia da partição $\mathcal{P}$ com respeito a μ** é*

$$H_\mu(\mathcal{P}) := \int I(x) d\mu = -\sum_{P \in \mathcal{P}} \mu(P) \log \mu(P)$$

onde $I(x) = -\log(\mu(\mathcal{P}(x)))$. Usualmente, fazemos a convenção de que

$$0 \log 0 = \lim_{x \to 0} x \log x = 0$$

Definição 2.2.3 *Chamamos **entropia condicional** de uma partição $\mathcal{P}$ com relação a uma partição $\mathcal{Q}$ ao número*

$$H_\mu(\mathcal{P}/\mathcal{Q}) = \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} -\mu(P \cap Q) \log(\frac{\mu(P \cap Q)}{\mu(Q)}).$$

Observação: 2.2.1 *É claro que $H_\mu(\mathcal{P}/M) = H_\mu(\mathcal{P})$*

Lema 2.2.1 *Sejam $\mathcal{P}, \mathcal{Q}$ e $\mathcal{R}$ partições com entropia finita. Então*

1. $H_\mu(P \vee \mathcal{Q}/\mathcal{R}) = H_\mu(\mathcal{P}/\mathcal{R}) + H_\mu(\mathcal{Q}/\mathcal{P} \vee \mathcal{R})$;
2. *Se $\mathcal{P} \prec \mathcal{Q}$ então $H_\mu(\mathcal{P}/\mathcal{R}) \leq H_\mu(\mathcal{Q}/\mathcal{R})$ e $H_\mu(\mathcal{R}/\mathcal{P}) \geq H_\mu(\mathcal{R}/\mathcal{Q})$*
3. *$\mathcal{P} \prec \mathcal{Q}$ se, e somente se, $H_\mu(\mathcal{P}/\mathcal{Q}) = 0$*

Demonstração. Por definição,

$$\begin{aligned} H_\mu(\mathcal{P} \vee \mathcal{Q}/\mathcal{R}) &= \sum_{P,Q,R} -\mu(P \cap Q \cap R) \log \frac{\mu(P \cap Q \cap R)}{\mu(R)} \\ &= \sum_{P,Q,R} -\mu(P \cap Q \cap R) \log \frac{\mu(P \cap Q \cap R)}{\mu(P \cap R)} \\ &+ \sum_{P,Q,R} -\mu(P \cap Q \cap R) \log \frac{\mu(P \cap R)}{\mu(R)}. \end{aligned}$$

A soma do lado direito pode ser reescrita como

$$\sum_{S\in\mathcal{P}\vee\mathcal{R},Q\in\mathcal{Q}} -\mu(S\cap Q)\log\frac{\mu(C\cap Q)}{\mu(S)} + \sum_{P\in\mathcal{P},R\in\mathcal{R}} -\mu(P\cap R)\log\frac{\mu(P\cap R)}{\mu(R)}$$
$$\stackrel{*}{=} H_\mu(\mathcal{Q}/\mathcal{P}\vee\mathcal{R}) + H_\mu(\mathcal{P}/\mathcal{R}).$$

Isto demonstra o item (1). Observe que em (*) usamos o fato de que

$$\sum_{P,R}\sum_{Q} -\mu(P\cap Q\cap R)\log\frac{\mu(P\cap R)}{\mu(R)} = \sum_{P\in\mathcal{P},R\in\mathcal{R}} -\mu(P\cap R)\log\frac{\mu(P\cap R)}{\mu(R)}$$

já que

$$\sum_{Q}\mu(P\cap Q\cap R) = \mu(P\cap R).$$

Agora, para demonstrar (2) observe que se $\mathcal{P}\prec\mathcal{Q}$ então

$$\begin{aligned} H_\mu(\mathcal{P}/\mathcal{R}) &= \sum_{P}\sum_{R}\sum_{Q\subset P} -\mu(Q\cap R)\log\frac{\mu(P\cap R)}{\mu(R)} \\ &\leq \sum_{P}\sum_{R}\sum_{Q\subset P} -\mu(Q\cap R)\log\frac{\mu(Q\cap R)}{\mu(R)} \\ &= H_\mu(\mathcal{Q}/\mathcal{R}). \end{aligned}$$

E isto prova a primeira parte de (2). A segunda parte de (2) é deixada como exercício para o leitor.
Finalmente, segue da definição que $H_\mu(\mathcal{P}/\mathcal{Q}) = 0$ se, e somente se, para todo $P\in\mathcal{P}$ e todo $Q\in\mathcal{Q}$,

$$\mu(P\cap Q) = 0 \text{ ou então } \frac{\mu(P\cap Q)}{\mu(Q)} = 1.$$

Em outras palavras, ou Q é disjunto de P (a menos de medida nula) ou Q está contido em P (a menos de medida nula). Isto quer dizer que $H_\mu(\mathcal{P}/\mathcal{Q}) = 0$ se, e somente se, $\mathcal{P}\prec\mathcal{Q}$ e isto demonstra (3). ■

Observe que se μ é invariante, então $H_\mu(f^{-1}(\mathcal{P})) = H_\mu(\mathcal{P})$. Além disso, tomando $\mathcal{R} = \mathcal{M}$(partição trivial) no item (1), vemos que

$$H_\mu(\mathcal{P}\vee\mathcal{Q}) = H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q}/\mathcal{P}) \leq H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q}).$$

Consideremos de agora em diante (a menos de menção explícita) $f : M \to M$ mensurável e μ invariante por f. Dada uma partição $\mathcal{P}$ com entropia finita, denotamos por $\mathcal{P}^n = \vee_{j=0}^{n-1} f^{-j}(\mathcal{P})$ para cada $n \geq 1$. Observe que o elemento $\mathcal{P}^n(x)$ que contém $x \in M$ está dado por:

$$\mathcal{P}^n(x) = \mathcal{P}(x) \cap f^{-1}(\mathcal{P}(f(x))) \cap f^{-2}(\mathcal{P}(f^2(x))) \cap \cdots \cap f^{-n+1}(\mathcal{P}(f^{n-1}(x))).$$

Teorema 2.2.1 $H_\mu(\mathcal{P}^{m+n}) \leq H_\mu(\mathcal{P}^m) + H_\mu(\mathcal{P}^n)$ *para todo* $m, n \geq 1$.

