Universidade Federal do Rio Grande do Sul

Instituto de Matemática e Estatística

Programa de Pós-Graduação em Matemática

Princípio Variacional para Sistemas com Entropia Topológica Infinita

Dissertação de Mestrado

Thomas Érico Jacobus

Porto Alegre, 5 abril de 2019.

Dissertação submetida por Thomas Érico Jacobus [1] como requisito parcial para a obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática e Estatística da Universidade Federal do Rio Grande do Sul.

Professor Orientador:
Fagner Bernardini Rodrigues (PPG-MAT-UFRGS)

Banca Examinadora:
Alexandre Tavares Baraviera (PPG-MAT-UFRGS)
Artur Oscar Lopes (PPG-MAT-UFRGS)
Rogério Ricardo Steffenon (UNISINOS)

Data da Apresentação: 5 de abril de 2019.

[1]Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

Agradecimentos

A meu orientador Fagner Bernardini Rodrigues pela amizade, conselhos, disposição e ensinamentos.

Aos professores Alexandre Tavares Baraviera, Artur Oscar Lopes e Rogério Ricardo Steffenon pelas valiosas sugestões e correções.

Aos meus professores e amigos Rogério Ricardo Steffenon e Rodrigo Orsini Braga pelos conselhos e incentivos.

A minha família, namorada e amigos pela paciência, compreensão e incentivo.

A todos os colegas e professores do Instituto de Matemática e Estatística.

Ao CNPq pelo auxílio financeiro.

Muito Obrigado!

Resumo

Nesta dissertação estudamos a dimensão métrica média que permite estender a noção de entropia para sistemas dinâmicos com entropia infinita. Este trabalho foi estruturado de acordo com o artigo de Anibal Velozo e Renato Velozo, estabelecendo um Princípio Variacional semelhante ao apresentado por Elon Lindenstrauss e Masaki Tsukamoto. Nos capítulos iniciais vamos apresentar alguns conceitos e pré-requisitos básicos para um melhor entendimento do principal capítulo deste trabalho.

Palavras-chave: Entropia, Dimensão Métrica Média, Princípio Variacional.

Abstract

In this dissertation we study the metric mean dimension that allows to extend the notion of entropy to dynamic systems with infinite entropy. This work was structured according to the article by Anibal Velozo and Renato Velozo, establishing a Variational Principle similar to that presented by Elon Lindenstrauss and Masaki Tsukamoto. In the initials chapters we will present some basic concepts and prerequisites for a better understanding of the main chapter of this work.

Keywords: Entropy, Metric Mean Dimension, Variational Principle.

Sumário

Capítulo 1

Introdução

No estudo de sistemas dinâmicos muitas vezes é útil conjugar um sistema dinâmico com outro que já possui propriedades invariantes por conjugação bem conhecidas, contribuindo para um melhor entendimento do sistema inicial. Uma das ferramentas que pode ser utilizada para afirmar se dois sistemas podem ou não serem conjugados é a entropia. Porém, com este invariante topológico não conseguimos distinguir sistemas dinâmicos que tenham entropia infinita.

Em [10] M. Gromov, com o objetivo de classificar sistemas com entropia topológica infinita, definiu o conceito de dimensão média. Este invariante topológico foi estudado mais tarde por E. Lindenstrauss e B. Weiss que, em [6], definiram a dimensão métrica média. Esta nova função pode ser utilizada como uma ferramenta para calcular cotas superiores para a dimensão média, além disso, foi provado por Lindenstrauss em [4] que ao considerar um sistema dinâmico $(\mathcal{X}, T)$ minimal (a órbita de qualquer ponto é densa) onde $\mathcal{X}$ é um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ contínua, existe uma distância definida em $\mathcal{X}$ tal que estas duas quantidades coincidem.

Intuitivamente, a dimensão média de um sistema dinâmico $(\mathcal{X}, T)$ conta o número médio de parâmetros necessários por iteração para descrever um ponto em $\mathcal{X}$. Por exemplo, ao considerarmos

$$\mathcal{X} = [0,1]^{\mathbb{Z}} = \cdots \times [0,1] \times [0,1] \times \cdots$$

e T a aplicação Shift, obtemos que este sistema dinâmico possui dimensão infinita e entropia infinita, mas sua dimensão média é igual a 1, significando que, intuitivamente, é necessário conhecer um parâmetro por iterado para descrever um ponto.

No presente trabalho focamos na dimensão métrica média e estudamos as ideias propostas por Anibal Velozo e Renato Velozo em [1]. Nesse artigo os autores obtém um tipo de princípio variacional que relaciona a dimensão métrica média e a entropia métrica, análogo ao provado por Elon Lindenstrauss e Masaki Tsukamoto em [5], porém, utilizando a função $h_\mu(\epsilon, T, \delta)$ ao invés da função taxa de distorção $R_\mu(\epsilon)$. Além disso, os autores também definiram uma nova função $\tilde{h}_\mu(\epsilon, T, \delta)$ e a relacionaram também com $R_\mu(\epsilon)$. Esta nova função é basicamente uma adaptação de $h_\mu(\epsilon, T, \delta)$ alterando a métrica envolvida, pois ao invés de considerar a distância máxima entre os segmentos de órbita passaram a considerar a distância média entre eles.

Esta dissertação é estruturada de tal forma que inicialmente vamos relembrar alguns conceitos importantes sobre Espaços Métricos, Teoria da Medida e Teoria Ergódica, cujos detalhes podem ser encontrados nas referências. Posteriormente, vamos abordar alguns resultados sobre entropia métrica e topológica, importantes para um melhor entendimento dos capítulos seguintes sobre o princípio variacional e a dimensão métrica média.

Capítulo 2

Preliminares

Este capítulo tem o objetivo de mostrar ou relembrar algumas definições e resultados importantes para o entendimento dos capítulos seguintes. Não faremos discussões profundas no assunto e omitiremos as demonstrações, encontradas nas referências.

2.1 Espaços Métricos

Um espaço métrico é um par $(\mathcal{X}, d)$ onde $\mathcal{X}$ é um conjunto e d uma métrica em $\mathcal{X}$. Uma métrica num conjunto $\mathcal{X}$ é uma função $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ que associa a cada par ordenado (x, y) de elementos de $\mathcal{X}$ um número real, chamado de distância de x a y. Além disso, esta função deve satisfazer as seguintes propriedades:

1. $d(x, x) = 0$;
2. Se $x \neq y$, então $d(x, y) > 0$;
3. $d(x, y) = d(y, x)$;
4. $d(x, z) \leqslant d(x, y) + d(y, z)$.

Dizemos que um conjunto $A \subset \mathcal{X}$ é limitado quando existe uma constante $D > 0$ tal que $\forall x, y \in A$ temos $d(x, y) \leqslant D$. Chamamos de diâmetro de um conjunto o menor dos números D tal que $d(x, y) \leqslant D$, ou seja,

$$diam(A) = \sup \{d(x, y) : x, y \in A\}.$$

Para x e $r > 0$ fixados,

$$B(x, r) = \{y \in \mathcal{X} : d(x, y) < r\}$$
$$\overline{B}(x, r) = \{y \in \mathcal{X} : d(x, y) \leqslant r\}$$

representam a bola aberta e fechada de centro x e raio r, respectivamente.

Em um espaço métrico $\mathcal{X}$ dizemos que o conjunto $A \subset \mathcal{X}$ é aberto se para todo elemento $a \in A$, existe um $\epsilon > 0$ tal que $B(x, \epsilon) \subset A$. Dizemos que um conjunto $B \subset \mathcal{X}$ é fechado se o complementar de B, ou seja, $B^c = \mathcal{X} \setminus B$ é aberto. A fronteira de $A \subset \mathcal{X}$, denotada por ∂A, é o conjunto formado pelos pontos $a \in A$, tais que toda bola aberta de centro a contém pelo menos um ponto de A e um ponto do complementar $\mathcal{X} \setminus A$.

Considerando $\mathcal{X}$ e $\mathcal{Y}$ espaços métricos. Dizemos que a aplicação $T : \mathcal{X} \to \mathcal{Y}$ é contínua no ponto $x \in \mathcal{X}$ se para todo $\epsilon > 0$ dado, é possível obter um $\delta > 0$ tal que $d(x, y) < \delta$ implica $d(T(x), T(y)) < \epsilon$. T é dita contínua se é contínua em todos os pontos de $\mathcal{X}$.

A próxima proposição será enunciada para subconjuntos abertos mas também temos um resultado análogo para subconjuntos fechados.

Proposição 2.1.1. *Sejam $\mathcal{X}$ e $\mathcal{Y}$ espaços métricos. A aplicação $T : \mathcal{X} \to \mathcal{Y}$ é contínua se, e somente se, a imagem inversa $T^{-1}(A)$ de todo subconjunto aberto de $\mathcal{Y}$ é um subconjunto aberto de $\mathcal{X}$.*

Dizemos que uma aplicação $T : \mathcal{X} \to \mathcal{Y}$ é uniformemente contínua se para todo $\epsilon > 0$, podemos encontrar um $\delta > 0$ tal que $\forall x, y \in \mathcal{X}$ com $d(x, y) < \delta$ implique $d(T(x), T(y)) < \epsilon$.

Uma topologia em um conjunto $\mathcal{X}$ é uma coleção Γ de subconjuntos de $\mathcal{X}$ chamados de abertos com as seguintes propriedades:

1. $\emptyset$ e $\mathcal{X}$ pertencem a $\mathit{\Gamma}$;

2. Se $A_1, \ldots, A_n \in \mathit{\Gamma}$, então $A_1 \cap \cdots \cap A_n \in \mathit{\Gamma}$;

3. Para qualquer família arbitrária de abertos $\{A_\lambda\}_{\lambda \in L}$ com $A_\lambda \in \mathit{\Gamma}$ para cada $\lambda \in L$, temos $\bigcup_{\lambda \in L} A_\lambda \in \mathit{\Gamma}$.

Um espaço topológico é um par $(\mathcal{X}, \mathit{\Gamma})$ onde $\mathcal{X}$ é um conjunto e $\mathit{\Gamma}$ uma topologia de $\mathcal{X}$. Observamos que todo espaço métrico $(\mathcal{X}, d)$ pode ser considerado um espaço topológico onde a topologia $\mathit{\Gamma}$ é a formada pelos subconjuntos abertos de $\mathcal{X}$.

Uma sequência $(x_n)_{n \in \mathbb{N}}$ definida em um conjunto $\mathcal{X}$ é uma aplicação $s : \mathbb{N} \to \mathcal{X}$ tal que o valor que a sequência assume em n é denotado por x_n. Dada $(x_n)_n$ uma sequência definida no espaço métrico $(\mathcal{X}, d)$, dizemos que o ponto $a \in \mathcal{X}$ é o limite da sequência $(x_n)_n$ se dado qualquer $\epsilon > 0$, podemos obter $n_0 \in \mathbb{N}$ tal que $n > n_0$ implica $d(x_n, a) < \epsilon$. Escrevemos então

$$\lim_{n \to \infty} x_n = a.$$

Definição 2.1.2. *Seja $(a_n)_{n \in \mathbb{N}}$ uma sequência de números reais. Então o limite superior e inferior de $(a_n)_n$ são definidos por*

$$\limsup a_n = \lim_{n \to \infty} \left[\sup \{a_k : k \geqslant n\}\right]$$

e

$$\liminf a_n = \lim_{n \to \infty} \left[\inf \{a_k : k \geqslant n\}\right]$$

respectivamente.

Para funções reais de uma variável real definimos:

$$\limsup_{x \to a} f(x) = \inf_{\delta > 0} \left\{ \sup_{0 < x - a < \delta} f(x) \right\}.$$

É possível mostrar que se $(x_n)_{n \in \mathbb{N}}$ é uma sequência tal que $x_n \neq a, \forall n \in \mathbb{N}$ e $\lim_{n \to \infty} x_n = a$, então

$$\limsup f(x_n) \leqslant \limsup_{x \to a} f(x).$$

Considerando $\mathcal{X}$ um espaço métrico compacto. Uma cobertura de $\mathcal{X}$ é uma família $\mathcal{C} = \{C_\lambda\}_{\lambda \in L}$ de subconjuntos de $\mathcal{X}$ tal que $\mathcal{X} = \bigcup_{\lambda \in L} C_\lambda$. Logo, para todo $x \in \mathcal{X}$ existe pelo menos um índice $\lambda \in L$ tal que $x \in C_\lambda$.

Se existe um subconjunto $L' \subset L$ tal que $\mathcal{X} \subseteq \bigcup_{\lambda \in L'} C_\lambda$, então dizemos que a subfamília $\mathcal{C}' = \{C_\lambda\}_{\lambda \in L'}$ é uma subcobertura de $\mathcal{C}$.

Um espaço métrico $(\mathcal{X}, d)$ é dito compacto quando toda cobertura aberta, ou seja, por abertos de $\mathcal{X}$, possui uma subcobertura finita.

2.2 Teoria da Medida

Definição 2.2.1. *Uma álgebra de $\mathcal{X}$ é uma família $\mathcal{B}$ de subconjunto de $\mathcal{X}$ que é fechada para as operações elementares de conjuntos e contém o conjunto vazio, ou seja,*

1. $\emptyset \in \mathcal{B}$;
2. *Se $A \in \mathcal{B}$, então $A^c \in \mathcal{B}$;*
3. *Se $A, B \in \mathcal{B}$, então $A \cup B \in \mathcal{B}$.*

Observamos que a união e a intersecção de qualquer número finito de elementos de $\mathcal{B}$ ainda pertence a $\mathcal{B}$.

Definição 2.2.2. *Uma álgebra é dita uma σ-álgebra de subconjuntos de $\mathcal{X}$ se também for fechada para as uniões enumeráveis, ou seja, se $A_i \in \mathcal{B}$ para todo $i \in \{1, 2, \ldots, n, \ldots\}$ implica $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$.*

Observamos que uma σ-álgebra também é fechada para intersecções enumeráveis.

Um espaço mensurável é uma dupla $(\mathcal{X}, \mathcal{B})$ onde $\mathcal{X}$ é um conjunto e $\mathcal{B}$ é uma σ-álgebra de subconjuntos de $\mathcal{X}$. Além disso, os elementos da σ-álgebra são chamados de conjuntos mensuráveis.

Proposição 2.2.3. *Considerando* $\{\mathcal{B}_i\}_{i\in\mathcal{I}}$ *uma família não vazia qualquer de* σ*-álgebras. Então* $\mathcal{B} = \bigcap_{i\in\mathcal{I}} \mathcal{B}_i$ *também é uma* σ*-álgebra.*

Definição 2.2.4. *A* σ*-álgebra gerada por uma família de subconjuntos de* $\mathcal{X}$ *é a menor* σ*-álgebra que contém esta família.*

Definição 2.2.5. *A* σ*-álgebra de Borel ou boreliana de um espaço topológico é a* σ*-álgebra gerada pela topologia, isto é, a menor* σ*-álgebra que contém todos os subconjuntos abertos de* $\mathcal{X}$*.*

Seja $(\mathcal{X}, \mathcal{B})$ um espaço mensurável. Uma medida em $(\mathcal{X}, \mathcal{B})$ é uma função $\mu : \mathcal{B} \to [0, +\infty]$ tal que $\mu(\emptyset) = 0$ e

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j)$$

para quaisquer conjuntos $A_j \in \mathcal{B}$ disjuntos dois a dois. A segunda propriedade é chamada de σ-aditividade.

A tripla $(\mathcal{X}, \mathcal{B}, \mu)$ é chamada espaço de medida. Além disso, quando temos $\mu(\mathcal{X}) < \infty$ dizemos que μ é finita e, se $\mu(\mathcal{X}) = 1$, dizemos que μ é uma probabilidade e $(\mathcal{X}, \mathcal{B}, \mu)$ é um espaço de probabilidade.

Teorema 2.2.6 (Extensão)**.** *Seja* $\mathcal{A}$ *uma álgebra de subconjuntos de* $\mathcal{X}$ *e seja* $\mu_0 : \mathcal{A} \to [0, +\infty]$ *uma função* σ*-aditiva com* $\mu(\mathcal{X}) < +\infty$*. Então existe uma única medida* μ *definida na* σ*-álgebra gerada por* $\mathcal{A}$ *que é uma extensão de* μ_0*, ou seja, tal que* $\mu(A) = \mu_0(A)$ *para todo elemento* $A \in \mathcal{A}$*.*

Definição 2.2.7. *Seja* $(\mathcal{X}, \Gamma)$ *um espaço topológico e* μ *uma medida definida na* σ*-álgebra de Borel de* $\mathcal{X}$*. O suporte de* μ *é o conjunto formado pelos pontos de* $\mathcal{X}$ *tais que* $\mu(B(x, \epsilon)) > 0$ *para todo* $\epsilon > 0$*.*

Definição 2.2.8. *Dados espaços mensuráveis* $(\mathcal{X}, \mathcal{B})$ *e* $(\mathcal{Y}, \mathcal{A})$*, dizemos que a aplicação* $T : \mathcal{X} \to \mathcal{Y}$ *é mensurável se* $T^{-1}(A) \in \mathcal{B}$ *para todo* $A \in \mathcal{A}$*.*

A próxima proposição mostra algumas propriedades básicas das funções mensuráveis.

Proposição 2.2.9. *Sejam $T, S : \mathcal{X} \to [-\infty, +\infty]$ funções mensuráveis e sejam $a, b \in \mathbb{R}$. Então são mensuráveis as funções*

$$(aT + bS)(x) = aT(x) + bS(x) \text{ e } (T \cdot S)(x) = T(x) \cdot S(x).$$

Além disso, também são mensuráveis

$$s(x) = \sup \{T_n(x) : n \geqslant 1\}\, ;\; i(x) = \inf \{T_n(x) : n \geqslant 1\}\, ,$$

$$T^*(x) = \limsup_{n\to\infty} T_n(x) \text{ e } T_*(x) = \liminf_{n\to\infty} T_n(x).$$

Se $T(x) = \lim_n T_n(x)$ existe, então T é mensurável.

Definição 2.2.10. *Dizemos que uma função $s : \mathcal{X} \to \mathbb{R}$ é simples se existem constantes $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ e conjuntos mensuráveis $A_1, \ldots, A_k$ disjuntos dois a dois, tais que*

$$s = \sum_{j=1}^{k} \alpha_j \mathcal{X}_{A_j},$$

onde $\mathcal{X}_{A_j}$ é a função característica do conjunto A_j dada por

$$\mathcal{X}_{A_j} = \begin{cases} 1, & \text{se } x \in A_j \\ 0, & \text{caso contrário} \end{cases}$$

Observamos que toda função simples é mensurável. Além disso, o próximo resultado afirma que toda função mensurável é limite de alguma sequência de funções simples.

Proposição 2.2.11. *Seja $T : \mathcal{X} \to [-\infty, +\infty]$ uma função mensurável. Então existe uma sequência $(s_n)_n$ de funções simples tal que $|s_n(x)| \leqslant |T(x)|$ para todo n e*

$$\lim_{n\to\infty} s_n(x) = T(x), \forall x \in \mathcal{X}.$$

Se T é não negativa, então podemos tomar $0 \leqslant s_1 \leqslant s_2 \leqslant \cdots \leqslant T$.

Agora, apresentaremos alguns resultados sobre integração. Inicialmente vamos definir a integral de funções simples em relação a uma medida e, posteriormente, estendê-la a funções mais gerais.

Definição 2.2.12. *Seja* $s = \sum_{j=1}^{k} \alpha_j \mathcal{X}_{A_j}$ *uma função simples. Então a integral de* s *em ralação à medida* μ *é dada por*

$$\int s \, d\mu = \sum_{j=1}^{k} \alpha_j \mu(A_j).$$

Para funções mensuráveis não negativas obtemos a seguinte definição:

Definição 2.2.13. *Seja* $T : \mathcal{X} \to [0, \infty]$ *uma função mensurável não negativa. Então a integral de* T *em relação à medida* μ *é dada por*

$$\int T \, d\mu = \lim_{n\to\infty} \int s_n \, d\mu,$$

onde $s_1 \leqslant s_2 \leqslant \cdots$ *é uma sequência de funções simples tal que*

$$\lim_{n\to\infty} s_n(x) = T(x), \forall x \in \mathcal{X}.$$

Observamos que dada uma função mensurável $T : \mathcal{X} \to [-\infty, +\infty]$ podemos escrever T da seguinte forma:

$$T = T^+ - T^-,$$

onde $T^+(x) = \max\{T(x), 0\}$ e $T^-(x) = \max\{-T(x), 0\}$.

É fácil ver que as funções T^+ e T^- são não negativas. Pela Proposição 2.2.9 obtemos também que estas funções são mensuráveis. Logo, podemos definir a integral de qualquer função mensurável.

Definição 2.2.14. *Seja* $T : \mathcal{X} \to [-\infty, +\infty]$ *uma função mensurável. Então a integral de* T *com respeito à medida* μ *é dada por*

$$\int T \, d\mu = \int T^+ \, d\mu - \int T^- \, d\mu,$$

desde que pelo menos uma das integrais do lado direito seja finita. Caso uma das integrais não seja finita, vale que $(+\infty) + a = +\infty$ *e* $a - (+\infty) = -\infty$ *para todo* $a \in \mathbb{R}$.

Dizemos que uma função $T : \mathcal{X} \to [-\infty, +\infty]$ é integrável se ela for mensurável e sua integral for um número real. Além disso, denotamos por $\mathcal{L}^1(\mathcal{X}, \mathcal{B}, \mu)$ o conjunto das funções integráveis.

Proposição 2.2.15. *O conjunto $\mathcal{L}^1(\mathcal{X}, \mathcal{B}, \mu)$ é um espaço vetorial real. Além disso, a aplicação $I : \mathcal{L}^1(\mathcal{X}, \mathcal{B}, \mu) \to \mathbb{R}$ dada por $I(T) = \int T \, d\mu$ é um funcional linear positivo, ou seja, para $a, b \in \mathbb{R}$ e T, S integráveis, obtemos*

1. $\int (aT + bS) d\mu = a \int T \, d\mu + b \int S \, d\mu$ *e*

2. $\int T \, d\mu \geqslant \int S \, d\mu$, *se* $T(x) \geqslant s(x)$, *para todo* $x \in \mathcal{X}$.

Definição 2.2.16. *Dizemos que uma propriedade é válida em μ quase todo ponto, abreviadamente, μ-q.t.p., se é válida em todo $\mathcal{X}$ exceto, possivelmente, em um conjunto de medida nula.*

Teorema 2.2.17 (Convergência monótona)**.** *Seja $T_n : \mathcal{X} \to [-\infty, +\infty]$ uma sequência monótona de funções mensuráveis não negativas e T dada por $T(x) = \lim_n T_n(x)$. Então*

$$\lim_{n\to\infty} \int T_n \, d\mu = \int T \, d\mu.$$

Teorema 2.2.18 (Convergência dominada)**.** *Seja $T_n : \mathcal{X} \to [-\infty, +\infty]$ uma sequência de funções mensuráveis e supondo que existe uma função integrável S tal que $|T_n(x)| \leqslant |S(x)|$ para μ-quase todo ponto $x \in \mathcal{X}$. Supondo também que a sequência $(T_n)_n$ converge em μ-quase todo ponto para uma função T. Então T é integrável e*

$$\lim_{n\to\infty} \int T_n \, d\mu = \int T \, d\mu.$$

Definição 2.2.19. *Seja $(\mathcal{X}, \mathcal{B}, \mu)$ um espaço de medida e $T : \mathcal{X} \to \mathcal{X}$ uma transformação mensurável. Dizemos que a medida μ é T-invariante se*

$$\mu(E) = \mu(T^{-1}(E))$$

para todo conjunto mensurável $E \subset \mathcal{X}$. Também dizemos que T preserva μ.

Lema 2.2.20. *Seja $T : \mathcal{X} \to \mathcal{X}$ uma transformação mensurável e μ uma medida finta em $\mathcal{X}$. Supondo que existe uma álgebra $\mathcal{A}$ de subconjuntos mensuráveis de $\mathcal{X}$ tal que esta álgebra gera a σ-álgebra e $\mu(E) = \mu(T^{-1}(E))$ para todo $E \in \mathcal{A}$. Então μ é uma medida T-invariante.*

Exemplo 2.2.21. *Considerando o deslocamento* $\sigma : \Sigma \to \Sigma$, *onde* $\Sigma = \{1, \dots, d\}^{\mathbb{N}} = \{(x_1, x_2, \dots) : x_i \in \{1, \dots, d\}\}$ *e* $\sigma(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots)$. *Seja* $\mu = \nu^{\mathbb{N}}$ *a medida produto , onde* $\nu(\{i\}) = p_i \geqslant 0, \forall i \in \{1, \dots, d\}$ *e* $p_1 + \cdots + p_d = 1$.

Denotamos por cilindro de tamanho n *o elemento da forma:*

$$[x_1, \dots, x_n] = \{z \in \Sigma : z = (z_1, z_2, \dots) \text{ e } z_1 = x_1, z_2 = x_2, \dots, z_n = x_n\}.$$

Observamos que a família das uniões finitas de cilindros disjuntos dois a dois é uma álgebra que gera a σ*-álgebra de Borel e a extensão da medida* μ *a esta* σ*-álgebra é a medida de Bernoulli.*

Além disso, a medida de Bernoulli é invariante pelo deslocamento.

De fato, dado $E = [x_1, \dots, x_n]$, *temos*

$$\begin{aligned}
\mu(\sigma^{-1}(E)) &= \mu\left(\bigcup_{i=1}^{d}[x_i, x_1, \dots, x_n]\right) \\
&= \sum_{i=1}^{d} \mu[x_i, x_1, \dots, x_n] \\
&= \sum_{i=1}^{d} p_i \cdot p_{x_1} \cdots p_{x_n} \\
&= p_{x_1} \cdots p_{x_n} \cdot \sum_{i=1}^{d} p_i \\
&= p_{x_1} \cdots p_{x_n} \\
&= \mu(E).
\end{aligned}$$

Pelo Lema 2.2.20, obtemos que μ *é invariante por* σ.

Proposição 2.2.22. *Seja* $T : \mathcal{X} \to \mathcal{X}$ *uma transformação mensurável e* μ *uma medida em* $\mathcal{X}$. *Então* μ *é* T*-invariante se, e somente se,*

$$\int \phi \, d\mu = \int \phi \circ T \, d\mu,$$

para toda função μ-integrável $\phi : \mathcal{X} \to \mathbb{R}$.

Teorema 2.2.23 (Riesz-Markov)**.** *Seja $\mathcal{X}$ um espaço métrico compacto e $\phi : C^0(\mathcal{X}) \to \mathbb{R}$ um funcional linear positivo, onde $C^0(\mathcal{X})$ representa o conjunto das funções contínuas de $\mathcal{X}$ em $\mathbb{R}$. Então existe uma única medida boreliana finita μ em $\mathcal{X}$ tal que*

$$\phi(\varphi) = \int \varphi \, d\mu,$$

para toda função $\varphi \in C^0(\mathcal{X})$.

Agora, vamos apresentar alguns resultados sobre existência de medidas invariantes em espaços métricos. Para isso, inicialmente vamos definir a topologia fraca* no conjunto das medidas borelianas de probabilidade em $\mathcal{X}$, denotado por $\mathcal{M}_1(\mathcal{X})$. Observamos que a partir de agora estamos considerando $\mathcal{X}$ espaço métrico.

Dada uma medida $\mu \in \mathcal{M}_1(\mathcal{X})$, um conjunto $\Phi = \{\phi_1, \ldots, \phi_N\}$ de funções contínuas limitadas $\phi_i : \mathcal{X} \to \mathbb{R}$ e $\epsilon > 0$, definimos

$$V(\mu, \Phi, \epsilon) = \left\{ \nu \in \mathcal{M}_1(\mathcal{X}) : \left| \int \phi \, d\nu - \int \phi \, d\mu \right| < \epsilon, \forall i \in \{1, \ldots, N\} \right\}.$$

A topologia fraca* é formada pelos conjuntos $\mathcal{A} \subset \mathcal{M}_1(\mathcal{X})$ tais que para todo elemento $\mu \in \mathcal{A}$ existe algum $V(\mu, \Phi, \epsilon) \subset \mathcal{A}$.

Lema 2.2.24. *Uma sequência de medidas $(\mu_n)_{n\in\mathbb{N}}$ converge para uma medida $\mu \in \mathcal{M}_1(\mathcal{X})$ na topologia fraca* se, e somente se,*

$$\lim_{n\to\infty} \int \phi \, d\mu_n = \int \phi \, d\mu,$$

para toda função contínua limitada $\phi : \mathcal{X} \to \mathbb{R}$.

Uma variação da definição da topologia fraca* é obtida considerando $\mathcal{A} = \{A_1, \ldots, A_N\}$ uma família de abertos de $\mathcal{X}$, $\epsilon > 0$ e

$$V_a(\mu, \mathcal{A}, \epsilon) = \{\nu \in \mathcal{M}_1(\mathcal{X}) : \nu(A_i) > \mu(A_i) - \epsilon, \forall i \in \{1, \ldots, N\}\}.$$

Teorema 2.2.25. *As topologias definidas pelas bases de vizinhanças $V(\mu, \Phi, \epsilon)$ e $V_a(\mu, \mathcal{A}, \epsilon)$ são equivalentes.*

Agora, supondo $(\mathcal{X}, d)$ um espaço métrico compacto é possível mostrar que $\mathcal{M}_1(\mathcal{X})$ munido da topologia fraca* é compacto. Logo, obtemos que toda sequência $(\mu_n)_{n\in\mathbb{N}}$ em $\mathcal{M}_1(\mathcal{X})$ admite alguma subsequência convergente na topologia fraca*.

Seja $T : \mathcal{X} \to \mathcal{X}$ mensurável e qualquer medida μ em $\mathcal{X}$. Para todo conjunto mensurável E definimos $T_*\mu$ como $T_*\mu(E) = \mu(T^{-1}(E))$. Então, μ é T-invariante se $T_*\mu = \mu$.

Lema 2.2.26. *Seja μ uma medida e ϕ uma função mensurável limitada. Então*

$$\int \phi \, dT_*\mu = \int \phi \circ T \, d\mu.$$

Lema 2.2.27. *Todo ponto de acumulação de uma sequência $(\mu_n)_{n\in\mathbb{N}}$ onde*

$$\mu_n = \frac{1}{n}\sum_{j=0}^{n-1} T_*^j \nu$$

e ν é qualquer medida de probabilidade é uma probabilidade invariante por T.

Teorema 2.2.28. *Seja $T : \mathcal{X} \to \mathcal{X}$ uma transformação contínua num espaço métrico compacto. Então existe pelo menos uma medida de probabilidade em $\mathcal{X}$ que é invariante por T.*

Observamos que a continuidade de T e a compacidade de $\mathcal{X}$ são importantes para garantirmos a existência de uma medida invariante, pois sem uma destas hipóteses é possível obtermos exemplos onde este Teorema não vale. Para mais detalhes sugerimos [9].

2.3 Teoria Ergódica

Em poucas palavras, a Teoria Ergódica, cujas origens remontam a mecânica estatística, estuda o comportamento de sistemas dinâmicos em relação a medidas que são invariantes pela ação da dinâmica.

Nesta seção vamos apresentar alguns resultados importantes para o presente trabalho, porém para mais detalhes sobre este assunto sugerimos [9], [12] e [13].

Teorema 2.3.1 (Birkhoff). *Seja $T : \mathcal{X} \to \mathcal{X}$ uma transformação mensurável e μ uma medida de probabilidade invariante por T. Então para qualquer função integrável $\varphi : \mathcal{X} \to \mathbb{R}$ o limite*

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

existe em μ-quase todo ponto $x \in \mathcal{X}$. Além disso, a função $\tilde{\varphi}$ assim definida é integrável e satisfaz

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

Definição 2.3.2. *Considerando $(\mathcal{X}, \mathcal{B}, \mu)$ um espaço de probabilidade e $T : \mathcal{X} \to \mathcal{X}$ uma transformação que preserva a medida μ. Um conjunto $A \subset \mathcal{X}$ é dito invariante se $A = T^{-1}(A)$ em μ-quase todo ponto.*

Definição 2.3.3. *Uma medida de probabilidade T-invariante é dita ergódica se todo conjunto invariante $A \subset \mathcal{X}$ satisfaz $\mu(A) = 0$ ou $\mu(A) = 1$.*

Exemplo 2.3.4. *A medida de Bernoulli no espaço $\Sigma = \{1, \ldots, d\}^{\mathbb{N}}$ considerando o deslocamento $\sigma : \Sigma \to \Sigma$ é uma medida ergódica.*

Denotamos por $\mathcal{M}_T(\mathcal{X})$ o conjunto das medidas de probabilidade invariantes pela aplicação T e observamos que se μ_1 e μ_2 pertencem a $\mathcal{M}_T(\mathcal{X})$, então $(1-t)\mu_1 + t\mu_2$ também é uma medida de probabilidade T-invariante, qualquer que seja $t \in [0, 1]$, ou seja, $\mathcal{M}_T(\mathcal{X})$ é um conjunto convexo. A próxima proposição afirma que as medidas ergódicas são os elementos extremais do conjunto $\mathcal{M}_T(\mathcal{X})$.

Proposição 2.3.5. *Uma medida de probabilidade μ é ergódica se, e somente se, não é possível escrevermos $\mu = (1-t)\mu_1 + t\mu_2$ com $t \in (0, 1)$ e μ_1,μ_2 medidas de probabilidade invariantes distintas.*

O próximo teorema afirma que em um espaço métrico compacto $\mathcal{X}$, podemos decompor qualquer medida invariante como combinação convexa de medidas ergódicas, valendo também para espaços métricos mais gerais um resultado semelhante.

Teorema 2.3.6 (Decomposição ergódica)**.** *Seja $\mathcal{X}$ um espaço métrico compacto, $T : \mathcal{X} \to \mathcal{X}$ uma transformação mensurável e μ uma medida de probabilidade T-invariante. Então existe um conjunto $\mathcal{X}_0 \subset \mathcal{X}$ com $\mu(\mathcal{X}_0) = 1$, uma partição $\mathcal{P}$ de $\mathcal{X}_0$ em conjuntos mensuráveis e uma família de probabilidades ergódicas da forma*

$$\mu_x = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T^j(x)},$$

tal que para $E \subset \mathcal{X}$ mensurável

$$\mu(E) = \int_{\mathcal{X}} \mu_x(E)\, d\mu.$$

Agora, vamos apresentar a noção de equivalência ergódica entre dois sistemas, que é uma noção fundamental para decidir se dois sistemas podem ser considerados o "mesmo"sob o ponto de vista da Teoria Ergódica.

Sejam μ e ν duas medidas de probabilidade invariantes por $T : \mathcal{X} \to \mathcal{X}$ e $S : \mathcal{Y} \to \mathcal{Y}$, respectivamente. Dizemos que os sistemas (T, μ) e (S, ν) são ergodicamente equivalentes se existem $X \subset \mathcal{X}$ e $Y \subset \mathcal{Y}$ com $\mu(X) = 1$, $\nu(Y) = 1$ e uma função bijetiva $\phi : X \to Y$ mensurável com inversa mensurável tal que

$$\phi \circ T = S \circ \phi$$

e

$$\phi_* \mu = \nu.$$

Capítulo 3

Entropia

Neste capítulo vamos definir a entropia métrica e a entropia topológica de um sistema dinâmico, bem como alguns resultados básicos sobre cada uma delas. Apesar do conceito de entropia ser algo conhecido em Dinâmica, muitos dos resultados e demonstrações deste trabalho tem inspiração nos resultados deste capítulo. Além disso, no capítulo seguinte será provado o Princípio Variacional que relaciona estas duas noções de entropia. Este resultado nos diz que ao considerarmos $\mathcal{X}$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ uma transformação contínua, então a entropia topológica é o supremo da entropia métrica onde o supremo é tomado sobre todas as medidas de probabilidade invariantes por T.

3.1 Entropia Métrica

Antes de definirmos formalmente a entropia métrica, vamos apresentar um exemplo que motiva a definição da função informação.

Como exemplo, consideramos o conjunto $\mathcal{X} = \{1, 2, \ldots, 6\}$ e $\mathbb{P}$ a medida de probabilidade que dá pesos $1/6$ a cada elemento de $\mathcal{X}$. Ao considerarmos $A = \{1, 2, \ldots, 6\}$ e $B = \{1, 2\}$, obtemos $\mathbb{P}(A) = 1$ e $\mathbb{P}(B) = 1/3$. Logo, se quisermos definir uma função I que a cada elemento do conjunto $\mathcal{X}$ associamos um número e este número representaria de certa forma o "valor"da informação, é razoável que $\mathbb{P}(A) = 1$ implique $I(A) = 0$ e $\mathbb{P}(A) > \mathbb{P}(B)$ implique $I(A) < I(B)$. Intuitivamente, poderíamos pensar em um lançamento de um dado de seis faces, se alguém afirmasse que ao jogar o dado o conjunto de resultados possíveis é A, ou seja, pode sair qualquer face, não é uma informação tão relevante quanto se alguém afirmasse que os resultados seriam somente $\{1, 2\}$.

Além disso, como $\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A/B)$ também é de se esperar que $I(A \cap B) = I(B) + I(A/B)$.

Com isso, obtemos que $I(A) = -\log \mathbb{P}(A)$.

Agora, ao considerar $(\mathcal{X}, \mathcal{B}, \mu)$ um espaço de probabilidade, podemos associar a cada partição $\mathcal{P}$ de conjuntos mensuráveis a respectiva função informação, que é definida por

$$I_{\mathcal{P}}(x) = -\log \mu(\mathcal{P}(x)),$$

onde $\mathcal{P}(x)$ representa o elemento da partição que contém x.

Então, isto nos motiva a seguinte definição.

Definição 3.1.1 (Entropia de uma partição). *Seja $\mathcal{P}$ uma partição finita ou enumerável de conjuntos mensuráveis e μ uma medida de probabilidade. Então, chamamos de entropia ou informação média da partição $\mathcal{P}$ o número*

$$H_{\mu}(\mathcal{P}) = \int I_{\mathcal{P}}\, d\mu = \sum_{P \in \mathcal{P}} -\mu(P) \log \mu(P).$$

Além disso, por convenção, assumimos que $0 \log 0 = \lim_{x \to 0} x \log x = 0$.

Exemplo 3.1.2. *Se $\mathcal{X} = \{1, 2\}$, μ é a medida de probabilidade que dá pesos $1/2$ a cada elemento de $\mathcal{X}$ e $\mathcal{P} = \{\{1\}, \{2\}\}$. Então*

$$H_{\mu}(\mathcal{P}) = -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} = \log 2.$$

Exemplo 3.1.3. *Se $\mathcal{X} = \{1, 2\}$, μ é a medida de probabilidade que dá pesos $1/3$ e $2/3$, respectivamente a 1 e 2. Então*

$$H_{\mu}(\mathcal{P}) = -\frac{1}{3} \log \frac{1}{3} - \frac{2}{3} \log \frac{2}{3} = \log 3 - \log 2 \cdot \frac{2}{3}.$$

Observamos que a entropia do primeiro exemplo é maior que a do segundo. Intuitivamente, isto significa que é mais difícil realizarmos previsões no primeiro caso do que o segundo. De fato, a chance de obtermos 1 ou 2 no primeiro caso são iguais, já no segundo, temos mais chance de que ocorra 2.