Demonstração: $\mathcal{P}^{m+n} = \vee_{i=0}^{m+n-1} f^{-i}(\mathcal{P}) = \mathcal{P}^m \vee f^{-m}(\mathcal{P}^n)$. Portanto

$$H_\mu(\mathcal{P}^{m+n}) \leq H_\mu(\mathcal{P}^m) + H_\mu(f^{-m}(\mathcal{P}^n))$$

Como μ é invariante, temos que $H_\mu(f^{-m}(\mathcal{P}^n)) = H_\mu(\mathcal{P}^n)$ para todo m, n. Assim, concluímos que

$$H_\mu(\mathcal{P}^{m+n}) \leq H_\mu(\mathcal{P}^m) + H_\mu(\mathcal{P}^n).$$

Podemos então definir a **entropia da transformação f com respeito a partição $\mathcal{P}$ e a probabilidade invariante μ** por

$$h_\mu(f, \mathcal{P}) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{P})).$$

Por fim, a **entropia métrica de f com respeito a μ** é

$$h_\mu(f) := sup\{h_\mu(f, \mathcal{P}) : \mathcal{P} \text{ partição finita}\},$$

Observe que se $\mathcal{P} \prec \mathcal{Q}$, então $\mathcal{P}^n \prec \mathcal{Q}^n$ e portanto $H_\mu(\mathcal{P}^n) \leq H_\mu(\mathcal{Q}^n)$, o que mostra que $h_\mu(f, \mathcal{P}) \leq h_\mu(f, \mathcal{Q})$

Exemplo 2.2.1 *Considere a expansão decimal* $f : [0, 1] \to [0, 1]$ *dada por* $f(x) = 10x - [10x]$. *O leitor pode verificar que a medida* μ *de Lebesgue é invariante por* f. *Seja* $\mathcal{P}$ *a partição nos intervalos da forma* $(\frac{(i-1)}{10}, \frac{i}{10})$ *com* $i = 1, 2, ..., 10$. *Então* $\mathcal{P}^n$ *é a partição nos intervalos da forma* $(\frac{(i-1)}{10^n}, \frac{i}{10^n})$ *com* $i = 1, 2, ..., 10^n$. *Então*

$$H_\mu(\mathcal{P}^n) = \sum_{i=1}^{10^n} -10^{-n}\log 10^{-n} = n\log 10.$$

e daí

$$h_\mu(f,\mathcal{P}) = \lim_{n\to\infty}\frac{1}{n}H_\mu(\mathcal{P}^n) = \lim_{n\to\infty}\frac{1}{n}n\log 10 = \log 10.$$

Lema 2.2.2 *$h_\mu(f,\mathcal{P}) = \lim_n H_\mu(\mathcal{P}/\vee_{j=1}^n f^{-j}(\mathcal{P}))$ para qualquer partição $\mathcal{P}$ com entropia finita.*

Demonstração: Lembre pelo Lema 2.2.1(1) que $H_\mu(\mathcal{P}'\vee\mathcal{Q}'/\mathcal{R}) = H\mu(\mathcal{P}'/\mathcal{R}) + H_\mu(\mathcal{Q}'/\mathcal{P}'\vee\mathcal{R})$, assim tomando $\mathcal{R} = M, \mathcal{Q}' = \mathcal{P}, \mathcal{P}' = \vee_{j=1}^{n-1} f^{-j}(\mathcal{P})$ e o fato de que a medida μ é invariante, temos

$$H_\mu(\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})) = H_\mu(\bigvee_{j=1}^{n-1} f^{-j}(\mathcal{P})) + H_\mu(\mathcal{P}/\bigvee_{j=1}^{n-1} f^{-j}(\mathcal{P})) =$$
$$H_\mu(\bigvee_{j=0}^{n-2} f^{-j}(\mathcal{P}) + H_\mu(\mathcal{P}/\bigvee_{j=1}^{n-1} f^{-j}(\mathcal{P}))$$

para todo n. Assim, por recorrência temos

$$H_\mu(\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})) = H_\mu(\mathcal{P}) + \sum_{k=1}^{n-1} H_\mu(\mathcal{P}/\bigvee_{j=1}^{k} f^{-j}(\mathcal{P}))$$

Portanto, $h_\mu(f,\mathcal{P})$ é dada pelo limite Cesaro

$$h_\mu(f,\mathcal{P}) = \lim_n \frac{1}{n}H_\mu(\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})) = \lim_n \frac{1}{n}\sum_{k=1}^{n-1}(\mathcal{P}/\bigvee_{j=1}^{k} f^{-j}(\mathcal{P}))$$

Por outro lado, o Lema 2.2.1(2) garante que a sequência $H_\mu(\mathcal{P}/\vee_{j=1}^n f^{-j}(\mathcal{P}))$ é decrescente e limitada, portanto $\lim_n H_\mu(\mathcal{P}/\vee_{j=1}^n f^{-j}(\mathcal{P}))$ existe e é igual a $\lim_n \frac{1}{n}\sum_{k=1}^{n-1}(\mathcal{P}/\bigvee_{j=1}^k f^{-j}(\mathcal{P}))$. Concluímos assim que

$$h_\mu(f,\mathcal{P}) = H_\mu(\mathcal{P}/\vee_{j=1}^n f^{-j}(\mathcal{P})).$$

Lema 2.2.3 *Se $\mathcal{P}$ é partição com entropia finita, então $h_\mu(f,\mathcal{P}) = h_\mu(f,\mathcal{P}^k)$ para todo $k\geq 1$.*

Demonstração: Observe que dado qualquer $n \geq 1$

$$\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P}^k) = \bigvee_{j=0}^{n-1} f^{-j}(\bigvee_{i=0}^{k-1} f^{-i}(\mathcal{P})) = \bigvee_{l=0}^{n+k-2} f^{-l}(\mathcal{P}) = \mathcal{P}^{n+k-1}.$$

Portanto

$$h_\mu(f, \mathcal{P}^k) = \lim_n \frac{1}{n} H_\mu(\mathcal{P}^{n+k-1}) = \lim_n \frac{1}{n} H_\mu(\mathcal{P}^n) = h_\mu(f, \mathcal{P}).$$

A principal dificuldade no cálculo da entropia reside no fato de calcular o supremo. O resultado a seguir facilita essa dificuldade