É possível mostrar que toda partição finita tem entropia finita e $H_\mu(\mathcal{P}) \leqslant \log|\mathcal{P}|$, onde $|\mathcal{P}|$ representa a *cardinalidade* do conjunto $\mathcal{P}$, valendo a igualdade quando $\mu(P) = 1/|\mathcal{P}|$ para todo elemento de $\mathcal{P}$.

Exemplo 3.1.4. *Seja $\mathcal{X} = \{1, \ldots, d\}^{\mathbb{N}}$, μ a medida de Bernoulli que dá peso $1/d$ a cada elemento de $\{1, \ldots, d\}$ e $\mathcal{P}$ a partição dada pelos cilindros de tamanho n. Então*

$$H_\mu(\mathcal{P}) = -\sum_{P \in \mathcal{P}} \left(\frac{1}{d}\right)^n \log \left(\frac{1}{d}\right)^n = n \log d,$$

pois esta partição contém d^n elementos e a medida de cada um deles é $1/d^n$.

Denotamos por $\mathcal{P} \vee \mathcal{Q} = \{P \cap Q : P \in \mathcal{P}, Q \in \mathcal{Q}\}$ o *refinamento* das partições $\mathcal{P}$ e $\mathcal{Q}$. Dizemos que $\mathcal{P}$ é menos fina que $\mathcal{Q}$, em notação $\mathcal{P} \prec \mathcal{Q}$, quando para todo elemento $Q \subset \mathcal{Q}$ existe um elemento $P \subset \mathcal{P}$ tal que $Q \subset P$ μ-q.t.p.

Definição 3.1.5 (Entropia Condicional)**.** *Chamamos de entropia condicional de $\mathcal{P}$ em relação à $\mathcal{Q}$ o número*

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

Intuitivamente, a entropia condicional mede a informação adicional fornecida pela partição $\mathcal{P}$ dado o conhecimento de $\mathcal{Q}$. Dizemos que duas partições $\mathcal{P}$ e $\mathcal{Q}$ são *independentes* se $\mu(P \cap Q) = \mu(P) \cdot \mu(Q)$ para todo $P \in \mathcal{P}$ e $Q \in \mathcal{Q}$.

Logo

$$I_{\mathcal{P} \vee \mathcal{Q}}(x) = -\log \mu((P \cap Q)(x)) = -\log \mu\left(P(x)Q(x)\right) = I_{\mathcal{P}}(x) + I_{\mathcal{Q}}(x)$$

e

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

Além disso, é fácil ver que quando $\mathcal{P}$ e $\mathcal{Q}$ são independentes, obtemos

$$H_\mu(\mathcal{P}/\mathcal{Q}) = H_\mu(\mathcal{P})$$

Afirmamos que $H_\mu(\mathcal{P}/\mathcal{Q}) = H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q})$.

De fato,

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

Também é possível mostrar que $H_\mu(\mathcal{P}/\mathcal{Q}) \leqslant H_\mu(\mathcal{P})$.

Consequentemente

$$H_\mu(\mathcal{P} \vee \mathcal{Q}) \leqslant H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q}).$$

Lema 3.1.6. *Sejam $\mathcal{P}$ e $\mathcal{Q}$ partições mensuráveis com entropia finita tais que $\mathcal{P} \prec \mathcal{Q}$, então*

$$H_\mu(\mathcal{P}) \leqslant H_\mu(\mathcal{Q}).$$

Demonstração. Sejam $\mathcal{P}$ e $\mathcal{Q}$ partições mensuráveis com entropia finita tais que $\mathcal{P} \prec \mathcal{Q}$, ou seja, todo elemento de $\mathcal{Q}$ está contido em algum elemento de $\mathcal{P}$ a menos de medida nula.

Logo

$$
\begin{aligned}
H_\mu(\mathcal{P}) &= -\sum_{P\in\mathcal{P}} \mu(P)\log\mu(P) \\
&= -\sum_{P}\sum_{Q\subset P} \mu(P\cap Q)\log\mu(P) \\
&\leqslant -\sum_{P}\sum_{Q\subset P} \mu(Q)\log\mu(Q) \\
&= -\sum_{Q} \mu(Q)\log\mu(Q) \\
&= H_\mu(\mathcal{Q}).
\end{aligned}
$$

□

Além disso, também é possível mostrar que se $\mathcal{P} \prec \mathcal{Q}$, então

$$H_\mu(\mathcal{P}/\mathcal{Q}) = 0.$$

Observamos que até agora não utilizamos a dinâmica no cálculo da entropia. Porém, para fazermos isso, consideramos $T : \mathcal{X} \to \mathcal{Y}$ uma aplicação mensurável, μ uma medida de probabilidade definida em $\mathcal{X}$ e $\mathcal{P}$ uma partição de $\mathcal{Y}$. Então, podemos definir uma medida de probabilidade ν em $\mathcal{Y}$ e uma partição $\mathcal{Q}$ em $\mathcal{X}$ escolhendo

$$\nu = T_*\mu \text{ e } \mathcal{Q} = T^{-1}(\mathcal{P}).$$

Proposição 3.1.7. *Seja $T : \mathcal{X} \to \mathcal{Y}$ uma aplicação mensurável, μ uma medida de probabilidade definida em $\mathcal{X}$ e $\mathcal{P}$ uma partição de $\mathcal{Y}$. Então*

$$H_{T_*\mu}(\mathcal{P}) = H_\mu(T^{-1}(\mathcal{P}))$$

Demonstração.

$$
\begin{aligned}
H_{T_*\mu}(\mathcal{P}) &= -\sum_{P\in\mathcal{P}} \nu(P)\log\nu(P) \\
&= -\sum_{P\in\mathcal{P}} \mu(T^{-1}(P))\log\mu(T^{-1}(P)) \\
&= H_\mu(T^{-1}(\mathcal{P})).
\end{aligned}
$$

□

Se considerarmos $T : \mathcal{X} \to \mathcal{X}$ mensurável, μ uma medida de probabilidade T-invariante e $\mathcal{P}$ uma partição de $\mathcal{X}$, podemos mostrar que

$$H_\mu(\mathcal{P}) = H_\mu(T^{-1}(\mathcal{P})).$$

A partir de agora, ao longo desta seção, vamos sempre considerar T uma transformação que preserva a medida μ.

Dada uma partição $\mathcal{P}$ com entropia finita, denotamos

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

Antes de definirmos a entropia métrica de um sistema dinâmico vamos defini-la com respeito a uma partição. Para isso, precisamos do seguinte lema técnico:

Lema 3.1.8. *Seja $(a_n)_{n\in\mathbb{N}}$ uma sequência de números positivos subaditiva, ou seja, uma sequência tal que $a_{m+n} \leqslant a_n + a_m, \forall m, n \geqslant 1$. Então*

$$\lim_n \frac{a_n}{n} = \inf_n \frac{a_n}{n}.$$

Demonstração. Fixado $p \in \mathbb{N}$, para todo $n \in \mathbb{N}$ tal que $n > p$ podemos escrever $n = pq + r$, onde $0 \leqslant r \leqslant p - 1$. Então

$$a_n = a_{pq+r} \leqslant a_{pq} + a_r,$$

mas $a_{pq} \leqslant a_p + \cdots + a_p \leqslant qa_p$.

Logo

$$\frac{a_n}{n} \leqslant \frac{qa_p}{n} + \frac{a_r}{n}.$$

Como $n \geqslant pq$, obtemos

$$\frac{a_n}{n} \leqslant \frac{qa_p}{pq} + \frac{a_r}{n} = \frac{a_p}{p} + \frac{a_r}{n}.$$

Como p está fixo, tomando o limite superior em ambos os lados, temos

$$\limsup_{n\to\infty} \frac{a_n}{n} \leqslant \limsup_{n\to\infty} \frac{a_p}{p} + \limsup_{n\to\infty} \frac{a_r}{n} = \frac{a_p}{p},$$

Então, como p é qualquer número natural, isto implica que

$$\limsup_{n\to\infty} \frac{a_n}{n} \leqslant \inf_{p\in\mathbb{N}} \frac{a_p}{p} \leqslant \liminf_{n\to\infty} \frac{a_n}{n}.$$

Portanto, sabendo que $\liminf_{n\to\infty} \frac{a_n}{n} \leqslant \limsup_{n\to\infty} \frac{a_n}{n}$, obtemos

$$\lim_n \frac{a_n}{n} = \inf_n \frac{a_n}{n}.$$

□

Lema 3.1.9. *Seja $\mathcal{P}$ uma partição com entropia finita e μ uma medida de probabilidade invariante por $T : \mathcal{X} \to \mathcal{X}$. Então*

$$H_\mu(\mathcal{P}^{m+n}) \leqslant H_\mu(\mathcal{P}^n) + H_\mu(\mathcal{P}^m), \forall m, n \geqslant 1,$$

ou seja, a sequência $H_\mu(\mathcal{P}^n)$ é subaditiva.

Demonstração. Como $\mathcal{P}^{m+n} = \mathcal{P}^m \vee T^{-m}(\mathcal{P}^n)$, temos que

$$\begin{aligned} H_\mu(\mathcal{P}^{m+n}) &= H_\mu(\mathcal{P}^m \vee T^{-m}(\mathcal{P}^n)) \\ &\leqslant H_\mu(\mathcal{P}^m) + H_\mu(T^{-m}(\mathcal{P}^n)) \\ &= H_\mu(\mathcal{P}^m) + H_\mu(\mathcal{P}^n). \end{aligned}$$

□

Logo, definimos a entropia de T com respeito à medida μ e à partição $\mathcal{P}$ como o limite

$$h_\mu(T, \mathcal{P}) = \lim_n \frac{1}{n} H_\mu(\mathcal{P}^n) = \inf_n \frac{1}{n} H_\mu(\mathcal{P}^n).$$

A entropia do sistema (T, μ) é definida como

$$h_\mu(T) = \sup_{\mathcal{P}} h_\mu(T, \mathcal{P}).$$

onde o supremo é tomado sobre todas as partições de $\mathcal{X}$ com entropia finita.

No exemplo 3.1.4 calculamos a entropia sobre a partição dos cilindros de tamanho n com respeito a medida de Bernoulli que dá peso $1/d$ a cada elemento de $\{1, \ldots, d\}$, porém, agora vamos considerar um caso mais geral. Considerando $T = \sigma$, onde σ representa o deslocamento e $\mathcal{P}$ a partição dos cilindros de tamanho 1, é possível mostrar que $\mathcal{P}^n$ é a partição dos cilindros de tamanho n.

Então, considerando $p_i = \mu([i]), \forall i \in \{1, \ldots, d\}$, obtemos

$$\begin{aligned} h_\mu(T, \mathcal{P}) &= \lim_{n \to \infty} \frac{1}{n} H_\mu(\mathcal{P}^n) \\ &= -\lim_{n \to \infty} \frac{1}{n} \cdot n \sum_{i=1}^{d} p_i \log p_i \\ &= -\sum_{i=1}^{d} p_i \log p_i \end{aligned}$$

Para mais detalhes sobre entropia métrica sugerimos a leitura de [9], [11] e [12], porém, vamos citar alguns resultados importantes.

Teorema 3.1.10 (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 $\bigcup_{n=1}^{\infty} \mathcal{P}_n$ gera a σ-álgebra dos conjuntos mensuráveis a menos de medida nula. Então*

$$h_\mu(T) = \lim_{n \to \infty} h_\mu(T, \mathcal{P}_n).$$

Corolário 3.1.11. *Seja $\mathcal{P}$ uma partição com entropia finita tal que $\mathcal{P}^n$ gera a σ-álgebra dos conjuntos mensuráveis. Então*

$$h_\mu(T) = h_\mu(T, \mathcal{P}).$$

Com isso, temos que

$$h_\mu(\sigma) = -\sum_{i=1}^{d} p_i \log p_i,$$

onde μ é a medida de Bernoulli.

Além disso, também é possível mostrar que se $(\mathcal{X}, T, \mu)$ e $(\mathcal{Y}, S, \nu)$ são ergodicamente equivalentes, então

$$h_\mu(T) = h_\nu(S).$$

3.2 Entropia Topológica

A entropia topológica é um invariante topológico que mede a taxa de crescimento exponencial de um número de diferentes segmentos de órbita de comprimento n. Além disso, pode ser utilizada para afirmar se dois sistemas dinâmicos podem ser topologicamente conjugados ou não.

Definição 3.2.1. *Seja $T : \mathcal{X} \to \mathcal{X}$ uma aplicação contínua e $(\mathcal{X}, d)$ um espaço métrico compacto. Então, $\forall n \in \mathbb{N}$ e $x, y \in \mathcal{X}$, definimos a métrica*

$$d_n(x, y) = \max_{0 \leqslant i < n} \left\{ d(T^i(x), T^i(y)) \right\}$$

que mede a distância máxima entre os primeiros n iterados dos pontos x e y. Além disso, note que $d_n \geqslant d_{n-1} \geqslant \cdots \geqslant d_1 = d$. Denotamos a bola de raio ϵ na métrica d_n como uma (n, ϵ)-bola dinâmica.

Teorema 3.2.2. *Todas as métricas d_n são topologicamente equivalentes.*

Demonstração. Como $d_n(x, y) < \epsilon$ implica que $d(x, y) < \epsilon$, então $\forall y \in B_n(x, \epsilon)$, onde $B_n(x, \epsilon) = \{y \in \mathcal{X} : d_n(x, y) < \epsilon\}$, obtemos que $y \in B(x, \epsilon)$, logo

$$B_n(x, \epsilon) \subseteq B(x, \epsilon).$$

Por outro lado, se $y \in B_n(x, \epsilon)$ então

$$d(T^i(x), T^i(y)) < \epsilon, \forall i \in \{0, 1, \ldots, n-1\}.$$

Como T^i é contínua $\forall i \in \mathbb{N}$, isto implica que T^i é uniformemente contínua em $\mathcal{X}$, pois $\mathcal{X}$ é compacto.

Então, $\forall i \in \{0, 1, \ldots, n-1\}$, existe $\epsilon_i > 0$ tal que $d(x,y) < \epsilon_i$ implica que $d(T^i(x), T^i(y)) < \epsilon$. Escolhendo $\overline{\epsilon} = \min\{\epsilon_i\}$, obtemos

$$B(x, \epsilon) \subseteq B_n(x, \epsilon).$$

Portanto, mostramos que $\forall n \in \mathbb{N}$ as métricas d e d_n são equivalentes, por transitividade, concluímos que todas as métricas d_n são duas a duas equivalentes, ou seja, induzem a mesma topologia em $(\mathcal{X}, T)$. $\square$

Agora, vamos definir alguns conceitos importantes que são utilizados na definição de entropia topológica.

Considerando $n \in \mathbb{N}$, $\epsilon > 0$, $T : \mathcal{X} \to \mathcal{X}$ contínua e $(\mathcal{X}, d)$ espaço métrico compacto, dizemos que um conjunto $A \subset \mathcal{X}$ é um (n, ϵ)-gerador se para todo $x \in \mathcal{X}$, existe $a \in A$ tal que $d_n(x, a) < \epsilon$. Denotamos por $G_d(n, \epsilon) = \min\{|A| : A \text{ é } (n, \epsilon)\text{-gerador}\}$.

Dizemos que um conjunto $B \subset \mathcal{X}$ é (n, ϵ)-separado se $d_n(x, y) \geqslant \epsilon, \forall x, y \in B$. Denotamos $N_d(n, \epsilon) = \max\{|B| : B \text{ é } (n, \epsilon)\text{-separado}\}$.

Por último, denotamos por $C_d(n, \epsilon)$ o número mínimo de elementos de uma cobertura de $\mathcal{X}$ cujo diâmetro dos elementos desta cobertura é menor do que ϵ com respeito a métrica d_n.

Como $\mathcal{X}$ é um espaço métrico compacto, os números $C_d(n, \epsilon)$, $G_d(n, \epsilon)$ e $N_d(n, \epsilon)$ são finitos. Além disso, estes números contam o número de segmentos de órbitas de comprimento n distinguíveis em uma escala ϵ.

Lema 3.2.3. *Para todo $n \in \mathbb{N}$, $\epsilon > 0$, T contínua e $(\mathcal{X}, d)$ compacto, temos*

$$C_d(n, 2\epsilon) \leqslant G_d(n, \epsilon) \leqslant N_d(n, \epsilon) \leqslant C_d(n, \epsilon).$$

Demonstração. Para provar a primeira desigualdade, consideramos $A \subset \mathcal{X}$ um conjunto (n, ϵ)-gerador tal que $|A| = G_d(n, \epsilon)$. Um conjunto ser (n, ϵ)-gerador implica que para todo $x \in \mathcal{X}$, existe um elemento $a \in A$ tal que $d_n(x, a) < \epsilon$. Logo

$$\mathcal{X} = \bigcup_{a \in A} B_n(a, \epsilon).$$

Com isso, obtemos uma cobertura por abertos de $\mathcal{X}$ tal que o diâmetro com respeito a métrica d_n é menor ou igual a 2ϵ e

$$C_d(n, 2\epsilon) \leqslant G_d(n, \epsilon).$$

Para a segunda desigualdade, consideramos $B \subset \mathcal{X}$ um conjunto (n, ϵ)-separado tal que $|B| = N_d(n, \epsilon)$. Então, isto implica que B também é um conjunto (n, ϵ)-gerador.

De fato, se existisse um elemento $x \in \mathcal{X}$ tal que $d_n(x, b) \geqslant \epsilon$ para todo $b \in B$, teríamos que $B \cup \{x\}$ seria um conjunto (n, ϵ)-separado, contrariando a maximalidade de B. Logo, isto implica que $\forall x \in \mathcal{X}$, $\exists b \in B$ tal que $d_n(x, b) < \epsilon$. Consequentemente

$$G_d(n, \epsilon) \leqslant N_d(n, \epsilon).$$

Por último, seja $\mathcal{C}$ uma cobertura de $\mathcal{X}$ tal que $|\mathcal{C}|= C_d(n, \epsilon)$ e o diâmetro dos elementos da cobertura é menor do que ϵ. Logo, $\forall C \in \mathcal{C}$, vai existir no máximo um elemento $b \in B$ em C, onde B é um conjunto (n, ϵ)-separado. Então, considerando $|B| = N_d(n, \epsilon)$, obtemos

$$N_d(n, \epsilon) \leqslant C_d(n, \epsilon).$$

$\square$

Observamos que se $\epsilon_1 < \epsilon_2$, então $C_d(n, \epsilon_1) \geqslant C_d(n, \epsilon_2)$. De fato, pois toda cobertura de $\mathcal{X}$ com diâmetro menor do que ϵ_1 têm diâmetro menor que ϵ_2, em particular, a cobertura com menor número de elementos.

Seja

$$h_\epsilon(T) = \limsup_{n\to\infty} \frac{1}{n} \log C_d(n, \epsilon).$$

Pela observação anterior, é fácil ver que se $\epsilon_1 < \epsilon_2$, então

$$h_{\epsilon_1}(T) \geqslant h_{\epsilon_2}(T).$$

Portando, obtemos que $h_\epsilon(T)$ é uma função monótona não-decrescente quando $\epsilon \to 0$.

Definição 3.2.4 (Entropia Topológica)**.** *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ contínua. Então a entropia topológica de T é definida por*

$$h_{top}(T) = \lim_{\epsilon \to 0} h_\epsilon(T).$$

Pelo lema anterior obtemos

$$\begin{aligned} h(T) &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log C_d(n,\epsilon) \\ &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log G_d(n,\epsilon) \\ &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_d(n,\epsilon). \end{aligned}$$

Teorema 3.2.5. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ contínua. Fixado $\epsilon > 0$, então*

$$\lim_{n\to\infty} \frac{1}{n} \log C_d(n,\epsilon) = h_\epsilon(T).$$

Demonstração. Sejam $\mathcal{A}$ e $\mathcal{B}$ duas coberturas de $\mathcal{X}$ tais que o diâmetro dos elementos de $\mathcal{A}$ e $\mathcal{B}$ é menor do que ϵ com respeito as métricas d_m e d_n, respectivamente. Além disso, consideramos $\mathcal{A}$ e $\mathcal{B}$ tais que $|\mathcal{A}| = C_d(m,\epsilon)$ e $|\mathcal{B}| = C_d(n,\epsilon)$.