Teorema 2.2.2 *(Kolmogorov-Sinai) Seja $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \prec \mathcal{P}_n \prec \cdots$ uma sequência não-decrescente de partições com entropia finita tal que $\cup_{n=1}^{\infty} \mathcal{P}_n$ gera a σ-álgebra dos conjuntos mensuráveis a menos de medida nula. Então*

$$h_\mu(f) = \lim_n h_\mu(f, \mathcal{P}_n).$$

Corolário 2.2.1 *Seja $\mathcal{P}$ uma partição com entropia finita tal que a união dos seus iterados $\mathcal{P}^n = \vee_{j=0}^{n-1} f^{-j}(\mathcal{P})$, $n \geq 1$ gera a σ-álgebra dos conjuntos mensuráveis. Então $h_\mu(f) = h_\mu(f, \mathcal{P})$.*

Demonstração: Sabemos que $\mathcal{P}^1 \prec \mathcal{P}^2 \prec \cdots \prec \mathcal{P}^n \prec \cdots$. Pelo Lema 2.2.3, $h_\mu(f, \mathcal{P}) = h_\mu(f, \mathcal{P}^k)$. Assim pelo teorema de Kolgomorov-Sinai, $h_\mu(f) = \lim_n h_\mu(f, \mathcal{P}^n) = \lim_n h_\mu(f, \mathcal{P}) = h_\mu(f, \mathcal{P})$.

Exemplo 2.2.2 *Considere a transformação expansão decimal $f : [0,1] \to [0,1]$, dada por $f(x) = 10x - [10x]$, f preserva a medida de Lebesgue. Seja $\mathcal{P}$ a partição $[0,1]$ nos intervalos da forma $(\frac{i-1}{10}, \frac{i}{10}]$ com $i = 1, ..., 10$. Então $\mathcal{P}^n$ é a partição nos intervalos da forma $(\frac{i-1}{10^n}, \frac{i}{10^n}]$ com $i = 1, ..., 10^n$. Pelo Exemplo 2.2.1 vimos que*

$$h_\mu(f, \mathcal{P}) = \lim_n \frac{1}{n} H_\mu(\mathcal{P}^n) = \log 10.$$

Agora, como $\mathcal{P}$ gera a σ-álgebra dos conjuntos mensuráveis, pelo Corolário 2.2.1 podemos concluir que a entropia da transformação expansão decimal é

$$h_\mu(f) = h_\mu(f, \mathcal{P}) = \log 10.$$

3 Fórmula de Rokhlin

3.1 Jacobiano

A noção de Jacobiano é claramente uma generalização da ideia de Jacobiano vista em cálculo, o Jacobiano de um difeomorfismo com respeito ao volume de $\mathbb{R}^n$. Neste capítulo iremos definir o que vem a ser o Jacobiano de uma função localmente invertível relativamente a uma medida η. Mais adiante, mostraremos o quão importante é a existência de Jacobiano para facilitar o cálculo da entropia.

Definição 3.1.1 *Seja $f : M \to M$ uma transformação mensurável. Diremos que f é localmente invertível se existe alguma cobertura enumerável $\{U_k : k \geq 1\}$ de M por conjuntos mensuráveis tais que a restrição de f a cada U_k é uma bijeção sobre a sua imagem, a qual é um conjunto mensurável, e a inversa dessa bijeção também é mensurável. Os subconjuntos mensuráveis destes conjuntos U_k serão chamados domínios de injetividade.*

Observação: 3.1.1 *Observe que se f é localmente invertível então a pré-imagem $f^{-1}(y)$ de qualquer $y \in M$ é enumerável, de fato, como a restrição a cada U_k é uma bijeção, a pré imagem de y contém no máximo um ponto em cada U_k, que por sua vez é enumerável.*

Definição 3.1.2 *Seja η uma probabilidade em M, não necessariamente invariante por f. Uma função mensurável $\xi : M \to [0, \infty)$ é um jacobiano de f relativamente a η se a restrição de ξ a qualquer domínio de injetividade A é integrável com relação a η e satisfaz*

$$\eta(f(A)) = \int_A \xi \; d\eta$$

Observação: Note que a definição não depende da escolha da cobertura $\{U_k : k \geq 1\}$

Dizemos que uma medida é **não singular** com relação a transformação f se a imagem de qualquer domínio de invertibilidade com medida nula também tem medida nula: Se $\eta(A) = 0$ então $\eta(f(A)) = 0$. Por exemplo, se $f : U \to U$ é um difeomorfismo local num aberto de $\mathbb{R}^d$ e η é a medida de Lebesgue, então η é não singular (fórmula de mudança de variáveis).
É fácil ver que se f admite jacobiano com relação a uma medida η então essa medida é não singular. Vamos mostrar que a recíproca também é verdadeira.

Teorema 3.1.1 *Seja $f : M \to M$ uma transformação localmente invertível e seja η uma medida boreliana em M, não singular com relação a f. Então, existe algum jacobiano de f com relação a η e ele é essencialmente único: dois jacobianos quaisquer coincidem em η-quase todo ponto.*

Demonstração. : Provaremos inicialmente a existência. Seja $\{U_k : k \geq 1\}$ uma cobertura enumerável arbitrária de M, onde cada U_k é um domínio de invertibilidade de f, defina $P_1 = U_1$ e $P_k = U_k \backslash (U_1 \cup \cdots \cup U_{k-1})$ para cada $k > 1$. Então, $\mathcal{P} = \{P_k : k \geq 1\}$ é uma partição de M formada por domínios de invertibilidade. Para cada $P_k \in \mathcal{P}$, represente por η_k a medida definida em P_k por $\eta_k(A) = \eta(f(A))$. Em outras palavras, η_k é a imagem por $(f|P_k)^{-1}$ da medida η restrita a $f(P_k)$. A hipótese de que η é não singular implica que cada η_k é absolutamente contínua com relação a η restrita a P_k:

$$\eta(A) = 0 \Rightarrow \eta_k(A) = \eta(f(A)) = 0$$

para todo conjunto mensurável $A \subset P_k$. Seja $\xi_k = d\eta_k / d(\eta|P_k)$ a derivada de Radón-Nykodim. Então ξ_k é uma função definida em P_k, integrável com relação a η e satisfazendo

$$\eta(f(A)) = \eta_k(A) = \int_A \xi_k \, d\eta$$

para todo conjunto mensurável $A \subset P_k$. Considere a função $\xi : M \to [0, \infty)$ cuja restrição a cada $P_k \in \mathcal{P}$ está dado por ξ_k. Todo subconjunto de U_k pode ser escrito como união disjunta de subconjuntos de $P_1, ..., P_k$. Aplicando a igualdade anterior a cada um desses subconjuntos e somando as respectivas igualdades, obtemos que

$$\eta(f(A)) = \int_A \xi \, d\eta \text{ para todo conjunto mensurável } A \subset U_k \text{ e } k \geq 1.$$

Isto prova que ξ é um jacobiano de f relativamente a η.