Para construir uma cobertura com diâmetro em relação a d_{n+m} menor do que ϵ, basta considerar

$$\mathcal{C} = \left\{ A \cap T^{-m}(B) : A \in \mathcal{A}, B \in \mathcal{B} \right\}.$$

De fato, como $\mathcal{B}$ é uma cobertura de $\mathcal{X}$, pela continuidade de T temos que $T^{-1}(\mathcal{B})$ também é uma cobertura de $\mathcal{X}$. Logo, sucessivamente, obtemos que $T^{-m}(\mathcal{B})$ é uma cobertura e, consequentemente, a intersecção dos elementos de $\mathcal{A}$ e $T^{-m}(\mathcal{B})$ também cobrem $\mathcal{X}$.

Agora, consideramos $C \in \mathcal{C}$, então $\forall x, y \in C$

$$
\begin{aligned}
d_{n+m}(x,y) &= \max_{0 \leqslant i \leqslant n+m-1} \left\{ d(T^i(x), T^i(y)) \right\} \\
&= \max \left\{ d_m(x,y), d_n(T^m(x), T^m(y)) \right\}
\end{aligned}
$$

Como $x, y \in A$ para algum $A \in \mathcal{A}$, temos que $d_m(x,y) < \epsilon$. Além disso, como $T^m(x), T^m(y) \in B$ para algum $B \in \mathcal{B}$, obtemos também que $d_n(T^m(x), T^m(y)) < \epsilon$. Logo $d_{n+m}(x,y) < \epsilon, \forall x, y \in C$. Como C é qualquer elemento de $\mathcal{C}$, concluímos que o diâmetro dos elementos de $\mathcal{C}$ com respeito a métrica d_{n+m} é menor do que ϵ. Então

$$C_d(n+m, \epsilon) \leqslant |\mathcal{A}| \cdot |\mathcal{B}| = C_d(m, \epsilon) \cdot C_d(n, \epsilon).$$

Consequentemente

$$
\begin{aligned}
\log C_d(n+m, \epsilon) &\leqslant \log \left(C_d(m, \epsilon) \cdot C_d(n, \epsilon) \right) \\
&= \log C_d(m, \epsilon) + \log C_d(n, \epsilon).
\end{aligned}
$$

Portando, isso mostra que $\log C_d(n, \epsilon)$ é uma sequência subaditiva e

$$\lim_{n \to \infty} \frac{1}{n} \log C_d(n, \epsilon)$$

existe.

Pela definição de $h_\epsilon(T)$ concluímos que

$$\lim_{n \to \infty} \frac{1}{n} \log C_d(n, \epsilon) = h_\epsilon(T).$$

□

Por meio deste teorema podemos escrever a entropia topológica em termos do limite inferior, ou seja

Definição 3.2.6. *Seja* $(\mathcal{X}, d)$ *um espaço métrico compacto e* $T : \mathcal{X} \to \mathcal{X}$ *contínua. Então a entropia topológica de* T *é definida por*

$$\begin{aligned}
h(T) &= \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log C_d(n, \epsilon) \\
&= \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log G_d(n, \epsilon) \\
&= \lim_{\epsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} \log N_d(n, \epsilon).
\end{aligned}$$

Exemplo 3.2.7. *Considerando o deslocamento* $\sigma : \{1,2\}^{\mathbb{N}} \to \{1,2\}^{\mathbb{N}}$*, a entropia topológica de* σ *é* $\log 2$*.*

De fato, considerando $k, n \in \mathbb{N}$*, a cobertura de* $\{1,2\}^{\mathbb{N}}$ *em cilindros de tamanho* $k + n$ *e a métrica* $d((x_1, x_2, x_3, \dots), (y_1, y_2, y_3, \dots)) = \frac{1}{2^i}$*, onde* $i = \min\{i : x_i \neq y_i\}$*, temos que esta cobertura cobre o espaço com* 2^{k+n} *elementos e a distância* d_n *entre cada elemento de uma partição é menor do que* 2^{-k-1}*. Com isso, obtemos que* $C_d(n, 2^{-k-1}) \leqslant 2^{k+n}$*. Além disso, escolhendo um elemento de cada cilindro da cobertura formamos um conjunto* $(n, 2^{-k-1})$*-separado, logo* $2^{n+k} \leqslant N_d(n, 2^{-k-1})$*. Pelo Lema 3.2.3, concluímos que* $C_d(n, 2^{-k-1}) = N_d(n, 2^{-k-1}) = 2^{k+n}$*. Logo*

$$\begin{aligned}
h_{top}(\sigma) &= \lim_{k \to \infty} \lim_{n \to \infty} \frac{1}{n} \log C_d(n, 2^{-k-1}) \\
&= \lim_{k \to \infty} \lim_{n \to \infty} \frac{1}{n} \log 2^{k+n} \\
&= \lim_{k \to \infty} \lim_{n \to \infty} \frac{k+n}{n} \log 2 \\
&= \lim_{k \to \infty} \log 2 = \log 2.
\end{aligned}$$

A partir de agora vamos apresentar algumas propriedades importantes da entropia topológica que muitas vezes podem servir para facilitar o seu cálculo.

Proposição 3.2.8. *Seja* $(\mathcal{X}, d)$ *um espaço métrico compacto. Então a entropia topológica de uma aplicação contínua* $T : \mathcal{X} \to \mathcal{X}$ *não depende da escolha de uma métrica particular que gera a topologia de* $\mathcal{X}$

Demonstração. Sejam d e d' duas métricas equivalentes em $\mathcal{X}$. Como as métricas d_n são duas a duas equivalentes para todo $n \in \mathbb{N}$, então obtemos que as métricas d_n e d' são equivalentes e, consequentemente, d_n e d'_n também são equivalentes. Fixado $\epsilon > 0$, consideramos A um conjunto (n, ϵ)-gerador com respeito a métrica d e com cardinalidade $G_d(n, \epsilon)$. Isto implica que $\mathcal{X}$ pode ser coberto por $G_d(n, \epsilon)$ (n, ϵ)-bolas dinâmicas centradas nos pontos de A. Como as métricas d_n e d'_n são equivalentes, então existe um $\delta > 0$ tal que $d'_n(x, a) < \delta$ implica $d_n(x, a) < \epsilon$ para todo elemento de A. Logo, toda (n, δ)-bola dinâmica centrada nos pontos de A está contida na (n, ϵ)-bola dinâmica centrada no mesmo ponto de A, com respeito as métricas d'_n e d_n, respectivamente. Com isso, obtemos que $G_d(n, \epsilon) \leqslant G_{d'}(n, \delta)$ e $h(T) \leqslant h'(T)$. Analogamente obtemos que $h(T) \geqslant h'(T)$ e, consequentemente, $h(T) = h'(T)$.

□

Teorema 3.2.9. *Sejam $(\mathcal{X}, T)$ e $(\mathcal{Y}, S)$ sistemas dinâmicos topologicamente conjugados tais que a conjugação $\varphi : (\mathcal{X}, T) \to (\mathcal{Y}, S)$ é uma isometria, ou seja, $d^{\mathcal{X}}(x, y) = d^{\mathcal{Y}}(\varphi(x), \varphi(y)), \forall x, y \in \mathcal{X}$. Então $h(T) = h(S)$.*

Demonstração. Como os sistemas dinâmicos são conjugados, ou seja, existe um homeomorfismo (bijeção contínua com inversa contínua) $\varphi : \mathcal{X} \to \mathcal{Y}$ tal que $\forall n \in \mathbb{N}$,

$$\varphi \circ T^n = S^n \circ \varphi$$

e φ é uma isometria, $\forall x, y \in \mathcal{X}$ obtemos

$$\begin{aligned}
d_n^{\mathcal{X}}(x, y) &= \max_{0 \leqslant i \leqslant n-1} \left\{ d(T^i(x), T^i(y)) \right\} \\
&= \max_{0 \leqslant i \leqslant n-1} \left\{ d(\varphi(T^i(x)), \varphi(T^i(y))) \right\} \\
&= \max_{0 \leqslant i \leqslant n-1} \left\{ d(S^i(\varphi(x)), S^i(\varphi(y))) \right\} \\
&= d_n^{\mathcal{Y}}(\varphi(x), \varphi(y)).
\end{aligned}$$

Portanto, considerando $A \subset \mathcal{X}$ um conjunto (n, ϵ)-separado tal que $|A| = N_d(n, \epsilon)$, obtemos que $\varphi(A)$ é um conjunto (n, ϵ)-separado em $\mathcal{Y}$. Além disso, como φ é uma bijeção, a cardinalidade de A e $\varphi(A)$ é a mesma. Com isso, concluímos que

$$N_d^{\mathcal{X}}(n,\epsilon) = |A| = |\varphi(A)| \leqslant N_d^{\mathcal{Y}}(n,\epsilon).$$

Por outro lado, observamos que φ^{-1} também é uma isometria e, analogamente

$$N_d^{\mathcal{Y}}(n,\epsilon) \leqslant N_d^{\mathcal{X}}(n,\epsilon).$$

Logo

$$N_d^{\mathcal{X}}(n,\epsilon) = N_d^{\mathcal{Y}}(n,\epsilon).$$

Portanto

$$\begin{aligned} h(T) &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_d^{\mathcal{X}}(n,\epsilon) \\ &= \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_d^{\mathcal{Y}}(n,\epsilon) \\ &= h(S). \end{aligned}$$

□

Corolário 3.2.10. *A entropia topológica é um invariante por conjugação topológica.*

Demonstração. Seja $\varphi : (\mathcal{X}, T) \to (\mathcal{Y}, S)$ a conjugação topológica entre os sistemas dinâmicos. Então, definindo uma métrica em $\mathcal{X}$ da seguinte forma

$$d^{\mathcal{X}}(x,y) = d^{\mathcal{Y}}(\varphi(x), \varphi(y)), \forall x, y \in \mathcal{X},$$

obtemos que φ é uma isometria. Logo $h(T) = h(S)$. □

Quando nos referirmos a um sistema dinâmico topológico, estamos considerando uma dupla $(\mathcal{X}, T)$ onde $\mathcal{X}$ é um espaço topológico e T uma transformação contínua.

Corolário 3.2.11. *Sejam $(\mathcal{X}, T)$ e $(\mathcal{Y}, S)$ sistemas dinâmicos topológicos com $h(T) \neq h(S)$. Então os sistemas não podem ser topologicamente conjugados.*

Proposição 3.2.12. *Seja $(\mathcal{X}, d)$ espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ contínua. Então $h(T^m) = m \cdot h(T), \forall m \in \mathbb{N}$. Além disso, se T é invertível, então $h(T^m) = |m| \cdot h(T), \forall m \in \mathbb{Z}$.*

Teorema 3.2.13. *Seja $(\mathcal{X}, T)$ um sistema dinâmico topológico tal que T é uma isometria. Então $h(T) = 0$.*

Demonstração. Como T é uma isometria, para todo $x, y \in \mathcal{X}$ temos que

$$\begin{aligned} d(x, y) &= d(T(x), T(y)) \\ &= d(T^2(x), T^2(y)) \\ &= d(T^n(x), T^n(y)), \forall n \in \mathbb{N}. \end{aligned}$$

Logo, $C_d(n, \epsilon) = C_d(1, \epsilon)$. Então

$$\begin{aligned} h_\epsilon(T) &= \lim_{n \to \infty} \frac{1}{n} \log C_d(n, \epsilon) \\ &= \lim_{n \to \infty} \frac{1}{n} \log C_d(1, \epsilon) \\ &= 0. \end{aligned}$$

Portanto

$$h(T) = \lim_{\epsilon \to 0} h_\epsilon(T) = 0.$$

□

O próximo exemplo é a rotação, que é uma aplicação bem conhecida em sistemas dinâmicos e é muito rica dinamicamente, porém, por se tratar de uma isometria, tem entropia nula.

Exemplo 3.2.14 (Rotação)**.** *Seja $S^1 = [0, 1]/\sim$, onde $\sim$ indica que 0 e 1 estão identificados. Então, para todo $\theta \in \mathbb{R}$, denotamos por rotação de ângulo $2\pi\theta$ a aplicação $R_\theta : S^1 \to S^1$ dada por*

$$R_\theta(x) = x + \theta \mod 1.$$

Capítulo 4

Princípio Variacional Clássico

Neste capítulo vamos apresentar o Princípio Variacional Clássico que relaciona entropia métrica e entropia topológica. A demonstração deste resultado é feita com o auxílio de um lema encontrado em [1] que também será provado neste trabalho.

Para exemplificar um pouco, anteriormente calculamos a entropia métrica do deslocamento σ em relação a medida de Bernoulli no espaço $\{1, \ldots, d\}^{\mathbb{N}}$. Considerando $d = 2$ e a medida de Bernoulli com pesos $1/2$ e $1/2$, obtemos que a entropia métrica é $\log 2$, porém, quando considerando a medida de Bernoulli com pesos $1/3$ e $2/3$ a entropia métrica é $\log 3 - 2/3 \log 2$. No exemplo 3.2.7, calculamos a entropia topológica do deslocamento neste mesmo contexto e obtivemos como resultado $\log 2$, ou seja, a entropia topológica coincidiu com a entropia métrica quando considerada a medida de Bernoulli com os pesos igualmente distribuídos no conjunto $\{1, 2\}$.

Agora, antes de enunciarmos e provarmos o lema, vamos apresentar uma definição e dois resultados importantes, cujas demonstrações são omitidas mas podem ser encontradas em [3].

Definição 4.0.1. *Seja μ uma medida de probabilidade T-invariante. Para $\delta \in (0, 1), n \in \mathbb{N}$ e $\epsilon > 0$ definimos*

$$h_\mu(\epsilon, T, \delta) = \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n, \epsilon, \delta),$$

onde $N_\mu(n, \epsilon, \delta)$ é o número mínimo de (n, ϵ)-bolas dinâmicas necessárias *para cobrir um conjunto de medida estritamente maior que $1-\delta$ com respeito à μ.*

Lema 4.0.2. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto, $T : \mathcal{X} \to \mathcal{X}$ uma aplicação contínua, μ uma medida de probabilidade ergódica T-invariante e $\delta \in (0,1)$. Então*

$$h_\mu(\mathcal{P}, T,) \geqslant \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n, \epsilon, \delta),$$

onde $\mathcal{P}$ é uma partição de diâmetro menor do que ϵ.

Teorema 4.0.3. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto, $T : \mathcal{X} \to \mathcal{X}$ uma aplicação contínua e μ uma medida de probabilidade ergódica T-invariante. Então*

$$h_\mu(T) = \lim_{\epsilon\to 0} h_\mu(\epsilon, T, \delta),$$

onde $h_\mu(T)$ é a entropia métrica com respeito à medida μ. Em particular o limite acima não depende de $\delta \in (0,1)$.

Lema 4.0.4. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ uma aplicação contínua. Dado $\epsilon > 0$, existe uma medida de probabilidade T-invariante μ_ϵ tal que*

$$h_{\mu_\epsilon}(\epsilon, T, \delta) \geqslant S(\mathcal{X}, d, 2\epsilon) - 3\delta S(\mathcal{X}, d, \epsilon).$$

Demonstração. Seja $E_n = \left\{x_1, x_2, \ldots, x_{N_d(n,\epsilon)}\right\}$ um conjunto de pontos (n, ϵ)-separados em $\mathcal{X}$ tal que $|E_n| = N_d(n, \epsilon)$. Definimos a medida

$$\sigma_n = \frac{1}{|E_n|} \sum_{x\in E_n} \delta_x,$$

onde δ_x é a medida de probabilidade suportada no ponto x. Além disso, definimos também a medida

$$\overline{\sigma}_n = \frac{1}{n} \sum_{k=0}^{n-1} T_*^k \sigma_n.$$

onde $T_*^k \sigma_n(A) = \sigma_n(T^{-k}(A))$ para qualquer A pertencente a σ-álgebra.

Consideramos $(n_k)_{k\in\mathbb{N}}$ uma subsequência tal que

$$S(\mathcal{X}, d, \epsilon) = \lim_{k\to\infty} \frac{1}{n_k} \log N_d(n_k, \epsilon).$$

Como $\mathcal{X}$ é compacto, o espaço $\mathcal{M}_1(\mathcal{X})$ das medidas de probabilidades em $\mathcal{X}$ munido da topologia fraca* é compacto, logo toda sequência de probabilidades em $\mathcal{M}_1(\mathcal{X})$ admite alguma subsequência convergente na topologia fraca*. Pelo Lema 2.2.27 todo ponto de acumulação de uma sequência $(\mu_n)_n$ com $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} T_*^k \nu$, onde ν é uma probabilidade qualquer é uma probabilidade invariante por T. Portanto, podemos encontrar uma subsequência de (n_k) também denotada por $(n_k)_{k\in\mathbb{N}}$, tal que $(\overline{\sigma}_{n_k})_{k\in\mathbb{N}}$ converge para uma medida de probabilidade T-invariante μ e os termos desta sequência são escolhidos de tal forma que $\lim_{k\to\infty} \frac{n_k}{k} = \infty$.

Seja K um subconjunto de $\mathcal{X}$ tal que $\mu(K) > 1-\delta$ e $N(K, n, \epsilon/2) = N_\mu(n, \epsilon/2, \delta)$, onde $N(K, n, \epsilon/2)$ é definido como sendo o número mínimo de $(n, \epsilon/2)$-bolas dinâmicas necessárias para cobrir K. Além disso, podemos tomar K um conjunto aberto, pois como $\overline{\sigma}_{n_k} \to \mu$ na topologia fraca* obtemos que para qualquer $\epsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que $\forall k \geqslant k_0$ temos $\overline{\sigma}_{n_k}(A) > \mu(A) - \epsilon$, para qualquer aberto de $\mathcal{X}$. Logo, para ϵ suficientemente pequeno, obtemos $\overline{\sigma}_{n_k}(K) > \mu(K) - \epsilon > 1 - \delta, \forall k \geqslant k_0$.

Agora, definimos $L_n = \{(i,j) \in \mathbb{N}^2 : 0 \leqslant i \leqslant n-1, 1 \leqslant j \leqslant N_d(n,\epsilon)\}$ onde se $T^i x_j \in K$ então associamos 1 ao par (i,j) e 0 caso contrário, ou seja, $(i,j) = 0$ se $T^i x_j \notin K$. Pela definição de $\overline{\sigma}_{n_k}$ sabemos que o número de uns em L_{n_k} é maior ou igual que $n_k N_d(n_k, \epsilon)(1-\delta)$. De fato,

$$\begin{aligned}
\overline{\sigma}_{n_k}(K) &= \frac{1}{n_k} \sum_{l=0}^{n_k-1} T_*^l \sigma_{n_k}(K) \\
&= \frac{1}{n_k} \sum_{l=0}^{n_k-1} \sigma_{n_k}(T^{-l}(K)) \\
&= \frac{1}{n_k N_d(n_k, \epsilon)} \sum_{l=0}^{n_k-1} \sum_{x\in E_{n_k}} \delta_x(T^{-l}(K)),
\end{aligned}$$

denotando o número de uns em L_{n_k} por τ, obtemos

$$\tau = \sum_{l=0}^{n_k-1} \sum_{x \in E_{n_k}} \delta_x(T^{-l}(K)),$$

pois, para todo $n \in \mathbb{N}$, $T^n(x) \in K$ se, e somente se, $x \in T^{-n}(K)$. Portanto,

$$\begin{aligned}\tau &= \overline{\sigma}_{n_k}(K) n_k N_d(n_k, \epsilon)\\ &\geqslant n_k N_d(n_k, \epsilon)(1-\delta)\end{aligned}$$

já que $\overline{\sigma}_{n_k}(K) \geqslant 1-\delta$.

Para $s \in \mathbb{R}$ e $s > 1$, definimos $L_{n_k}(s)$ como sendo o conjunto de pontos de L_{n_k} com as primeiras coordenadas no intervalo $\left[\left[\frac{n_k}{k}\right], n_k - \left[\frac{n_k}{s}\right] - 1\right]$. Logo, o número de uns em $L_{n_k}(s)$ é pelo menos

$$n_k N_d(n_k,\epsilon)\left(1-\delta-\frac{1}{n_k}\left[\frac{n_k}{s}\right]-\frac{1}{n_k}\left[\frac{n_k}{k}\right]\right) \geqslant n_k N_d(n_k,\epsilon)\left(1-\delta-\frac{1}{s}-\frac{1}{k}\right).$$

De fato, denotando o número de uns em $L_{n_k}(s)$ por $\tau(s)$, temos que

$$\begin{aligned}\tau(s) &= \tau - \sum_{l=0}^{\left[\frac{n_k}{k}\right]-1} \sum_{x\in E_{n_k}} \delta_x(T^{-l}(K)) - \sum_{l=n_k-\left[\frac{n_k}{s}\right]}^{n_k-1} \sum_{x\in E_{n_k}} \delta_x(T^{-l}(K))\\ &\geqslant \tau - \left[\frac{n_k}{k}\right] N_d(n_k,\epsilon) - \left[\frac{n_k}{s}\right] N_d(n_k,\epsilon)\\ &= n_k N_d(n_k,\epsilon)(1-\delta) - n_k N_d(n_k,\epsilon)\left(\frac{1}{n_k}\left[\frac{n_k}{k}\right]-\frac{1}{n_k}\left[\frac{n_k}{s}\right]\right)\\ &= n_k N_d(n_k,\epsilon)\left(1-\delta-\frac{1}{n_k}\left[\frac{n_k}{k}\right]-\frac{1}{n_k}\left[\frac{n_k}{s}\right]\right)\\ &\geqslant n_k N_d(n_k,\epsilon)\left(1-\delta-\frac{1}{s}-\frac{1}{k}\right).\end{aligned}$$

Agora, assumimos que $s > \frac{1}{1-2\delta-\frac{1}{k}}$, implicando que $\frac{1}{s} < 1-2\delta-\frac{1}{k}$. Então $-\frac{1}{s} > 2\delta + \frac{1}{k} - 1$ e, com isso, obtemos que $1-\delta-\frac{1}{s}-\frac{1}{k} > \delta$. Logo, para esta escolha de s obtemos $\tau(s) \geqslant n_k N_d(n_k,\epsilon)\delta$.

Pela definição de L_{n_k}, podemos associar uma matriz (l_{ij}) de zeros e uns onde para cada entrada l_{ij} atribuímos o valor referente ao par (i,j) correspondente. Desse modo, a matriz correspondente a $L_{n_k}(s)$ tem $N_d(n_k,\epsilon)$ colunas e $\left(n_k - \left[\frac{n_k}{s}\right] - \left[\frac{n_k}{k}\right]\right)$ linhas.

Calculando o número médio de uns por linhas de $L_{n_k}(s)$, obtemos que existe um índice m_k tal que a m_k-ésima linha tem pelo menos

$$\frac{n_k N_d(n_k,\epsilon)\delta}{\left(n_k - \left[\frac{n_k}{s}\right] - \left[\frac{n_k}{k}\right]\right)}$$

uns e $\left[\frac{n_k}{k}\right] \leqslant m_k < n_k - \left[\frac{n_k}{s}\right]$.

Além disso, escolhendo B como sendo um conjunto de pontos $(m_k,\epsilon/2)$-separados com cardinalidade igual a $N_d(m_k,\epsilon/2)$, isto implica que podemos cobrir $\mathcal{X}$ com bolas dinâmicas centradas nos pontos de B e raio $\epsilon/2$, pois, como $d_{m_k}(x,y) \geqslant \epsilon/2, \forall x,y \in B$, temos que B é um conjunto $(m_k,\epsilon/2)$-gerador, ou seja, $\forall x \in \mathcal{X}, \exists b \in B$, tal que $d_{m_k}(x,b) < \epsilon/2$. Se supormos por absurdo que existe pelo menos um ponto $x \in \mathcal{X}$ que não está contido em alguma bola, implicaria que $B \cup \{x\}$ seria um conjunto $N_d(m_k,\epsilon/2)$-separado, contrariando a maximalidade de B, já que $|B| = N_d(m_k,\epsilon/2)$.

Observamos que se $i \neq j$ e $d_{m_k}(x_i,x_j) \leqslant \epsilon$, então

$$d_{n_k-m_k}(T^{m_k}(x_i), T^{m_k}(x_j) > \epsilon,$$

pois, pela definição do conjunto E_{n_k}, sabemos que $d_{n_k}(x_i,x_j) > \epsilon$. Logo, podemos concluir que existe um subconjunto $I \subset \{1,2,\ldots,N_d(n_k,\epsilon)\}$ tal que para cada $i \in I$, temos $T^{m_k}(x_i) \in K$ e

$$|I| \geqslant \frac{n_k N_d(n_k,\epsilon)\delta}{\left(n_k - \left[\frac{n_k}{s}\right] - \left[\frac{n_k}{k}\right]\right)} > N_d(n_k,\epsilon)\delta.$$

Além disso, podemos obter um conjunto $A \subset \{T^{m_k}(x_i)\}_{i\in I}$ tal que o diâmetro de A com respeito d_{m_k} é menor que ϵ e

$$|A| \geqslant \frac{N_d(n_k,\epsilon)\delta}{N_d(m_k,\epsilon/2)}.$$

Para fazer isso, basta dividir os elementos de $\{T^{m_k}(x_i)\}_{i\in I}$ nas $N_d(m_k,\epsilon/2)$ $(m_k,\epsilon/2)$-bolas dinâmicas, então vai existir pelo menos um conjunto com a cardinalidade esperada. Portanto, se $a, b \in A$ e $a \neq b$, então $d_{n_k}(a, b) \geqslant d_{n_k - m_k}(a, b) \geqslant \epsilon$. Consequentemente,

$$N_\mu \left(n_k, \epsilon/2, \delta\right) = N \left(K, n_k, \epsilon/2\right) \geqslant \frac{N_d(n_k, \epsilon)\delta}{N_d(m_k, \epsilon/2)}.$$

De fato, como $A \subset K$ e $|A| \leqslant N_d(n_k, \epsilon)$, isto implica que

$$|A| \leqslant C_d(n_k, \epsilon) \leqslant C_d(n_k, \epsilon/2)$$

para alguma cobertura de K, em particular, temos que

$$C_d(n_k, \epsilon/2) \leqslant N(K, n_k, \epsilon/2) = N_\mu(n, \epsilon/2, \delta).$$

Logo

$$\begin{aligned} h_\mu(\epsilon/2, T, \delta) &= \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n, \epsilon/2, \delta) \\ &\geqslant \limsup_{k\to\infty} \left(\frac{1}{n_k} \log N_d(n_k, \epsilon) - \frac{m_k}{n_k} \frac{1}{m_k} \log N_d(m_k, \epsilon/2) \right). \end{aligned}$$

Como $m_k < n_k - \left[\frac{n_k}{s}\right]$, então $\frac{m_k}{n_k} \leqslant 1 - \frac{1}{s} + \frac{1}{n_k}$ e

$$\begin{aligned} h_\mu(\epsilon/2, T, \delta) &\geqslant \limsup_{k\to\infty} \left(\frac{\log N_d(n_k, \epsilon)}{n_k} - \left(1 - \frac{1}{s} + \frac{1}{n_k}\right) \frac{\log N_d(m_k, \epsilon/2)}{m_k} \right) \\ &\geqslant \limsup_{k\to\infty} \left(\frac{1}{n_k} \log N_d(n_k, \epsilon) - \left(3\delta + \frac{1}{n_k}\right) \frac{1}{m_k} \log N_d(m_k, \epsilon/2) \right), \end{aligned}$$

pois podemos tomar $s < \frac{1}{1-3\delta}$, basta escolher k suficientemente grande de modo que $\delta > 1/k$. Além disso, como $\frac{n_k}{k} \to \infty$ e $\left[\frac{n_k}{k}\right] \leqslant m_k$, isto implica que $m_k \to \infty$ quando $k \to \infty$.

Antes de continuar a demonstração do lema vamos provar dois resultados que envolvem o limite superior de sequências.

Lema 4.0.5. *Sejam $(x_n)_n$ e $(y_n)_n$ sequências tais que $\lim_{n\to\infty} x_n$ existe. Então*

$$\limsup_{n\to\infty}(x_n - y_n) \geqslant \lim_{n\to\infty} x_n - \limsup_{n\to\infty} y_n$$

Demonstração. Supondo que x_n e y_n são sequências tais que $\lim_{n\to\infty} x_n$ existe. Então, dado $\epsilon > 0$, para n suficientemente grande obtemos

$$x - \epsilon - y_n < x_n - y_n < x + \epsilon - y_n$$

tomando o limite superior em ambos os lados, temos que

$$x - \epsilon + \limsup_{n\to\infty} -y_n \leqslant \limsup_{n\to\infty}(x_n - y_n) \leqslant x + \epsilon + \limsup_{n\to\infty} -y_n$$

usando que $\limsup_{n\to\infty} -y_n = -\liminf_{n\to\infty} y_n$, temos

$$\limsup_{n\to\infty}(x_n - y_n) = \lim_{n\to\infty} x_n - \liminf_{n\to\infty} y_n$$

como $\liminf_{n\to\infty} y_n \leqslant \limsup_{n\to\infty} y_n$, concluímos que

$$\limsup_{n\to\infty}(x_n - y_n) \geqslant \lim_{n\to\infty} x_n - \limsup_{n\to\infty} y_n$$

□

Lema 4.0.6. *Sejam* $(a_n)_n$ *e* $(b_n)_n$ *sequências tais que* $\lim_{n\to\infty} a_n = a, a \in \mathbb{R}, a > 0$ *e* $b_n \geqslant 0$*. Então*

$$\limsup_{n\to\infty} (a_n \cdot b_n) = a \cdot \limsup_{n\to\infty} b_n.$$

Demonstração. Pela definição de limite e limite superior, dado $\epsilon > 0$, podemos encontrar um $n_0 \in \mathbb{N}$ tal que $\forall n > n_0$, temos

$$a - \epsilon < a_n < a + \epsilon \text{ e } b_n < L + \epsilon,$$

onde $L = \limsup_{n\to\infty} b_n$. Além disso, também podemos construir uma sequência $(n_k)_{k\in\mathbb{N}}$ tal que $L - \epsilon < b_{n_k}$.

Portanto, por um lado obtemos

$$\begin{aligned}\limsup_{n\to\infty}(a_n \cdot b_n) &\leqslant (a + \epsilon)(L + \epsilon)\\ &= aL + a\epsilon + L\epsilon + \epsilon^2.\end{aligned}$$

Fazendo $\epsilon \to 0$, temos $a_n \cdot b_n \leqslant aL$.

Por outro lado, temos que

$$\begin{aligned}\limsup_{k\to\infty}(a_{n_k}\cdot b_{n_k}) &\geqslant (a-\epsilon)(L-\epsilon)\\ &\geqslant aL - a\epsilon - L\epsilon + \epsilon^2,\end{aligned}$$

fazendo $\epsilon \to 0$, concluímos que

$$\limsup_{n\to\infty}(a_n\cdot b_n) = a\cdot\limsup_{n\to\infty} b_n.$$

Caso $L=\infty$, para todo $\epsilon>0$, existe $n_0\in\mathbb{N}$ tal que para todo $n>n_0$, obtemos $a_n>a-\epsilon$ e $b_n>\frac{M}{a-\epsilon}>0$, onde M é qualquer número real positivo. Então, $a_n\cdot b_n>M, \forall n>n_0$ e, consequentemente,

$$\limsup_{n\to\infty}(a_n\cdot b_n) = \infty = aL.$$

□

Portanto, continuando a demonstração do lema, obtemos

$$\begin{aligned}h_\mu(\epsilon/2,T,\delta) &\geqslant \lim_{k\to\infty}\frac{\log N_d(n_k,\epsilon)}{n_k} - \limsup_{k\to\infty}\left(\left(3\delta+\frac{1}{n_k}\right)\frac{\log N_d(m_k,\epsilon/2)}{m_k}\right)\\ &= \lim_{k\to\infty}\frac{1}{n_k}\log N_d(n_k,\epsilon) - 3\delta\limsup_{k\to\infty}\frac{1}{m_k}\log N_d(m_k,\epsilon/2)\\ &= S(\mathcal{X},d,\epsilon) - 3\delta S(\mathcal{X},d,\epsilon/2).\end{aligned}$$

Logo,

$$h_\mu(\epsilon,T,\delta) \geqslant S(\mathcal{X},d,2\epsilon) - 3\delta S(\mathcal{X},d,\epsilon).$$

Em particular, obtemos que

$$\sup_{\mu\in\mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon,T,\delta) \geqslant S(\mathcal{X},d,2\epsilon) - 3\delta S(\mathcal{X},d,\epsilon).$$

□

A demonstração do próximo lema será omitida, porém pode ser encontrada em [11].

Lema 4.0.7. *Seja E_n um conjunto (n,ϵ)-separado, $\nu_n = \frac{1}{|E_n|}\sum_{x\in E_n}\delta_x$ e $\mu_n = \frac{1}{n}\sum_{i=0}^{n-1} T_*^i \nu_n$. Se μ é um ponto de acumulação de $(\mu_n)_{n\in\mathbb{N}}$ então μ é uma medida T-invariante e*

$$\limsup_{n\to\infty}\frac{1}{n}\log|E_n| \leqslant h_\mu(T,\mathcal{P}),$$

onde $\mathcal{P}$ é uma partição finita mensurável tal que seus elementos têm diâmetro menor do que ϵ e $\mu\left(\bigcup_{P\in\mathcal{P}}\partial P\right) = 0$.

Corolário 4.0.8 (Princípio Variacional Clássico)**.** *Seja $(\mathcal{X},d)$ um espaço métrico compacto e $T:\mathcal{X}\to\mathcal{X}$ contínua. Então*

$$h_{top}(T) = \sup_{\mu\in\mathcal{M}_T(\mathcal{X})} h_\mu(T).$$

Demonstração. Considerando E_n um conjunto (n,ϵ)-separado tal que $|E_n| = N_d(n,\epsilon)$. Pelo Lema 4.0.7, existe uma medida T-invariante ν tal que

$$\limsup_{n\to\infty}\frac{1}{n}\log N_d(n,\epsilon) \leqslant h_\nu(T,\mathcal{P})$$

e $\mathcal{P}$ é uma partição finita de $\mathcal{X}$ em conjunto mensuráveis tal que seus elementos possuem diâmetro menor do que ϵ e $\nu\left(\bigcup_{P\in\mathcal{P}}\partial P\right) = 0$.

Para construir uma partição que satisfaz esta condição, seja $\{B_1,\ldots B_k\}$ uma cobertura de $\mathcal{X}$ por bolas de raio menor do que $\epsilon/2$ e $\nu(\partial B_i) = 0$. Definimos $\mathcal{P} = \{P_1,\ldots,P_k\}$, onde

$$P_1 = \overline{B}_1 \text{ e } P_i = \overline{B}_i \setminus \bigcup_{j=1}^{i-1}\overline{B}_j.$$

Então, cada elemento da partição

$$\mathcal{P}^n = \mathcal{P}\vee\cdots\vee T^{-n+1}(\mathcal{P})$$

possui diâmetro em relação a métrica d_n menor do que ϵ e $\nu(\partial P_i) = 0$ para todo $i \in \{1, \ldots, k\}$. Para mostrar que $\nu(\partial P_i) = 0$, basta notar que $\bigcup_{\epsilon' \in [0, \epsilon/2)} \partial B(x, \epsilon') = B(x, \epsilon/2)$ tem medida finita e é uma união disjunta não enumerável, logo existe pelo menos um $\epsilon' < \epsilon/2$ tal que $\nu(\partial B(x, \epsilon')) = 0$.

De fato, supondo por absurdo que existe uma quantidade não enumerável de parcelas da soma com medida positiva. Definindo

$$A = \left\{ \epsilon' \in \left[0, \frac{\epsilon}{2}\right) : \nu(\partial B(x, \epsilon')) > 0 \right\}$$

e

$$A_n = \left\{ \epsilon' \in \left[0, \frac{\epsilon}{2}\right) : \nu(\partial B(x, \epsilon')) \geqslant \frac{1}{n} \right\},$$

temos que

$$A = \bigcup_{n=1}^{\infty} A_n.$$

Como A é não enumerável isto implica que existe pelo menos um índice n tal que A_n é não enumerável, consequentemente infinito e

$$\mu\left(\bigcup_{\epsilon' \in A_n} \nu(\partial B(x, \epsilon')) \right) = \sum_{\epsilon' \in A_n} \nu(\partial B(x, \epsilon')) \geqslant |A_n| \cdot \frac{1}{n} = \infty.$$

Porém, a medida de A deve ser finita. Logo, existe uma quantidade enumerável de parcelas com medida positiva, implicando que existe pelo menos um $\epsilon' < \epsilon/2$ tal que $\nu(\partial B(x, \epsilon')) = 0$. Além disso, como $\partial P_i \subset \bigcup_{j=1}^{k} \partial B_j$, isto implica que $\nu(\partial P_i) = 0$.

Como podemos cobrir $\mathcal{X}$ por $|E_n|$ (n, ϵ)-bolas dinâmicas, isto implica que o número mínimo de (n, ϵ)-bolas dinâmicas necessárias para cobrir $\mathcal{X}$ é menor do que $N_d(n, \epsilon)$. Em particular, vale também que para todo $\delta \in (0, 1)$ e qualquer medida $\mu \in \mathcal{M}_T(\mathcal{X})$

$$N_\mu(n, \epsilon, \delta) \leqslant N_d(n, \epsilon).$$

Logo,

$$\limsup_{n \to \infty} \frac{1}{n} \log N_\mu(n, \epsilon, \delta) \leqslant \limsup_{n \to \infty} \frac{1}{n} \log N_d(n, \epsilon) \leqslant h_\nu(T, \mathcal{P}).$$

Portanto, obtemos que

$$h_\mu(\epsilon, T, \delta) \leqslant h_\nu(T, \mathcal{P}) \leqslant h_\nu(T)$$

Agora, consideramos $\nu = \int_{\mathcal{X}} \nu_x \; d\nu(x)$ a decomposição ergódica de ν e, analogamente a esta decomposição, também podemos escrever

$$h_\nu(T) = \int h_{\nu_x}(T) d\nu(x).$$

Logo

$$h_\mu(\epsilon, T, \delta) \leqslant h_\nu(T) \leqslant \sup_{x \in \mathcal{X}} h_{\nu_x}(T) \leqslant \sup_{\nu \text{ ergódicas}} h_\nu(T).$$

Como μ é qualquer medida de probabilidade T-invariante, escolhemos μ tal que

$$h_\mu(\epsilon, T, \delta) \geqslant S(\mathcal{X}, d, 2\epsilon) - 3\delta S(\mathcal{X}, d, \epsilon).$$

Pelo Teorema 4.0.3, temos que $h_\nu(T) = \lim_{\epsilon \to 0} h_\nu(\epsilon, T, \delta)$ para toda medida ergódica T-invariante e $\delta \in (0, 1)$. Então, como $N_\nu(n, \epsilon, \delta) \leqslant N_d(n, \epsilon)$, obtemos $h_\nu(T) \leqslant h_{top}(T)$. Organizando as desigualdades, concluímos que

$$S(\mathcal{X}, d, 2\epsilon) - 3\delta S(\mathcal{X}, d, \epsilon) \leqslant h_\mu(\epsilon, T, \delta) \leqslant \sup_{\nu \text{ ergódicas}} h_\nu(T) \leqslant h_{top}(T).$$

Fazendo $\epsilon \to 0$, obtemos

$$h_{top}(T) - 3\delta h_{top}(T) \leqslant \sup_{\mu \text{ ergódicas}} h_\mu(T) \leqslant h_{top}(T)$$

e com $\delta \to 0$ concluímos que

$$\sup_{\mu \text{ ergódicas}} h_\mu(T) = h_{top}(T).$$

□

Note que, como estamos considerando o caso em que $h_{top}(T)$ pode ser infinita, estamos adotando a aritmética da reta estendida $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$.