Agora, suponha que ξ e ζ são jacobianos de f relativamente a η e que existe $B \subset M$ com $\eta(B) > 0$ tal que $\xi(x) \neq \zeta(x)$ para todo $x \in B$. A menos de substituir B por um subconjunto adequado, e permutar os papéis de ξ e ζ é necessário, podemos supor que $\xi(x) < \zeta(x)$ para todo $x \in B$. De modo similar, podemos supor que B está contigo em algum U_k. Então,

$$\eta(f(B)) = \int_B \xi \, d\eta < \int_B \zeta \, d\eta = \eta(f(B)).$$

Contradição, logo o jacobiano é único. ■

Usaremos a notação $J_\eta f$ para representar o jacobiano de f com relação a η, quando exista. Por definição, $J_\eta f$ é integrável em cada domínio de invertibilidade.

Corolário 3.1.1 *Seja $f : M \to M$ uma transformação localmente invertível e η uma probabilidade boreliana em M não singular com relação a f. Então valem as seguintes fórmulas de mudanças de variáveis.*

1. *$\int_{f(A)} \varphi \, d\eta = \int_A (\varphi \circ f) J_\eta f \, d\eta$ para todo domínio de invertibilidade $A \subset M$ e toda função mensurável $\varphi : f(A) \to \mathbb{R}$ tal que as integrais estão definidas (podendo ser $\pm\infty$)*

2. *$\int_A \psi \, d\eta = \int_{f(A)} (\psi / J_\eta f) \circ (f|A)^{-1} \, d\eta$ para qualquer função mensurável $\psi : A \to \mathbb{R}$ tal que as integrais estão definidas (podendo ser $\pm\infty$)*

Demonstração:

1. Como η é não singular, existe $J_\eta f$. Usando o teorema anterior temos

$$\int_{f(A)} \mathcal{X}_{f(A)} \, d\eta = \eta(f(A)) = \int_A J_\eta f \, d\eta = \int_A (\mathcal{X}_{f(A)} \circ f) J_\eta f \, d\eta$$

   E portanto a equação se verifica para funções características. Pelo mesmo argumento se verifica para funções simples. Usando o Teorema da Convergência Dominada verifica-se que o lema é válido para toda $\varphi : f(A) \to \mathbb{R}$ mensurável, já que podemos aproximar qualquer função mensurável integrável por funções simples.

2. Aplicando o item (1) à função $\varphi = (\psi / J_\eta f) \circ (f|A)^{-1}$ temos justamente o resultado desejado.

Não é verdade que toda probabilidade invariante η é não singular, é fácil construir contra-exemplos para isto.

3.2 Fórmula de Rokhlin

Teorema 3.2.1 *Seja $f : M \to M$ uma transformação localmente invertível e seja μ uma probabilidade invariante por f. Suponha que existe alguma partição finita ou enumerável $\mathcal{P}$ tal que $\cup_n \mathcal{P}^n$ gera a σ-álgebra de M e todo $P \in \mathcal{P}$ é domínio de invertibilidade de f. Então $h_\mu(f) = \int log J_\mu f \ d\mu$.*

Demonstração: Consideremos a sequência de partições $\mathcal{Q}_n = \vee_{j=1}^n f^{-j}(\mathcal{P})$. Pelo Corolário 2.2.1 e pelo Lema 2.2.2,

$$h_\mu(f) = h_\mu(f, \mathcal{P}) = lim_n H_\mu(\mathcal{P}/\mathcal{Q}_n). \tag{3.1}$$

Por definição (como anteriormente, $\phi(x) = -xlogx$)

$$H_\mu(\mathcal{P}/\mathcal{Q}_n) = \sum_{P\in\mathcal{P}} \sum_{Q_n\in\mathcal{Q}_n} -\mu(P\cap \mathcal{Q}_n) log \frac{\mu(P\cap\mathcal{Q}_n)}{\mu(Q_n)} = \sum_{P\in\mathcal{P}} \sum_{Q_n\in\mathcal{Q}_n} \mu(\mathcal{Q}_n)\phi(\frac{\mu(P\cap\mathcal{Q}_n)}{\mu(Q_n)}) \tag{3.2}$$

Seja $e_n(\phi, x)$ a esperança condicional de uma função ϕ relativamente à partição $\mathcal{Q}_n$ e seja $e(\phi, x)$ o seu limite quando n vai para o infinito. É claro da definição que

$$\frac{\mu(P\cap\mathcal{Q}_n)}{\mu(\mathcal{Q}_n)} = e_n(\mathcal{X}_P, x) \text{ para todo } x \in Q_n \text{ e todo } Q_n \in \mathcal{Q}_n.$$

Portanto,

$$\sum_{P\in\mathcal{P}} \sum_{Q_n \mathcal{Q}_n} \mu(Q_n)\phi(\frac{\mu(P\cap Q_n)}{\mu(Q_n)}) = \sum_{P\in\mathcal{P}} \int \phi(e_n(\mathcal{X}_P, x)) \ d\mu(x). \tag{3.3}$$

Pelo lema 1.2.1, o limite $e(\mathcal{X}_P, x) = \lim_n e_n(\mathcal{X}_P, x)$ existe para μ-quase todo ponto x. Então, observando que a função ϕ é limitada, podemos usar o teorema da convergência dominada para deduzir das relações (3.1) - (3.3) que

$$h_\mu(f) = \sum_{P\in\mathcal{P}} \int \phi(e(\mathcal{X}_P, x)) \ d\mu(x). \tag{3.4}$$

Precisaremos do seguinte lema para relacionar o jacobiano com o integrando do lado direito.