Capítulo 5

Dimensão Métrica Média

Neste capítulo vamos apresentar a dimensão métrica média seguindo as ideias propostas por A. Velozo e R. Velozo em [1]. Definimos um Princípio Variacional para sistemas dinâmicos com entropia topológica infinita que, posteriormente, é utilizado para relacionar a dimensão métrica média com a dimensão de Minkowski. Ao estabelecer esta relação, conseguimos generalizar o caso em que consideramos a aplicação deslocamento σ no espaço $[0,1]^{\mathbb{Z}}$ para o caso $([0,1]^d)^{\mathbb{Z}}$.

Inicialmente vamos estabelecer uma notação que será útil na apresentação dos resultados seguintes. Considerando $N_d(n,\epsilon)$ a cardinalidade máxima de um conjunto (n,ϵ)-separado em $(\mathcal{X},d)$, denotamos

$$S(\mathcal{X},d,\epsilon) = \limsup_{n\to\infty} \frac{1}{n}\log N_d(n,\epsilon).$$

Definimos a dimensão métrica média superior como

$$\overline{mdim}(\mathcal{X},d,T) = \limsup_{\epsilon\to 0} \frac{S(\mathcal{X},d,\epsilon)}{|\log(\epsilon)|}.$$

Analogamente, definimos a dimensão métrica média inferior como

$$\underline{mdim}(\mathcal{X},d,T) = \liminf_{\epsilon\to 0} \frac{S(\mathcal{X},d,\epsilon)}{|\log(\epsilon)|}.$$

Observamos que se $\mathcal{X}$ é um espaço métrico compacto, então sabemos que entropia topológica é independente da métrica, desde que ela gere a topologia de $\mathcal{X}$, porém, a dimensão métrica média depende da métrica ambiante. Além disso, observamos que se a entropia topológica é finita, então a dimensão métrica média é nula. Um caso em que a entropia topológica é finita é o Exemplo 3.2.7, que possui entropia topológica $\log 2$ e $S(\mathcal{X}, d, \epsilon) = \log 2$. Calculando a dimensão métrica média superior deste exemplo, obtemos que

$$\overline{mdim}(\mathcal{X}, d, T) = \limsup_{\epsilon \to 0} \frac{S(\mathcal{X}, d, \epsilon)}{|\log(\epsilon)|} = \limsup_{\epsilon \to 0} \frac{\log 2}{\log \epsilon} = 0.$$

5.1 Princípio Variacional

Nesta seção vamos apresentar um princípio variacional para sistemas dinâmicos com entropia topológica infinita e um corolário que permite calcular a dimensão métrica média de espaços métricos compactos mais gerais. Além disso, também calculamos a dimensão métrica média superior do deslocamento em $[0,1]^{\mathbb{Z}}$.

Teorema 5.1.1 (Princípio Variacional)**.** *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ contínua. Então*

$$\overline{mdim}(\mathcal{X}, d, T) = \lim_{\delta \to 0} \limsup_{\epsilon \to 0} \frac{\sup_{\mu \in \mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon, T, \delta)}{|\log(\epsilon)|}.$$

Demonstração. Pelo Lema anterior obtemos

$$\sup_{\mu \in \mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon, T, \delta) \geqslant S(\mathcal{X}, d, 2\epsilon) - 3\delta S(\mathcal{X}, d, \epsilon).$$

Logo,

$$\begin{aligned} \frac{\sup_{\mu \in \mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon, T, \delta)}{|\log(\epsilon)|} &\geqslant \frac{S(\mathcal{X}, d, 2\epsilon) - 3\delta S(\mathcal{X}, d, \epsilon)}{|\log(\epsilon)|} \\ \limsup_{\epsilon \to 0} \frac{\sup_{\mu \in \mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon, T, \delta)}{|\log(\epsilon)|} &\geqslant \limsup_{\epsilon \to 0} \left(\frac{S(\mathcal{X}, d, 2\epsilon) - 3\delta S(\mathcal{X}, d, \epsilon)}{|\log(\epsilon)|} \right). \end{aligned}$$

Escolhemos $(\epsilon_k)_{k \in \mathbb{N}}$ uma sequência tal que

$$\lim_{k \to \infty} \frac{S(\mathcal{X}, d, 2\epsilon_k)}{|\log(\epsilon_k)|} = \limsup_{\epsilon \to 0} \frac{S(\mathcal{X}, d, 2\epsilon)}{|\log(\epsilon)|}.$$

Além disso, considerando $m = 2\epsilon$, temos que

$$\limsup_{m\to 0} \frac{S(\mathcal{X},d,m)}{|\log(\frac{m}{2})|} = \limsup_{m\to 0} \frac{S(\mathcal{X},d,m)}{|\log(m)-\log(2)|} = \overline{mdim}(\mathcal{X},d,T).$$

Então,

$$\limsup_{\epsilon\to 0} \frac{\sup_\mu h_\mu(\epsilon,T,\delta)}{|\log(\epsilon)|} \geqslant \lim_{k\to\infty} \frac{S(\mathcal{X},d,2\epsilon_k)}{|\log(\epsilon_k)|} - \limsup_{k\to\infty} 3\delta \cdot \frac{S(\mathcal{X},d,\epsilon_k)}{|\log(\epsilon_k)|}.$$

Denotando $S(\mathcal{X},d,\epsilon) = S(\epsilon)$ e fazendo $\delta \to 0$ obtemos

$$\begin{aligned}
\lim_{\delta\to 0}\limsup_{\epsilon\to 0} \frac{\sup_\mu h_\mu(\epsilon,T,\delta)}{|\log(\epsilon)|} &\geqslant \lim_{\delta\to 0}\lim_{k\to\infty} \frac{S(2\epsilon_k)}{|\log(\epsilon_k)|} - \lim_{\delta\to 0} 3\delta \cdot \limsup_{k\to\infty} \frac{S(\epsilon_k)}{|\log(\epsilon_k)|} \\
&\geqslant \lim_{\delta\to 0}\lim_{k\to\infty} \frac{S(2\epsilon_k)}{|\log(\epsilon_k)|} - \lim_{\delta\to 0} 3\delta \cdot \limsup_{\epsilon\to 0} \frac{S(\epsilon)}{|\log(\epsilon)|} \\
&= \lim_{\delta\to 0} \overline{mdim}(\mathcal{X},d,T) - \lim_{\delta\to 0} 3\delta \cdot \overline{mdim}(\mathcal{X},d,T) \\
&= \overline{mdim}(\mathcal{X},d,T).
\end{aligned}$$

Por outro lado, $\forall \epsilon > 0, \delta \in (0,1)$ e $\mu \in \mathcal{M}_T(\mathcal{X})$, obtemos a seguinte desigualdade:

$$N_\mu(n,\epsilon,\delta) \leqslant N_d(n,\epsilon).$$

Consequentemente,

$$\sup_{\mu\in\mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon,T,\delta) \leqslant S(\epsilon)$$

e

$$\limsup_{\epsilon\to 0} \frac{\sup_{\mu\in\mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon,T,\delta)}{|\log(\epsilon)|} \leqslant \limsup_{\epsilon\to 0} \frac{S(\epsilon)}{|\log(\epsilon)|}.$$

Fazendo $\delta \to 0$

$$\lim_{\delta \to 0} \limsup_{\epsilon \to 0} \frac{\sup_{\mu \in \mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon, T, \delta)}{|\log(\epsilon)|} \leqslant \lim_{\delta \to 0} \limsup_{\epsilon \to 0} \frac{S(\epsilon)}{|\log(\epsilon)|}.$$

Logo

$$\lim_{\delta \to 0} \limsup_{\epsilon \to 0} \frac{\sup_{\mu \in \mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon, T, \delta)}{|\log(\epsilon)|} \leqslant \overline{mdim}(\mathcal{X}, d, T).$$

Com isso, concluímos que

$$\lim_{\delta \to 0} \limsup_{\epsilon \to 0} \frac{\sup_{\mu \in \mathcal{M}_T(\mathcal{X})} h_\mu(\epsilon, T, \delta)}{|\log(\epsilon)|} = \overline{mdim}(\mathcal{X}, d, T).$$

□

Proposição 5.1.2. *A dimensão métrica média superior da aplicação shift (deslocamento) em* $[0,1]^{\mathbb{Z}}$ *com a métrica* $d_T(x,y) = \sum_{k \in \mathbb{Z}} \frac{1}{2^{|k|}} d(x_k, y_k)$ *é dada por*

$$\overline{mdim}([0,1]^{\mathbb{Z}}, d_T, T) = 1,$$

onde $x = (\dots, x_{-1}, x_0, x_1, \dots)$*,* $y = (\dots, y_{-1}, y_0, y_1, \dots)$ *e* $d = |x_k - y_k|$ *em* $[0,1]$*.*

Demonstração. Para $k \geqslant 1$ consideramos o conjunto $P_k = \{p_1, p_2, \dots, p_k\}$, onde $p_i = \frac{2i-1}{2k}$. Definimos λ_k como uma medida de probabilidade em $[0,1]$ tal que $\lambda_k(p_i) = \frac{1}{k}, \forall i \in \{1, 2, \dots, k\}$, definimos também a medida produto $\mu_k = (\lambda_k)^{\otimes \mathbb{Z}}$. Agora, consideramos o conjunto

$$A_{i_0, i_1, \dots, i_{n-1}} = \left\{ x \in [0,1]^{\mathbb{Z}} : x_0 = p_{i_0}, \dots, x_{n-1} = p_{i_{n-1}} \right\}.$$

Para prosseguir a demonstração precisamos do seguinte lema:

Lema 5.1.3. *Seja $r < \frac{1}{2k}$ e $q \in [0,1]^{\mathbb{Z}}$. Então existe um único conjunto $A_{i_0,i_1,\ldots,i_{n-1}}$ tal que*

$$supp(\mu_k) \cap B_n(q,r) \subset A_{i_0,i_1,\ldots,i_{n-1}}.$$

Demonstração. Seja $x \in supp(\mu_k) \cap B_n(q,r)$. Logo, por definição, temos que

$$d(x_j, y_j) \leqslant d_T(T^j x, T^j q) \leqslant d_n(x,q) < r < \frac{1}{2k}, \forall j \in \{0, 1, \ldots, n-1\}.$$

De fato, como $B_n(q,r) = \{x \in [0,1]^{\mathbb{Z}} : d_n(x,q) < r\}$, onde

$$d_n(x,q) = \max_{0 \leqslant j < n} \left\{ d_T(T^j x, T^j q) \right\},$$

obtemos que $|x_j - q_j| < r, \forall j$, pois o primeiro termo da soma de $d_T(T^j x, T^j q)$ é $|x_j - q_j|$ e o restante dos termos são positivos. Logo

$$|x_j - q_j| = d(x_j, q_j) \leqslant d_T(T^j x, T^j q) \leqslant d_n(x,q) < r.$$

Como $x \in supp(\mu_k)$ e μ_k é um produto de medidas, nenhuma das parcelas da multiplicação deve ser nula, logo $x_j \in P_k, \forall j \in \mathbb{Z}$, em particular para $j \in \{0, 1, \ldots, n-1\}$. Como, pela definição de P_k sabemos que $d(p_i, p_j) \geqslant \frac{1}{2k}, \forall i \neq j$, isto implica que só temos uma escolha para x_j em P_k, denotada por p_{i_j}, provando a unicidade de $A_{i_0,i_1,\ldots,i_{n-1}}$ e concluindo a demonstração do lema.

□

Continuando a demonstração da proposição, observamos que, em particular, o lema prova que

$$\mu_k(B_n(q,r)) \leqslant \mu_k(A_{i_0,i_1,\ldots,i_{n-1}}) = \frac{1}{k^n}, \forall q \in [0,1]^{\mathbb{Z}}.$$

Logo, se A é um conjunto tal que $\mu_k(A) > 1 - \delta$ e

$$A \subset \bigcup_{i=1}^{L} B_n(z_i, r),$$

ou seja, A é coberto por L bolas dinâmicas, então

$$1 - \delta < \mu_k(A) \leqslant \mu_k\left(\bigcup_{i=1}^{L} B_n(z_i, r)\right) \leqslant \sum_{i=1}^{L} B_n(z_i, r) \leqslant \frac{L}{k^n},$$

implicando que $N_{\mu_k}(n, r, \delta) > (1-\delta)k^n, \forall n \geqslant 1, \delta \in (0,1)$. Portanto, pelo Lema, obtemos que para $\delta \in (0,1)$ fixado e $r < 1/2k$

$$\begin{aligned}
h_{\mu_k}(r, T, \delta) &= \limsup_{n\to\infty} \frac{1}{n} \log N_{\mu_k}(n, r, \delta) \\
&\geqslant \limsup_{n\to\infty} \frac{1}{n} \log\left((1-\delta)k^n\right) \\
&= \limsup_{n\to\infty} \frac{1}{n} \left(\log(1-\delta) + \log(k^n)\right) \\
&= \lim_{n\to\infty} \frac{1}{n} \log(1-\delta) + \limsup_{n\to\infty} \frac{1}{n} \log(k^n) \\
&= \limsup_{n\to\infty} \log(k) = \log(k).
\end{aligned}$$

Então, considerando a sequência $(r_k)_{k\in\mathbb{N}}$ tal que $r_k = 1/3k$, obtemos

$$\limsup_{r\to 0} \frac{\sup_\mu h_\mu(r, T, \delta)}{|\log(r)|} \geqslant \limsup_{k\to\infty} \frac{h_{\mu_k}(\frac{1}{3k}, T, \delta)}{\left|\log\left(\frac{1}{3k}\right)\right|} \geqslant \lim_{k\to\infty} \frac{\log(k)}{\log(3k)} = 1.$$

Portanto, pelo Teorema anterior, concluímos que

$$\overline{mdim}([0,1]^{\mathbb{Z}}, d_T, T) \geqslant 1.$$

Para mostrar a desigualdade oposta, utilizamos a ideia da demonstração encontrada em [5]. Inicialmente, como $N_d(n, \epsilon) \leqslant C_d(n, \epsilon)$ obtemos

$$\begin{aligned}
S(\epsilon) &= \limsup_{n\to\infty} \frac{1}{n} \log N_d(n, \epsilon) \\
&\leqslant \limsup_{n\to\infty} \frac{1}{n} \log C_d(n, \epsilon).
\end{aligned}$$

Agora, dado $\epsilon > 0$ e $l = \lceil \log_2(4/\epsilon) \rceil$, onde $\lceil \log_2(4/\epsilon) \rceil$ representa o menor inteiro maior ou igual que $\log_2(4/\epsilon)$, observamos que

$$\sum_{|n|>l} \frac{1}{2^{|n|}} \leqslant \frac{\epsilon}{2}.$$

De fato,

$$\begin{aligned}
\sum_{|n|>l} \frac{1}{2^{|n|}} &= 2\left(\sum_{n=l+1}^{\infty} \frac{1}{2^{|n|}}\right) \\
&= 2\left(\frac{1}{2^{l+1}} \cdot \frac{1}{2^{-1}}\right) \\
&= 2 \cdot \frac{1}{2^l} \\
&\leqslant \frac{2\epsilon}{4} = \frac{\epsilon}{2}.
\end{aligned}$$

Considerando a cobertura por abertos de $[0,1]$ dada por

$$I_k = \left(\frac{(k-1)\epsilon}{12}, \frac{(k+1)\epsilon}{12}\right), 0 \leqslant k \leqslant \lfloor 12/\epsilon \rfloor,$$

onde $\lfloor 12/\epsilon \rfloor$ representa o maior inteiro menor ou igual a $12/\epsilon$ e cada intervalo I_k tem comprimento igual a $\epsilon/6$. Então, para $n \geqslant 1$, consideramos a cobertura de $[0,1]^{\mathbb{Z}}$ dada por

$$[0,1]^{\mathbb{Z}} = \bigcup_{0 \leqslant k_{-l}, \ldots, k_{l+n}} \left\{x : x_{-l} \in I_{k_{-l}}, \ldots, x_{l+n} \in I_{k_{l+n}}\right\}.$$

Observamos que cada aberto desta cobertura tem diâmetro com respeito a métrica d_n menor que ϵ. De fato, para $|n| > l$ sabemos que $\sum_{|n|>l} 2^{-|n|} \leqslant \frac{\epsilon}{2}$. Para $|n| \leqslant l$,

$$
\begin{aligned}
\sum_{|n|\leqslant l} \frac{1}{2^{|n|}} &= 2\left(\sum_{n=1}^{l} \frac{1}{2^n}\right) + 1 \\
&= 2\left(\frac{1}{2}\cdot\frac{2^{-l}-1}{-2^{-1}}\right) + 1 \\
&= 2\left(1 - \frac{1}{2^l}\right) + 1 \\
&\leqslant 2 + 1 = 3.
\end{aligned}
$$

Como o supremo entre a distância de dois elementos de um mesmo intervalo I_k é $\epsilon/6$, obtemos que

$$
\sum_{|n|\leqslant l} \frac{1}{2^{|n|}} \cdot \frac{\epsilon}{6} \leqslant \frac{\epsilon}{2}
$$

e, portanto, o diâmetro com respeito a d_n dos abertos desta cobertura é menor ou igual que ϵ. Logo,

$$
C_d(n,\epsilon) \leqslant (1 + \lfloor 12/\epsilon \rfloor)^{n+2l+1} \leqslant (1 + 12/\epsilon)^{n+2l+1},
$$

implicando que

$$
\begin{aligned}
S(\epsilon) &\leqslant \limsup_{n\to\infty} \frac{1}{n} \log\left(1 + 12/\epsilon\right)^{n+2l+1} \\
&= \limsup_{n\to\infty} \frac{1}{n} \cdot \frac{n+2l+1}{1} \log\left(1 + 12/\epsilon\right) \\
&= \log\left(1 + 12/\epsilon\right).
\end{aligned}
$$

Fazendo $\epsilon \to 0$, obtemos

$$
\begin{aligned}
\lim_{\epsilon\to 0} \log\left(1 + 12/\epsilon\right) &= \lim_{\epsilon\to 0} \log\left(12/\epsilon\right) \\
&= \lim_{\epsilon\to 0} \left(\log(12) - \log(\epsilon)\right) \\
&= \lim_{\epsilon\to 0} -\log(\epsilon) = \lim_{\epsilon\to 0} |\log(\epsilon)|,
\end{aligned}
$$

para ϵ suficientemente pequeno.

Portanto,

$$\begin{aligned}\overline{mdim}([0,1]^{\mathbb{Z}}, d_T, T) &= \limsup_{\epsilon\to 0} \frac{S(\epsilon)}{|\log(\epsilon)|} \\ &\leqslant \limsup_{\epsilon\to 0} \frac{\log(1+12/\epsilon)}{|\log(\epsilon)|} = 1\end{aligned}$$

e, consequentemente, $\overline{mdim}([0,1]^{\mathbb{Z}}, d_T, T) = 1$.