Lema 3.2.1 *Para toda função mensurável limitada $\psi : M \to \mathbb{R}$ e toda probabilidade boreliana η invariante por f,*

$$e(\psi, x) = \hat{\psi}(f(x)) \text{ para } \eta\text{-quase todo } x, \text{ onde } \hat{\psi}(y) = \sum_{z \in f^{-1}(y)} \frac{\psi}{J_\eta f}(z).$$

Demonstração. Lembre que $\mathcal{Q}_n = \vee_{j=1}^{n} f^{-j}(\mathcal{P})$. Também usaremos a sequência de partições $\mathcal{P}^n = \vee_{j=0}^{n-1} f^{-j}(\mathcal{P})$. Observe que $\mathcal{Q}_n(x) = f^{-1}(\mathcal{P}^{n-1}(f(x)))$ e $\mathcal{P}^n(x) = \mathcal{P}^n(x) = \mathcal{P}(x) \cap \mathcal{Q}_n(x)$ para todo n e todo x. Então,

$$\int_{\mathcal{P}^{n-1}(f(x))} \hat{\psi}\, d\eta = \sum_{P \in \mathcal{P}} \int_{f(P) \cap \mathcal{P}^{n-1}(f(x))} \frac{\psi}{J_\eta f} \circ (f|P)^{-1}\, d\eta$$

Usando a fórmula de mudança de variáveis dada no Corolário 3.1.1, a expressão do lado direito pode ser reescrita como

$$\sum_{P \in \mathcal{P}} \int_{P \cap \mathcal{Q}_n(x)} \psi(z)\, d\eta(z) = \int_{\mathcal{Q}_n(x)} \psi\, d\eta.$$

Portanto,

$$\int_{\mathcal{P}^{n-1}(f(x))} \hat{\psi}\, d\eta = \int_{\mathcal{Q}_n(x)} \psi\, d\eta \tag{3.5}$$

A hipótese de que η é invariante dá que $\eta(\mathcal{P}^{n-1}(f(x)) = \eta(\mathcal{Q}_n(x))$. Dividindo ambos os lados de (3.5) por este número, obtemos que

$$e_n(\psi, x) = e'_{n-1}(\hat{\psi}, f(x)) \text{ para todo } x \text{ e todo } n > 1.$$

Então, passando o limite, $e(\psi, x) = e'(\hat{\psi}, f(x))$ para η-quase todo x. Por outro lado, é fácil ver que $e'(\hat{\psi}, y) = \hat{\psi}(y)$ para η-quase todo $y \in M$. ∎

Vamos aplicar este resultado a $\psi = \mathcal{X}_P$ e $\eta = \mu$. Como f é injetiva em todo elemento de $\mathcal{P}$, cada interseção $P \cap f^{-1}(y)$ ou é vazia ou contém exatamente um ponto. Portanto, segue do Lema anterior que $e(\mathcal{X}_P, x) = \hat{\mathcal{X}}_P(f(x))$, com

$$\hat{\mathcal{X}}_P(y) = \begin{cases} 1/J_\mu f((f \mid P)^{-1}(y)) & \text{se } y \in f(P); \\ 0 & \text{se } y \notin f(P). \end{cases}$$

Então, lembrando que a medida μ é invariante,

$$\int \phi(e(\mathcal{X}_P, x))\ d\mu(x) = \int \phi(\hat{\mathcal{X}}_P(y))\ d\mu(y) = \int_{f(P)} (\frac{1}{J_\mu f} \log J_\mu f) \circ (f|P)^{-1}\ d\mu = \int_P \log J_\mu f\ d\mu$$

(usamos na última igualdade a parte (2) do corolário 3.1.1). Substituindo essa expressão em (3.4), vem que

$$h_\mu(f) = \sum_{P \in \mathcal{P}} \int_P \log J_\mu f\ d\mu = \int \log J_\mu f\ d\mu.$$

Exemplo 3.2.1 *Considere a expansão decimal* $f : [0,1] \to [0,1]$ *dada por* $f(x) = 10x - [10x]$. *A medida* μ *de Lebesgue é invariante por* f. *Considere a cobertura* $U_i = (\frac{(i-1)}{10}, \frac{i}{10})$ *com* $i = 1, 2, 3, ..., 10$, *então* f *restrita a cada* U_j *é uma bijeção e sua inversa é mensurável, logo* f *é localmente invertível. Essa mesma cobertura* $\mathcal{P} = \cup_{i=1}^{10} U_i$ *é uma partição que gera a* σ*-álgebra de Borel em* $[0,1]$. *Portanto, como* f *satisfaz as hipóteses do Teorema, podemos aplicar a Fórmula de Rokhlin para esta transformação. O jacobiano de* f *relativamente a* μ *é*

$$J_\mu f(x) = 10 \ \textit{para todo}\ x \in \mathcal{P}.$$

Logo, concluímos através da fórmula de Rokhlin que a entropia da expansão decimal é dada por

$$h_\mu(f) = \int \log J_\mu f\ d\mu = \int \log 10\ d\mu = \log 10.$$

Um outro fato interessante é dado pelo seguinte exemplo:

Exemplo 3.2.2 *Seja* $f : S^1 \to S^1$ *uma transformação de classe* C^2 *tal que* $|f'(x)| > \sigma > 1$ *para todo* $x \in S^1$. *Nestas condições sabemos que existe uma única medida* μ *absolutamente contínua com respeito à medida de Lebesgue* m *que é invariante por* f. *Além disso,* μ *é ergódica (Recomenda-se uma leitura do capitulo 11 da referência [2]).*

Neste caso, o Teorema de Birkhoff nos permite definir o número

$$\lambda = \lim \frac{1}{n} \sum_{i=0}^{n-1} \log |f'(f^i(x))| = \lim \frac{1}{n} \log |(f^n)'(x)|.$$

Este número é chamado ***Expoente de Lyapunov*** *de μ com respeito a f.*

Pode-se mostrar que sob estas condições, existe $\delta > 0$ tal que para toda partição $\mathcal{P}$ com diâmetro menor que δ, $\mathcal{P}$ é geradora. Assim, podemos usar a fórmula de Rokhlin para concluir que

$$h_\mu(f) = \int \log J_\mu f \; d\mu$$

Vamos checar que $J_\mu f = \frac{h \circ f}{h}|f'|$. De fato, como $\mu = hm$, podemos concluir que

$$\mu(f(A)) = \int_{f(A)} 1 \; d\mu = \int_{f(A)} h \; dm = \int_A |f'|(h \circ f) \; dm = \int_A |f'| \frac{h \circ f}{h} \; d\mu,$$

perceba que na terceira igualdade usamos a fórmula de mudança de variáveis.

Portanto, concluímos que a entropia de f com respeito a μ é dada por

$$h_\mu(f) = \int \log(J_\mu f) \; d\mu = \int \log(|f'| \frac{h \circ f}{h}) \; d\mu.$$

Esta igualdade é uma caso particular da chamada Fórmula de Pesin que afirma que:

Teorema 3.2.2 *Seja $f : M \to M$ uma transformação expansora numa variedade Riemanniana compacta, tal que a derivada Df é Holder. Seja μ a única probabilidade invariante absolutamente contínua com respeito à medida de Lebesgue em M. Então*

$$h_\mu(f) = \sum_{i=1}^{k} d_i \lambda_i,$$

onde λ_i, $i = 1, \cdots, k$ são os expoentes de Lyapunov de f em μ-quase todo ponto e d_i, $i = 1, \cdots, k$ são as respectivas multiplicidades.

Exemplo 3.2.3 *Seja $G(x) = \frac{1}{x} - [\frac{1}{x}]$ a transformação de Gauss, pode-se checar que a medida*

$$\mu(E) = \int_E \frac{1}{(1+x)\log 2} dx$$

é invariante por G.

Seja $\mathcal{P}$ a partição nos intervalos $(\frac{1}{m+1}, \frac{1}{m})$ para $m \geq 1$. Perceba que G é localmente invertível e a partição $\mathcal{P}$ é geradora. Portanto, pelo resultado acima podemos concluir que o jacobiano de G é dado por

$$J_\mu G = \frac{h \circ G}{h} |G'|.$$

onde $h = \dfrac{1}{(1+x)\log 2}$. *Então, podemos concluir pela Fórmula de Rokhlin que*

$$h_\mu(G) = \int \log(\frac{h \circ G}{h} |G'|) \; d\mu = \int \log |G'| \; d\mu = \int_0^1 \frac{-2 \log x \; dx}{(1+x)\log 2} = \frac{\pi^2}{6 \log 2}.$$

onde a segunda igualdade vem do fato da medida μ *ser invariante por* G.

4 Sistemas Dinâmicos Aleatórios

Neste trabalho, um sistema dinâmico aleatório (RDS) consiste da seguinte dupla:

i. Um espaço métrico completo e separável $(X, \mathcal{F}, \mathbb{P})$ e uma transformação invertível $\theta : X \to X$ preservando uma probabilidade boreliana ergódica $\mathbb{P}$.

ii. Uma variedade Riemanniana Y conexa e compacta; uma transformação mensurável $F : X \times Y \to X \times Y$ da seguinte forma

$$F(x, y) = (\theta(x), f_x(y))$$

onde cada $f_x : Y \to Y$ é um difeomorfismo local C^1.

Denotamos por $\mathcal{J} = X \times Y$ e $\mathcal{J}_x = \{x\} \times Y$ a fibra de $x \in X$. Definimos para cada inteiro $n \geq 0$ e cada $x \in X$

$$f_x^n(y) := f_{\theta^{n-1}(x)} \circ \cdots \circ f_{\theta(x)} \circ f_x(y)$$

tal que $F^n(x, y) = (\theta^n(x), f_x^n(y))$.

Exemplo 4.0.1 *Considere $X = \{0, 1\}^{\mathbb{Z}}$, $\mathbb{P}$ =Bernoulli, $\theta = \sigma$ (descolamento de Bernoulli) e Y uma variedade Riemanniana conexa e compacta. Considere as funções*

$$f_x = \begin{cases} f_0 & \textit{se } x_0 = 0; \\ f_1 & \textit{se } x_0 = 1. \end{cases}$$

onde $f_i : Y \to Y$ são difeomorfismos. É um sistema dinâmico aleatório, consiste em "iterar dois difeomorfismos aleatoriamente de acordo com o lançamento de uma moeda".

Seja $\mathcal{M}^1_{\mathbb{P}}(\mathcal{J})$ conjunto das medidas de probabilidade sobre $(\mathcal{J}, \mathcal{B})$ tal que

$$\mu \circ \pi_X^{-1} = \mathbb{P}$$

onde $\pi_X : \mathcal{J} \to X$ é a projeção na primeira coordenada ($\pi_X(x,y) = x$), e seja

$$\mathcal{M}^1_{\mathbb{P}}(F) = \{\mu \in \mathcal{M}^1_{\mathbb{P}}(\mathcal{J}) : \mu \circ F^{-1} = \mu\}$$

Denotamos por ϵ_X a partição de X sobre pontos. A partição $\pi_X^{-1}(\epsilon_X)$ é mensurável (veja Exemplo 1.1.4), portanto, para cada $\mu \in \mathcal{M}^1_{\mathbb{P}}(\mathcal{J})$, pelo teorema da desintegração de Rokhlin, existe um sistema de medidas $(\mu_x)_{x \in X}$ tal que $\mu = \int \mu_x d\mathbb{P}(x)$, chamando-se sistema canônico de medidas condicionais. Por conveniência, dado uma medida $\mu \in \mathcal{M}^1_{\mathbb{P}}(F)$, denotaremos por $\mathcal{A}_\mu(F|\theta)$ o conjunto de todas partições mensuráveis α de $\mathcal{J}$ que são mais finas que $\pi_X^{-1}(\epsilon_X)$.

Definição 4.0.1 *Uma partição $\mathcal{P} \in \mathcal{A}_\mu(F|\theta)$ é geradora por F relativa a θ se e somente se*

$$\mathcal{P}^\infty := \bigvee_{j=0}^{\infty} F^{-j}(\mathcal{P}) \equiv_\mu \epsilon_J.$$

onde ϵ_J é a partição de sob $\mathcal{J} = \bigcup_{x \in X} \{x\} \times Y$ sobre elementos unitários.

A relação $\equiv_\mu$ significa que dado duas partições $\mathcal{P}_1$ e $\mathcal{P}_2$ de $\mathcal{J}$, então $\mathcal{P}_1 \equiv_\mu \mathcal{P}_2$ se existe um conjunto mensurável $W \subset \mathcal{J}$ com $\mu(\mathcal{J}/W) = 0$ tal que $\mathcal{P}_1|_W = \mathcal{P}_2|_W$.

Para um RDS podemos definir também a noção de entropia de uma forma bastante similar com a definição introduzida no capítulo anterior. Perceba que no capítulo anterior a noção de entropia era dada em cima de partições enumeráveis. Agora, podemos estender essa noção para uma partição mensurável (não necessariamente enumerável).

4.1 Entropia para sistemas dinâmicos aleatórios

Vamos generalizar algumas noções de entropia para partições mensuráveis. A tripla $(X, \mathcal{F}, \mu)$ é um espaço de Lebesgue. Para nós, um espaço de Lebesgue é um espaço métrico compacto e separável.