$\square$

Para $\left([0,1]^d\right)^{\mathbb{Z}}$, basta seguir um desenvolvimento análogo ao caso $[0,1]^{\mathbb{Z}}$ para obter

$$\overline{mdim}\left(\left([0,1]^d\right)^{\mathbb{Z}}, d_T, T\right) = d$$

ou utilizar a próxima proposição para isso.

O próximo resultado é uma generalização do resultado anterior para espaços métricos mais gerais, envolvendo a Dimensão de Minkowski superior ou *Upper Box Dimension*, definida por

$$\overline{dim}_B(Y,d) = \limsup_{\epsilon\to 0} \frac{\log N(\epsilon)}{|\log(\epsilon)|},$$

onde $N(\epsilon)$ denota a cardinalidade máxima de um conjunto ϵ-separado em (Y,d).

Proposição 5.1.4.

$$\overline{mdim}(Y^{\mathbb{Z}}, d_T, T) = \overline{dim}_B(Y,d).$$

Demonstração. Seja Y um espaço métrico compacto com a métrica d. Consideramos $(\epsilon_k)_{k\in\mathbb{N}}$ uma sequência decrescente convergindo a zero tal que

$$\lim_{k\to\infty} \frac{\log N(\epsilon_k)}{|\log(\epsilon_k)|} = \overline{dim}_B(Y,d).$$

Seja $P_k = \{x_1, x_2, \ldots, x_{N(\epsilon_k)}\}$ um conjunto de pontos ϵ_k-separados em Y e λ_k uma medida de probabilidade que dá pesos $1/N(\epsilon_k)$ a cada elemento de P_k. Pelo lema anterior, obtemos que para $r < \epsilon_k/2$

$$\mu_k(B_n(q,r)) < \frac{1}{N(\epsilon_k)^n}, \ \forall q \in Y^{\mathbb{Z}}.$$

Logo, para todo $\delta \in (0,1)$ consideramos a sequência $(r_k)_{k\in\mathbb{N}}$ onde $r_k = \epsilon_k/3$ e o supremo em relação as medidas $\mu \in \mathcal{M}_T(Y^{\mathbb{Z}})$, obtemos

$$\begin{aligned}
\limsup_{r\to 0} \frac{\sup_\mu h_\mu(r,T,\delta)}{|\log(r)|} &\geqslant \limsup_{r\to 0} \frac{h_{\mu_k}(r,T,\delta)}{|\log(r)|} \\
&\geqslant \limsup_{k\to\infty} \frac{h_{\mu_k}(\epsilon_k/3,T,\delta)}{|\log(\epsilon_k/3)|} \\
&\geqslant \limsup_{k\to\infty} \frac{\limsup_{n\to\infty} \dfrac{1}{n}\log\left((1-\delta)N(\epsilon_k)^n\right)}{|\log(\epsilon_k/3)} \\
&= \limsup_{k\to\infty} \frac{\lim_n \frac{1}{n}\log(1-\delta) + \limsup_n \frac{1}{n}\log(N(\epsilon_k)^n)}{|\log(\epsilon_k/3)|} \\
&= \limsup_{k\to\infty} \frac{\log(N(\epsilon_k))}{|\log(\epsilon_k/3)|} \\
&= \limsup_{k\to\infty} \frac{\log(N(\epsilon_k))}{|\log(\epsilon_k)-\log(3)|} \\
&= \lim_{k\to\infty} \frac{\log(N(\epsilon_k))}{|\log(\epsilon_k)|} = \overline{dim}_B(Y,d).
\end{aligned}$$

Portanto, $\overline{mdim}(Y^{\mathbb{Z}}, d_T, T) \geqslant \overline{dim}_B(Y,d)$.

Por outro lado, considerando $\epsilon \in (0,1)$, $l \in \mathbb{N}$ tal que

$$\sum_{|j|\geqslant l} 2^{-|j|} = \frac{1}{2^{l-2}} < \frac{\epsilon}{2\cdot diam(Y)}$$

e M a cardinalidade máxima de um conjunto ϵ-separado $B = \{x_i\}_{i=1}^M$ em Y. Supondo B um conjunto ϵ-separado de cardinalidade máxima, temos que $Y = \bigcup_{i=1}^M B(x_i,\epsilon)$. Logo, podemos considerar a partição de Y dada pelos conjuntos

$$\begin{aligned}
A_1 &= B(x_1, \epsilon)\\
A_2 &= B(x_2, \epsilon) \setminus B(x_1, \epsilon)\\
A_3 &= B(x_3, \epsilon) \setminus B(x_1, \epsilon) \cup B(x_2, \epsilon)\\
&\vdots\\
A_M &= B(x_M, \epsilon) \setminus \bigcup_{i=1}^{M-1} B(x_i, \epsilon).
\end{aligned}$$

Consideramos a aplicação $f : Y \to Y$ definida por $f(x) = x_i$ onde i representa o conjunto A_i tal que $x \in A_i$. Como cada A_i é mensurável e a união de mensuráveis é mensurável, temos que f é mensurável.

Agora, consideramos o conjunto

$$S_{i_{-l},\ldots,i_{n+l}} = \left\{ y \in Y^{\mathbb{Z}} : y_j \in A_{i_j} \text{ para todo } -l \leqslant j \leqslant n+l \right\},$$

onde $i_j \in \{1, 2, \ldots, M\}$ para cada j. Consequentemente, para cada $z, y \in S_{i_{-l},\ldots,i_{n+l}}$, obtemos que

$$\begin{aligned}
d(T^i z, T^i y) &\leqslant diam(Y) \cdot \sum_{|j| \geqslant l} \frac{1}{2^{|j|}} + \sum_{|j| \leqslant l} \frac{1}{2^{|j|}} \cdot d(z_{i+j}, y_{i+j})\\
&< \frac{\epsilon}{2} + 2\epsilon \sum_{|j| \leqslant l} \frac{1}{2^{|j|}}\\
&< 7\epsilon,
\end{aligned}$$

para todo $i \in \{0, 1, \ldots, n\}$, pois $diam(A_i) \leqslant 2\epsilon\ \forall i$ e

$$\sum_{|j| \leqslant l} \frac{1}{2^{|j|}} = 2 \sum_{i=1}^{l} \frac{1}{2^j} + 1 = 2(1 - 2^{-l}) + 1 = 3 - 2^{-l} < 3.$$

Concluímos então que $diam(S_{i_{-l},\ldots,i_{n+l}}) < 7\epsilon$ com respeito a métrica d_n. Além disso,

$$Y^{\mathbb{Z}} = \bigcup_{i_{-l},\ldots,i_{n+l}} S_{i_{-l},\ldots,i_{n+l}}.$$

Então, obtemos a seguinte desigualdade

$$N_{d_T}(n, 7\epsilon) \leqslant M^{n+2l+1} \leqslant N(\epsilon)^{n+2l+1}.$$

Considerando $m = 7\epsilon$ e $\delta \in (0,1)$, temos que $N_\mu(n, m, \delta) \leqslant C_d(n, m)$ para toda medida μ. Logo

$$\begin{aligned}
\limsup_{m\to 0} \frac{\sup_\mu h_\mu(m, T, \delta)}{|\log(m)|} &\leqslant \limsup_{m\to 0} \frac{\lim_{n\to\infty} \frac{1}{n} \log C_d(n, m)}{|\log(m)|} \\
&\leqslant \limsup_{\epsilon\to 0} \frac{\lim_{n\to\infty} \frac{1}{n} \log(N(\epsilon)^{n+2l+1})}{|\log(7\epsilon)|} \\
&= \limsup_{\epsilon\to 0} \frac{\log N(\epsilon)}{|\log(7\epsilon)|} \\
&= \limsup_{\epsilon\to 0} \frac{\log N(\epsilon)}{|\log(7) - \log(\epsilon)|} \\
&= \limsup_{\epsilon\to 0} \frac{\log N(\epsilon)}{|\log(\epsilon)|} = \overline{dim}_B(Y, d).
\end{aligned}$$

Portanto, concluímos que

$$\overline{mdim}(Y^{\mathbb{Z}}, d_T, T) \leqslant \overline{dim}_B(Y, d).$$

Logo,

$$\overline{mdim}(Y^{\mathbb{Z}}, d_T, T) = \overline{dim}_B(Y, d).$$

□

5.2 Função Taxa de Distorção e $\tilde{h}_\mu(\epsilon, T, \delta)$

Nesta seção vamos relacionar a função taxa de distorção $R_\mu(\epsilon)$ e a função $\tilde{h}_\mu(\epsilon, T, \delta)$. Além disso, vamos utilizar os resultados desta seção para provar o princípio variacional para a dimensão métrica média que E. Lindenstrauss e M. Tsukamoto estabeleceram em [5], onde relacionaram a dimensão métrica média com a função taxa de distorção.

Antes de definirmos a função taxa de distorção $R_\mu(\epsilon)$, vamos definir a informação mútua entre duas varáveis aleatórias que possuam imagem finita. Sejam Z_1 e Z_2 duas variáveis aleatórias com imagem finita agindo em $\mathcal{X}$, denotamos por $H_\mu(Z_i)$ a entropia da partição pré-imagem de Z_i com respeito à medida μ, para $i = \{1, 2\}$. Definimos a informação mútua de Z_1 e Z_2 por

$$I_\mu(Z_1, Z_2) = H_\mu(Z_1) + H_\mu(Z_2) - H_\mu(Z_1 \vee Z_2),$$

onde $H_\mu(Z_1 \vee Z_2)$ denota a entropia do refinamento das partições pré-imagem de Z_1 e Z_2.

Definição 5.2.1. *Dizemos que o par (X, Y) satisfaz a condição $(*)_{n,\epsilon}$ se as seguintes propriedades são satisfeitas:*

1. *Existe um espaço de probabilidade $(\Omega, \mathbb{P})$ tal que $X : \Omega \to \mathcal{X}$ e $Y = (Y_0, Y_1, \ldots, Y_{n-1}) : \Omega \to \mathcal{X}^n$ são aplicações mensuráveis;*
2. *A medida $X_*\mathbb{P}$ é uma medida de probabilidade T-invariante em $\mathcal{X}$;*
3. $\mathbb{E}\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^k X, Y_k)\right) \leqslant \epsilon.$

Dizemos que o par (X, Y) satisfaz a condição $(*)_{n,\epsilon,\mu}$ se, além disso, temos $X_*\mathbb{P} = \mu$.

Definição 5.2.2. *Dada uma medida de probabilidade T-invariante μ em $\mathcal{X}$, definimos a função taxa de distorção como*

$$R_\mu(\epsilon) = \inf \frac{1}{n} I_\mu(X, Y),$$

onde o ínfimo é tomado sobre todos os pares (X, Y) satisfazendo a condição $(*)_{n,\epsilon,\mu}$.

Agora, considerando a aplicação $T : \mathcal{X} \to \mathcal{X}$ e uma métrica d podemos definir mais uma nova família de métricas em $\mathcal{X}$, semelhante a métrica d_n, porém, usando a média da distância entre os n primeiros iterados ao invés da distância máxima entre eles.

Então, definimos

$$\tilde{d}_n(x,y) = \frac{1}{n}\sum_{k=0}^{n-1} d(T^k(x), T^k(y)).$$

Denotamos uma bola de raio ϵ na métrica $\tilde{d}_n$ como uma (n,ϵ)-bola dinâmica média.

Analogamente a Definição 4.0.1, onde usamos a métrica d_n para definir $h_\mu(\epsilon, T, \delta)$, a próxima definição é semelhante, porém, utilizamos a métrica $\tilde{d}_n$.

Definição 5.2.3. *Seja μ uma medida de probabilidade T-invariante em $\mathcal{X}$. Definimos*

$$\tilde{h}_\mu(\epsilon, T, \delta) = \limsup_{n\to\infty} \frac{1}{n}\log \tilde{N}_\mu(n,\epsilon,\delta),$$

onde $\tilde{N}_\mu(n,\epsilon,\delta)$ é o número mínimo de (n,ϵ)-bolas dinâmicas médias necessárias para cobrir um conjunto com medida estritamente maior que $1-\delta$ com respeito à μ. Definimos também a entropia métrica média com respeito à μ como sendo o limite

$$\tilde{h}_\mu(T,\delta) = \lim_{\epsilon\to 0} \tilde{h}_\mu(\epsilon, T, \delta).$$

Proposição 5.2.4. *Seja μ uma medida de probabilidade ergódica T-invariante e $(\mathcal{X}, d)$ um espaço métrico compacto com diâmetro D. Então*

$$R_\mu(\epsilon) \leqslant \tilde{h}_\mu(\epsilon/2, T, \epsilon/2D).$$

Se $\mathcal{P}$ é uma partição com diâmetro menor que ϵ. Então

$$R_\mu(\epsilon) \leqslant h_\mu(\mathcal{P}, T).$$

Demonstração. Por definição, temos que

$$R_\mu(\epsilon) = \inf \frac{1}{n} I_\mu(X,Y),$$

onde (X,Y) satisfazem a condição $(*)_{n,\epsilon,\mu}$.

Considerando $K \subset \mathcal{X}$ tal que $\tilde{N}(K, n, \epsilon/2) = \tilde{N}(n, \epsilon/2, \epsilon/2D) = N$, ou seja, o número mínimo de $(n, \epsilon/2)$-bolas dinâmicas médias necessárias para cobrir K é igual ao número mínimo de $(n, \epsilon/2)$-bolas dinâmicas médias necessárias para cobrir um conjunto de medida μ maior que $1 - \epsilon/2D$.

Logo, existem pontos $\{x_i\}_{i=1}^N$ em $\mathcal{X}$ tais que

$$K \subset \bigcup_{i=1}^{N} \tilde{B}_n \left(x_i, \frac{\epsilon}{2} \right)$$

e

$$\mu(K) > 1 - \frac{\epsilon}{2D}.$$

Além disso, consideramos uma partição de K $\{A_1, A_2, \ldots, A_N\}$, construída da seguinte forma:

$$\begin{aligned}
A_1 &= \tilde{B}_n \left(x_1, \frac{\epsilon}{2} \right) \cap K \\
A_2 &= \left(\tilde{B}_n \left(x_2, \frac{\epsilon}{2} \right) \cap K \right) \setminus \tilde{B}_n \left(x_1, \frac{\epsilon}{2} \right) \\
&\vdots \\
A_N &= \left(\tilde{B}_n \left(x_N, \frac{\epsilon}{2} \right) \cap K \right) \setminus \left(\tilde{B}_n \left(x_2, \frac{\epsilon}{2} \right) \cup \cdots \cup \tilde{B}_n \left(x_{N-1}, \frac{\epsilon}{2} \right) \right)
\end{aligned}$$

Então, considerando $A_{N+1} = \mathcal{X} \setminus \bigcup_{i=1}^N A_i$, obtemos uma partição de $\mathcal{X}$ com $N+1$ elementos.

Agora, definimos uma aplicação mensurável $f : \mathcal{X} \to \mathcal{X}$, análoga a definida na proposição anterior, onde $f(x) = x_i$ se $x \in A_i$ e $i \in \{1, 2, \ldots, N\}$ e $f(x) = p \in \mathcal{X} \setminus K$ se $x \in A_{N+1}$. Além disso, observamos que para $x \in K$, $\tilde{d}_n(x, f(x)) < \epsilon/2$ e, por construção, obtemos também que $|f(x)| = N + 1$.

Definimos $Y = (f(x), T(f(x)), T^2(f(x)), \ldots, T^{n-1}(f(x)))$. Como $f(x)$ pode tomar somente $N+1$ valores, obtemos também que $|Im(Y)| = N + 1$.

Logo

$$
\begin{aligned}
\mathbb{E}\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^k X, Y_k)\right) &= \int_{\mathcal{X}} \frac{1}{n}\sum_{k=0}^{n-1} d(T^k(x), T^k(f(x))d\mu(x) \\
&= \int_{K} \frac{1}{n}\sum_{k=0}^{n-1} d(T^k(x), T^k(f(x))d\mu(x) \\
&+ \int_{\mathcal{X}\setminus K} \frac{1}{n}\sum_{k=0}^{n-1} d(T^k(x), T^k(f(x))d\mu(x) \\
&\leqslant \int_{K} \frac{1}{n}\sum_{k=0}^{n-1} d(T^k(x), T^k(f(x))d\mu(x) + \mu(\mathcal{X}\setminus K)D,
\end{aligned}
$$

pois, como $d(T^k(x), T^k(f(x)) \leqslant D$, concluímos que

$$
\begin{aligned}
\int_{\mathcal{X}\setminus K} \frac{1}{n}\sum_{k=0}^{n-1} d(T^k(x), T^k(f(x))d\mu(x) &\leqslant \int_{\mathcal{X}\setminus K} Dd\mu(x) \\
&= D\int_{\mathcal{X}\setminus K} d\mu(x) = D\mu(\mathcal{X}\setminus K).
\end{aligned}
$$

Para a integral em K, sabendo que $\tilde{d}_n(x, f(x)) < \epsilon/2$, obtemos a seguinte desigualdade

$$
\begin{aligned}
\int_{K} \frac{1}{n}\sum_{k=0}^{n-1} d(T^k(x), T^k(f(x))d\mu(x) &\leqslant \int_{K} \frac{\epsilon}{2}d\mu(x) \\
&= \frac{\epsilon}{2}\mu(K) \\
&\leqslant \frac{\epsilon}{2}.
\end{aligned}
$$

Portanto

$$
\mathbb{E}\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^k X, Y_k)\right) \leqslant \frac{\epsilon}{2} + \frac{\epsilon}{2D}\cdot D = \epsilon,
$$

pois $\mu(K) > 1 - \epsilon/2D$ implica que $\mu(\mathcal{X} \setminus K) < \epsilon/2D$.

Agora, a partir da definição da informação mútua, obtemos as seguintes desigualdades:

$$\begin{aligned}
I_\mu(X,Y) &= H_\mu(X) + H_\mu(Y) - H_\mu(X \vee Y) \\
&= H_\mu(X) + H_\mu(Y) - (H_\mu(X) + H_\mu(Y|X)) \\
&= H_\mu(Y) - H_\mu(Y|X) \\
&\leqslant H_\mu(Y) \\
&\leqslant \log(|f(x)|) \\
&= \log(N+1) \\
&= \log\left(\tilde{N}_\mu\left(n, \frac{\epsilon}{2}, \frac{\epsilon}{2D}\right) + 1\right).
\end{aligned}$$

Logo

$$\begin{aligned}
R_\mu(\epsilon) &= \inf \frac{1}{n} I_\mu(X,Y) \\
&\leqslant \liminf_{n\to\infty} \frac{1}{n} I_\mu(X,Y) \\
&\leqslant \liminf_{n\to\infty} \frac{1}{n} \log\left(\tilde{N}_\mu\left(n, \frac{\epsilon}{2}, \frac{\epsilon}{2D}\right)\right) \\
&\leqslant \limsup_{n\to\infty} \frac{1}{n} \log\left(\tilde{N}_\mu\left(n, \frac{\epsilon}{2}, \frac{\epsilon}{2D}\right)\right) \\
&= \tilde{h}_\mu\left(\frac{\epsilon}{2}, T, \frac{\epsilon}{2D}\right).
\end{aligned}$$

Relacionando as duas métricas d_n e $\tilde{d}_n$, observamos que se o máximo entre as distâncias dos iterados é menor do que ϵ, ou seja, $d_n(x,y) < \epsilon$, então a distância média dos iterados também é menor do que ϵ, ou seja, $\tilde{d}_n(x,y) < \epsilon$.

Portanto, concluímos que

$$B_n(x,\epsilon) \subset \tilde{B}_n(x,\epsilon).$$

Então, considerando uma cobertura por bolas dinâmicas para um determinado conjunto, isto implica que podemos considerar uma cobertura para este mesmo conjunto por meio de bolas dinâmicas médias centradas nos mesmos pontos e com mesmo raio. Consequentemente $N_\mu(n,\epsilon,\delta) \geqslant \tilde{N}_\mu(n,\epsilon,\delta)$ e $h_\mu(\epsilon/2, T, \epsilon/2D) \geqslant \tilde{h}_\mu(\epsilon/2, T, \epsilon/2D)$.