Definição 4.1.1 *Se $\mathcal{A}$ é uma partição mensurável de X então a entropia é definida da seguinte maneira:*

i. $H(\mathcal{A}) = \infty$ *se $\mathcal{A}$ é uma partição não enumerável.*

ii. $H(\mathcal{A}) = -\sum_{A\in\mathcal{A}} \mu(A)\log\mu(A)$ *se $\mathcal{A}$ é uma partição enumerável.*

Definição 4.1.2 *Se $\mathcal{A}$ e $\mathcal{B}$ são partições mensuráveis de X, então a entropia condicional $H_\mu(\mathcal{A}|\mathcal{B})$ da partição $\mathcal{A}$ sujeita a $\mathcal{B}$ é definida por*

$$H_\mu(\mathcal{A}|\mathcal{B}) = \int_{X/\mathcal{B}} H_{\mu_B}(\mathcal{A}|B)d\hat{\mu}(B),$$

onde $(\mathcal{A}|B) = \{A\cap B|A\in\mathcal{A}\}$. É claro que podemos reescrever a definição acima como

$$H_\mu(\mathcal{A}|\mathcal{B}) = \int_X H_{\mu_{\mathcal{B}(x)}}(\mathcal{A}|\mathcal{B}(x))d\mu(x)$$

Portanto, podemos escrever como estamos habituados

$$H_\mu(\mathcal{A}|\mathcal{B}) = \int_X I(\mathcal{A}|\mathcal{B})d\mu,$$

onde $I(\mathcal{A}|\mathcal{B})$ é a ***função informação condicional*** *definida por:*

$$I(\mathcal{A}|\mathcal{B})(x) := -\log\mu_{\mathcal{B}(x)}(\mathcal{A}(x)\cap\mathcal{B}(x)).$$

Definição 4.1.3 *Se $\mu\in\mathcal{M}^1_{\mathbb{P}}(F)$ e $\mathcal{P}$ é uma partição mensurável de $\mathcal{J}$ mais fina que $\pi_X^{-1}(\epsilon_X)$, então*

$$h_\mu(F|\theta;\mathcal{P}) := \lim_{n\to\infty}\frac{1}{n}H_\mu(\mathcal{P}^n|\pi_X^{-1}(\epsilon_X))$$

onde

$$\mathcal{P}^n = \bigvee_{n=0}^{n-1} F^{-j}(\mathcal{P}).$$

Além do mais, seja

$$h_\mu(F|\theta) := \sup\{h_\mu(F|\theta; \mathcal{P})\}$$

onde o supremo é tomado sobre todas as partições mensuráveis $\mathcal{P}$ de $\mathcal{J}$ que são mais finas que $\pi_X^{-1}(\epsilon_X)$ e tem entropia finita relativa a $\pi_X^{-1}(\epsilon_X)$, isto é,

$$H_\mu(\mathcal{P}|\pi_X^{-1}(\epsilon_X)) := \int_X H_{\mu_x}(\mathcal{P}_x) d\mathbb{P}(x) < \infty$$

*onde $\mathcal{P}_x = \{P \cap \mathcal{J}_x : P \in \mathcal{P}\}$. O número $h_\mu(F|\theta)$ é chamado **entropia de F relativa a $\boldsymbol{\theta}$ com respeito a medida $\boldsymbol{\mu}$**.*

Daremos agora uma adaptação ao cenário aleatório do bem conhecido Teorema de Kolmogorov-Sinai. Sua declaração e um esboço de sua prova podem ser encontrados na referência [1].

Teorema 4.1.1 *Seja $F : \mathcal{J} \to \mathcal{J}$ é um sistema dinâmico aleatório métrico, se $\mu \in \mathcal{M}_{\mathbb{P}}^1(F)$, e se $\alpha \in \mathcal{A}_\mu(F|\theta)$ é uma partição geradora por F relativa a θ, então*

$$h_\mu(F|\theta) = h_\mu(F|\theta; \alpha) = H_\mu(\epsilon_{\mathcal{J}}|F^{-1}(\epsilon_{\mathcal{J}})).$$

Teorema 4.1.2 *Suponha que $\mathcal{A}$ e $\mathcal{B}$ são duas partições mensuráveis do Espaço de Lebesgue $(X, \mathcal{F}, \mu)$ tal que $\mathcal{A} \cap B$ é contável (mod 0 com respeito a μ_B) para quase todo $B \in \mathcal{B}$. Então existe uma partição contável $\gamma = \{\gamma_1, \gamma_2, ...\}$ de X (mod 0) tal que cada $\gamma_j \in \gamma$ intersecta quase todo B em não mais que um ponto, que é então um átomo de μ_B, em particular*

$$\mathcal{A} \vee \mathcal{B} = \gamma \vee \mathcal{B} \text{ (mod 0)}.$$

Observação: 4.1.1 *As demonstrações omitidos podem ser encontradas no capítulo 01 da referência [4].*

Definição 4.1.4 *Seja $(X, \mathcal{F}, \mu)$ um espaço de Lebesgue. Seja $T : X \to X$ uma transformação mensurável. Nós dizemos que T é **essencialmente contável** se as medidas μ_A do sistema canônico de medidas condicionais para a partição $\mathcal{A} := T^{-1}(\epsilon_X)$ são puramente atômicos (mod 0 com respeito a μ_A), para todo $A \in \mathcal{A}$.*

Lema 4.1.1 *Se T é essencialmente contável e preserva μ então existe um conjunto mensurável $Y \subset X$ de medida total tal que $T(Y) \subset Y$ e*

i. *$T^{-1}(x) \cap Y$ é contável para cada $x \in Y$. Além do mais, para cada $x \in Y$, $T^{-1}(x) \cap Y$ consiste somente de átomos da medida condicional $\mu_{T^{-1}(x)}$;*

ii. *$T(B)$ é mensurável se $B \subset Y$ é mensurável;*

iii. *$T\,|_Y$ é não singular, isto é, $\mu(B) = 0$ para $B \subset Y$ implica $\mu(T(B)) = 0$.*

Demonstração. A demonstração pode ser encontrada da referência [2]. ■

Teorema 4.1.3 *Seja $(X, \mathcal{F}, \mu)$ um Espaço de Lebesgue e $T : X \to X$ uma transformação mensurável preservando μ, essencialmente contável. Então existe o Jacobiano, ele é único μ-q.t.p..*

Demonstração. Vamos definir o Jacobiano num conjunto Y de medida total, no complementar de Y podemos definir o Jacobiano como sendo zero. Primeiramente, considere a partição $\gamma = \{\gamma_1, \gamma_2, ...\}$ dada pela Teorema 4.1.2 com $\mathcal{A} = \epsilon$ e $\mathcal{B} = T^{-1}(\epsilon)$, onde ϵ é a partição em pontos do conjunto X. Então para cada j, o mapa $T|_{\gamma_j \cap Y}$ é injetivo, onde Y é o conjunto de medida total dado pelo lema 4.1.1. Além do mais, $T|_Y$ é não singular, assim J existe em cada $\gamma_j \cap Y$ pelo Teorema de Radon-Nikodym, fazemos a mesma construção feita no Teorema 3.1.1 e concluímos o resultado. ■

Teorema 4.1.4 *Seja $(X, \mathcal{F}, \nu)$ um espaço de Lebesgue. Seja $T : X \to X$ uma transformação preservando ν, essencialmente contável. Então o Jacobiano J_ν tem logaritmo igual a $I_\nu(\epsilon | T^{-1}(\epsilon))$, onde $I(\mathcal{A}|\mathcal{B})(x) := -\log \mu_{\mathcal{B}(x)}(\mathcal{A}(x) \cap \mathcal{B}(x))$ é a função informação.*

Demonstração. Considere já T restrita a Y. Seja $Z \subset Y$ um conjunto mensurável tal que T é injetivo. Para cada $y \in Y$ denotamos por $A(y)$ o elemento de $\zeta = T^{-1}(\epsilon)$ contendo y. Nós obtemos

$$\begin{aligned}
\nu(T(Z)) &= \nu(T^{-1}(T(Z))) \\
&= \int_{T^{-1}(T(Z))} 1 \; d\nu(y) \\
&= \int_{T^{-1}(T(Z))} (1_Z(x)/\nu_A(y)\{x\}) \; d\nu_{A(y)}(x)) \; d\nu(y) \\
&= \int_{T^{-1}(T(Z))} (1_Z(y)/\nu_{A(y)}\{y\}) \; d\nu(y) \\
&= \int_Z (1/\nu_{A(y)}\{y\}) \; d\nu(y).
\end{aligned}$$

Assim sendo, $J_\nu(y) = 1/\nu_{A(y)}\{y\}$ e tem logaritmo igual a $I_\nu(\epsilon|T^{-1}(\epsilon))(y)$. De fato, $\log J_\nu(y) = \log(\frac{1}{\nu_{A(y)}\{y\}}) = -\log(\nu_{A(y)}\{y\})$, enquanto que $I(\epsilon|T^{-1}(\epsilon))(y) = -\log \nu_{A(y)}(\epsilon(y) \cap A(y)) = -\log(\nu_{A(y)}\{y\})$ ■

Por fim, podemos agora enunciar a Fórmula de Rokhlin para sistemas dinâmicos aleatórios métrico.

Teorema 4.1.5 *Se μ é uma medida ergódica invariante que admite uma partição $\mathcal{P}$ F-geradora com respeito a θ, então*

$$h_\mu(F|\theta) = \int \log J_\mu(F) d\mu.$$

Demonstração. Como F é essencialmente contável, podemos aplicar os teoremas anteriores. Daí pelo Teorema 4.1.1 e pelo Teorema 4.1.4 , como $\mathcal{P}$ é uma partição geradora por F relativa a θ, temos que

$$h_\mu(F|\theta) = H_\mu(\epsilon_{\mathcal{J}}|F^{-1}(\epsilon_{\mathcal{J}})) = \int I(\epsilon_{\mathcal{J}}|F^{-1}(\epsilon_{\mathcal{J}}))(x) d\mu(x) = \int \log J_\mu(F) d\mu.$$

■

Já sabemos que existe o Jacobiano em μ-q.t.p., isto é, temos uma função não-negativa e um conjunto $\mathcal{I} \subset \mathcal{J}$ de medida zero tal que para cada mensurável $A \subset J \backslash \mathcal{I}$ para qual F é injetora, vale a seguinte igualdade:

$$\mu(F(A)) = \int_A J_\mu F \; d\mu.$$

Agora, restringindo $J_\mu F$ a $\mathcal{J}_x$ e usando o conjunto de medida zero $\mathcal{I}$, consideremos a função $J_{\mu_x}(f_x) = J_\mu(F)|f_x$ sob a fibra $\mathcal{J}_x$ para $\mathbb{P}-$q.t.p. $x \in X$. Pela definição acima é claro que

$$\mu_{\theta(x)}(f_x(A_x)) = \int_{A_x} J_{\mu_x}(f_x)\ d\mu_x$$

para qualquer $A_x \subset \mathcal{J}_x \backslash \mathcal{I}_x$ mensurável tal que $f_x|A_x$ é injetora. Em particular, J_{μ_x} é o jacobiano de f_x relativa a μ_x. Portanto, podemos escrever o teorema acima como sendo

$$h_\mu(F|\theta) = \int \log J_\mu(F) d\mu = \int_X (\int_{\mathcal{J}_x} \log J_{\mu_x} f_x(y)\ d\mu_x(y))\ d\mathbb{P}(x).$$

Exemplo 4.1.1 *Seja* $X = \{x_0\}$, $\theta(x_0) = x_0$, $\mathbb{P} = \delta_{x_0}$ *e* Y *uma variedade Riemanniana conexa e compacta. Considere* $f_{x_0} : Y \to Y$ *um difeomorfismo local* C^1*. Sob as condições do Teorema 4.1.5, concluímos que a entropia de* F *relativa a* θ *é dada por*

$$h_\mu(F|\theta) = \int \log J_\mu F\ d\mu = \int_{x_0} \int_Y \log J_{\mu_{x_0}} f_{x_0}(y)\ d\mu_{x_0}(y))\ d\delta_{x_0} = \int_Y \log J_{\mu_{x_0}} f_{x_0}(y)\ d\mu_{x_0}(y) = h_{\mu_{x_0}}(f_{x_0})$$

Neste caso, a entropia no caso determinístico se iguala a entropia no caso aleatório.

Referências

[1.] Bartle, Robert G., Elements of Integration and Lebesgue Measure, John Wiley & Sons, 1995.

[2.] K. Oliveira, M. Viana *Fudamentos da Teoria Ergódica*, Colecao Fronteiras da Matematica, SBM, (2014).

[3.] Bilbao, Rafael Avarez. Medidas maximizantes em sistemas dinâmicos aleatórios. / Rafael Alvarez Bilbao. Maceió, 2015.

[4.] F. Przytycki and M. Urbanski, Conformal fractals- ergodic theory methods, London Mathematical Society Lecture Note 371 (2010).

[5.] D. Simmons and M. Urbanski, Relative equilibrium states and dimensions of

berwise invariant measures for distance expanding random maps, Stochastic and Dynamics 14,1 (2014).