Pelo Lema 4.0.2, temos

$$h_\mu(\mathcal{P}, T) \geqslant \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n,\epsilon,\delta),$$

para qualquer partição com diâmetro menor que ϵ e qualquer $\delta \in (0,1)$.

Com isso, concluímos que

$$R_\mu(\epsilon) \leqslant h_\mu(\epsilon/2, T, \epsilon/2D) \leqslant h_\mu(\mathcal{P}, T),$$

onde $\mathcal{P}$ é uma partição com diâmetro menor do que ϵ.

□

A demonstração do próximo teorema pode ser encontrado em [14].

Teorema 5.2.5. *Seja $X : (\Omega, \mathbb{P}) \to (\mathcal{X}, \mu)$ uma aplicação mensurável com $X_*\mathbb{P} = \mu$ onde μ é uma medida de probabilidade ergódica T-invariante em $\mathcal{X}$. Dado $\epsilon_0 > 0$, existe $n(\epsilon_0) \in \mathbb{N}$ tal que para cada $n > n(\epsilon_0)$, existe uma função mensurável $f_n : \mathcal{X} \to \mathcal{X}^n$ tal que o par $(X, f_n \circ X)$ satisfaz a condição $(*)_{n,\epsilon+\epsilon_0}$ e $|f_n(\mathcal{X})| \leqslant e^{n(R_\mu(\epsilon)+\epsilon_0)}$.*

Definição 5.2.6. *Uma aplicação $Z : \mathcal{X} \to \mathcal{X}^n$ é uma (n,ϵ)-aproximação de $(\mathcal{X}, T)$ se é uma aplicação mensurável com imagem finita e satisfaz*

$$\int_{\mathcal{X}} \frac{1}{n} \sum_{k=0}^{n-1} d(T^k(x), Z_k(x)) d\mu(x) \leqslant \epsilon.$$

Lema 5.2.7. *Dada Z uma (n,ϵ)-aproximação de $(\mathcal{X}, T)$, podemos encontrar uma $(n, 2\epsilon)$-aproximação de $(\mathcal{X}, T)$ $Z' = (Z'_0, \ldots, Z'_{n-1})$ tal que $Z'_k = T^k Z'_0$. Além disso, as partições geradas pelas pré-imagens de Z e Z' coincidem.*

Demonstração. Como Z é uma (n,ϵ)-aproximação de $(\mathcal{X}, T)$, temos que $Im(Z) = \{z_1, z_2, \ldots, z_T\}$, onde $z_j \in \mathcal{X}^n$, $\forall j \in \{1, \ldots, T\}$, pois Z tem imagem finita.

Logo, denotamos por $\mathcal{Q} = \{Q_1, \ldots, Q_T\}$ a partição pré-imagem de Z, ou seja, a partição tal que

$$Z_{|Q_j} \equiv z_j = (z_{j,0}, \ldots, z_{j,n-1}) \in \mathcal{X}^n.$$

Então, obtemos que

$$\int_{\mathcal{X}} \frac{1}{n} \sum_{k=0}^{n-1} d(T^k(x), Z_k(x)) d\mu(x) = \sum_{k=1}^{T} \int_{Q_k} \frac{1}{n} \sum_{i=0}^{n-1} d(T^i(x), z_{k,i}) d\mu(x),$$

Para simplificar a notação, definimos a função $g_k : Q_k \to \mathbb{R}$ dada por

$$g_k(x) = \frac{1}{n} \sum_{i=0}^{n-1} d(T^i(x), z_{k,i}).$$

Usando o fato de Z ser uma (n, ϵ)-aproximação, obtemos que

$$\sum_{k=1}^{T} \int_{Q_k} \frac{1}{n} \sum_{i=0}^{n-1} d(T^i(x), z_{k,i}) d\mu(x) = \sum_{k=1}^{T} \int_{Q_k} g_k(x) d\mu(x) \leqslant \epsilon.$$

Considerando $a_k = \int_{Q_k} g_k(x) d\mu(x)$, analisamos os casos $a_k \neq 0$ e $a_k = 0$. Então, supondo $a_k \neq 0$ para algum $k \in \{1, \ldots, T\}$, afirmamos que exite um ponto $x_k \in Q_k$ tal que

$$g_k(x_k) \leqslant \frac{a_k}{\mu(Q_k)}.$$

De fato, supondo por absurdo que

$$g_k(x_k) > \frac{a_k}{\mu(Q_k)}, \forall x \in Q_k,$$

então

$$
\begin{aligned}
\int_{Q_k} g_k(x)d\mu(x) &> \int_{Q_k} \frac{a_k}{\mu(Q_k)} d\mu(x) \\
&= \frac{a_k}{\mu(Q_k)} \int_{Q_k} d\mu(x) \\
&= \frac{a_k}{\mu(Q_k)} \cdot \mu(Q_k) = a_k,
\end{aligned}
$$

o que implica que $a_k > a_k$, absurdo. Logo, obtemos que

$$
g_k(x_k) \leqslant \frac{a_k}{\mu(Q_k)}.
$$

Agora, se $a_k = 0$, escolhemos x_k como sendo qualquer ponto de Q_k. Definimos

$$
Z_i'(x) = T^i(x_k), \forall x \in Q_k
$$

e

$$
\begin{aligned}
Z'(x) &= (Z_0', Z_1', \ldots, Z_{n-1}') \\
&= (x_k, T^i(x_k), \ldots, T^{n-1}(x_k)),
\end{aligned}
$$

sempre que $x \in Q_k$. Além disso, pela construção de Z', obtemos que a partição pré-imagem de Z e Z' coincidem.

Então,

$$
\int_{\mathcal{X}} \frac{1}{n} \sum_{i=0}^{n-1} d(T^i(x), Z_k'(x))d\mu(x) = \sum_{k=1}^{T} \int_{Q_k} \frac{1}{n} \sum_{i=0}^{n-1} d(T^i(x), T^i(x_k))d\mu(x). \quad (5.1)
$$

Usando a desigualdade triangular e a linearidade da integral, obtemos que 5.1 é menor um igual a

$$
\sum_{k=1}^{T} \left(\int_{Q_k} \frac{1}{n} \sum_{k=0}^{n-1} d(T^i(x), z_{k,i})d\mu(x) + \int_{Q_k} \frac{1}{n} \sum_{k=0}^{n-1} d(z_{k,i}, T^i(x_k))d\mu(x) \right),
$$

que por sua vez é igual a

$$\sum_{k=1}^{T} (a_k + \mu(Q_k) g_k(x_k)) . \tag{5.2}$$

Como $g_k(x_k) \leqslant \dfrac{a_k}{\mu(Q_k)}$, obtemos que 5.2 é menor ou igual a

$$\sum_{k=1}^{T} (2a_k) = 2 \cdot \sum_{k=1}^{T} (a_k) \leqslant 2\epsilon.$$

Logo, concluímos que Z' é uma $(n, 2\epsilon)$-aproximação para $(\mathcal{X}, T)$.

□

Proposição 5.2.8. *Seja μ uma medida de probabilidade ergódica T-invariante. Então*

$$\tilde{h}_\mu(4L\epsilon, T, 1/L) \leqslant R_\mu(\epsilon).$$

Demonstração. Fixado $\epsilon_0 > 0$ arbitrariamente pequeno. Para $n > n(\epsilon_0)$ consideramos a aplicação $f_n = (f_{n,0}, \ldots, f_{n,n-1}) : \mathcal{X} \to \mathcal{X}^n$ dada pelo Teorema 5.2.5.

Como $(X, f_n(X))$ satisfaz a condição $(*)_{n,\epsilon+\epsilon_0}$, onde $X : (\Omega, \mathbb{P}) \to (\mathcal{X}, \mu)$ é uma aplicação mensurável com $X_* \mathbb{P} = \mu$, obtemos que

$$\int_\Omega \frac{1}{n} \sum_{k=0}^{n-1} d(T^k \circ X, f_{n,k} \circ X) d\mathbb{P} = \int_{\mathcal{X}} \frac{1}{n} \sum_{k=0}^{n-1} d(T^k(x), f_{n,k}(x)) d\mu(x) \leqslant \epsilon + \epsilon_0 \tag{5.3}$$

De fato, considerando E um conjunto mensurável e $\mathcal{X}_E$ sua função característica, então

$$\int_{\mathcal{X}} \mathcal{X}_E \, d\mu = \mu(E) = \mathbb{P}(X^{-1}(E)) = \int_\Omega \mathcal{X}_E \circ X d\mathbb{P}.$$

Como a integral é um funcional linear, segue que

$$\int_{\mathcal{X}} s \, d\mu = \int_\Omega s \circ X d\mathbb{P} \tag{5.4}$$

para toda função simples $s = \sum_{j=1}^{k} \alpha_j \mathcal{X}_{A_j}$, onde $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ e $A_1, \dots, A_k$ são conjuntos mensuráveis disjuntos dois a dois.

Seja φ uma função mensurável não negativa e integrável, então existe uma sequência não decrescente de funções simples $0 \leqslant s_1 \leqslant s_2 \leqslant \cdots$ com $\lim_{n\to\infty} s_n(x) = \varphi(x), \forall x \in \mathcal{X}$ e, por definição de integral, $\int \varphi \; d\mu = \lim_{n\to\infty} \int s_n \; d\mu$. Além disso, obtemos também que $\lim_{n\to\infty} (s_n \circ X)(x) = (\varphi \circ X)(x)$, para todo $x \in \Omega$ e $0 \leqslant s_1 \circ X \leqslant s_2 \circ X \leqslant \cdots$. Logo, pelo Teorema da Convergência Monótona, temos que

$$\lim_{n\to\infty} \int_\Omega s_n \circ X \; d\mathbb{P} = \int_\Omega \varphi \circ X \; d\mathbb{P}.$$

Portanto, usando 5.4 concluímos que

$$\int_{\mathcal{X}} \varphi \; d\mu = \int_\Omega \varphi \circ X d\mathbb{P}$$

para toda função mensurável não negativa e integrável.

Como d é mensurável e não negativa, isto implica que vale 5.3.

Por meio desta igualdade, observamos que para todo $n > n(\epsilon_0)$ cada f_n é uma $(n, \epsilon + \epsilon_0)$-aproximação de $(\mathcal{X}, T)$ e, além disso, como $|f_n(\mathcal{X})| \leqslant e^{n(R_\mu(\epsilon)+\epsilon_0)}$, aplicando o logaritmo e dividindo por n esta desigualdade, obtemos que

$$\frac{1}{n} \log |f_n(\mathcal{X})| \leqslant R_\mu(\epsilon) + \epsilon_0.$$

Pelo Lema 5.2.7 , podemos encontrar aplicações $\tilde{f}_n : \mathcal{X} \to \mathcal{X}^n$ que são $(n, 2\epsilon + 2\epsilon_0)$-aproximações de $(\mathcal{X}, T)$ com $\tilde{f}_{n,k} = T^k \tilde{f}_{n,0}$. Definindo $g_n = \tilde{f}_{n,0}$, escrevemos $\tilde{f}_n$ como

$$\tilde{f}_n = (g_n, T g_n, \dots, T^{n-1} g_n).$$

Como as partições pré-imagem de f_n e $\tilde{f}_n$ coincidem, obtemos também que $H_\mu(f_n) = H_\mu(\tilde{f}_n)$ e $|f_n(\mathcal{X})| = |\tilde{f}_n(\mathcal{X})| = |g_n(\mathcal{X})|$, pois as coordenadas de $\tilde{f}_n$ dependem somente da primeira coordenada que é g_n. Logo, obtemos que

$$\frac{1}{n} \log |g_n(\mathcal{X})| \leqslant R_\mu(\epsilon) + \epsilon_0.$$

Antes de relacionar $R_\mu(\epsilon)$ com $\tilde{h}_\mu(4L\epsilon, T, 1/L)$, observamos que

$$\int_{\mathcal{X}} \tilde{d}_n(x, g_n(x)) d\mu(x) = \int_{\mathcal{X}} \frac{1}{n} \sum_{k=0}^{n-1} d(T^k(x), T^k g_n(x)) d\mu(x) \leqslant 2\epsilon + 2\epsilon_0.$$

Definindo $C = g_n(\mathcal{X}) \subset \mathcal{X}$ e $A = \bigcup_{p \in C} \tilde{B}_n(p, 2L(\epsilon + \epsilon_0))$ afirmamos que $\mu(A) \geqslant 1 - 1/L$.

De fato, se $x \in \mathcal{X} \setminus A$, então $\tilde{d}_n(x, g_n(x)) \geqslant 2L(\epsilon + \epsilon_0)$, caso contrário, $\tilde{d}_n(x, g_n(x)) < 2L(\epsilon + \epsilon_0)$ implicaria que $x \in \tilde{B}_n(p, 2L(\epsilon + \epsilon_0))$ para algum $p \in C$ e, consequentemente, $x \in A$.

Como

$$\int_{\mathcal{X}} \tilde{d}_n(x, g_n(x)) d\mu = \int_{\mathcal{X} \setminus A} \tilde{d}_n(x, g_n(x)) d\mu + \int_A \tilde{d}_n(x, g_n(x)) d\mu \leqslant 2(\epsilon + \epsilon_0),$$

obtemos

$$\int_{\mathcal{X} \setminus A} 2L(\epsilon + \epsilon_0) d\mu(x) \leqslant \int_{\mathcal{X} \setminus A} \tilde{d}_n(x, g_n(x)) d\mu(x) \leqslant 2(\epsilon + \epsilon_0),$$

implicando

$$2L(\epsilon + \epsilon_0)\mu(\mathcal{X} \setminus A) \leqslant 2(\epsilon + \epsilon_0).$$

Portanto, concluímos que

$$\mu(\mathcal{X} \setminus A) \leqslant \frac{1}{L}$$

consequentemente

$$\mu(A) \geqslant 1 - \frac{1}{L}.$$

Por construção, sabemos que o conjunto A pode ser coberto por $|C| = |g_n(\mathcal{X})|$ $(n, 2L(\epsilon + \epsilon_0))$-bolas dinâmicas médias. Logo, obtemos que

$$\tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L) \leqslant |g_n(\mathcal{X})|.$$

Aplicando o logaritmo e dividindo por n em ambos os lados da desigualdade, temos

$$\frac{1}{n}\log \tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L) \leqslant \frac{1}{n}\log|g_n(\mathcal{X})| \leqslant R_\mu(\epsilon)+\epsilon_0, \forall n > n(\epsilon_0).$$

Agora, para cada $\epsilon_0 > 0$ fixado, a desigualdade anterior implica que

$$\limsup_{n\to\infty}\frac{1}{n}\log \tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L) \leqslant R_\mu(\epsilon)+\epsilon_0.$$

Fazendo $\epsilon_0 \to 0$, em algum momento vamos ter $\epsilon_0 < \epsilon$ e $2L(\epsilon+\epsilon_0) \leqslant 4L\epsilon$, então

$$\tilde{N}_\mu(n, 4L\epsilon, 1/L) \leqslant \tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L)$$

e

$$\limsup_{n\to\infty}\frac{1}{n}\log \tilde{N}_\mu(n, 4L\epsilon, 1/L) \leqslant \limsup_{n\to\infty}\frac{1}{n}\log \tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L).$$

Portanto,

$$\tilde{h}_\mu(4L\epsilon, T, 1/L) = \limsup_{n\to\infty}\frac{1}{n}\log \tilde{N}_\mu(n, 4L\epsilon, 1/L) \leqslant R_\mu(\epsilon).$$

□

O próximo teorema resume as Proposições 5.2.4 e 5.2.8.

Teorema 5.2.9. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto com diâmetro D, $T : \mathcal{X} \to \mathcal{X}$ uma transformação contínua e μ uma medida de probabilidade ergódica T-invariante. Então*

$$\tilde{h}_\mu(4L\epsilon, T, 1/L) \leqslant R_\mu(\epsilon) \leqslant \tilde{h}_\mu(\epsilon, T, \epsilon/2D).$$

Definição 5.2.10. *Dado $x \in \mathcal{X}$, $n > 0$ e $r \in (0,1)$ definimos*

$$B'_n(x,\epsilon,r) = \left\{ y \in \mathcal{X} : \frac{1}{n} \# \left\{ 0 \leqslant i \leqslant n-1 : d(T^i(x), T^i(y)) \leqslant \epsilon \right\} \geqslant 1 - r \right\}.$$

Definimos também $N_\mu(n,\epsilon,\delta,r)$ como o número mínimo de bolas $B'_n(x,\epsilon,r)$ necessárias para cobrir um conjunto com medida maior ou igual que $1-\delta$.

A seguinte proposição segue de um resultado encontrado em [16].

Proposição 5.2.11. *Seja μ uma medida de probabilidade ergódica T-invariante. Então, para cada $\delta \in (0,1)$, temos a seguinte desigualdade*

$$h_\mu(T) \leqslant \lim_{r\to 0} \lim_{\epsilon \to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n,\epsilon,\delta,r).$$

Se assumirmos que $\mathcal{X}$ é compacto, então obtemos a igualdade.

Uma aplicação da Proposição 5.2.11 é o próximo resultado.

Proposição 5.2.12. *Seja μ uma medida de probabilidade ergódica T-invariante. Então, para $\delta \in (0,1)$ temos*

$$h_\mu(T) \leqslant \tilde{h}_\mu(T,\delta).$$

Demonstração. Inicialmente, observamos que se $y \in \tilde{B}_n(x,\epsilon)$, então $y \in B'_n(x, L\epsilon, 1/L)$. De fato, supondo por contradição que $y \in \tilde{B}_n(x,\epsilon)$ e

$$\# \left\{ 0 \leqslant i \leqslant n-1 : d(T^i(x), T^i(y) > L\epsilon \right\} > \frac{n}{L},$$

obtemos

$$\begin{aligned} \tilde{d}_n(x,y) &= \frac{1}{n} \sum_{i=0}^{n-1} d(T^i(x), T^i(y) \\ &\geqslant \frac{1}{n} \left(\frac{n}{L} \cdot L\epsilon \right) = \epsilon, \end{aligned}$$

implicando que $y \neq \tilde{B}_n(x,\epsilon)$.

Logo, concluímos que

$$\#\left\{0 \leqslant i \leqslant n-1 : d(T^i(x), T^i(y) > L\epsilon\right\} \leqslant \frac{n}{L}$$

que é equivalente à

$$\frac{1}{n}\#\left\{0 \leqslant i \leqslant n-1 : d(T^i(x), T^i(y)) \leqslant L\epsilon\right\} \geqslant 1 - \frac{1}{L}.$$

Com isso, concluímos que

$$\tilde{B}_n(x,\epsilon) \subset B'_n(x, L\epsilon, 1/L).$$

Então,

$$N_\mu(n, L\epsilon, \delta, 1/L) \leqslant \tilde{N}_\mu(n,\epsilon,\delta), \forall \delta \in (0,1)$$

e

$$\limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n, L\epsilon, \delta, 1/L) \leqslant \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_\mu(n,\epsilon,\delta).$$

Fazendo $\epsilon \to 0$, mas com L fixado, obtemos

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n, L\epsilon, \delta, 1/L) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_\mu(n,\epsilon,\delta).$$

Como

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_\mu(n,\epsilon,\delta) = \tilde{h}_\mu(T,\delta),$$

temos que

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n,\epsilon,\delta, 1/L) \leqslant \tilde{h}_\mu(T,\delta).$$

Fazendo $L \to \infty$ e usando a proposição anterior, obtemos

$$h_\mu(T) \leqslant \lim_{L\to\infty} \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_\mu(n,\epsilon,\delta,1/L) \leqslant \tilde{h}_\mu(T,\delta).$$

□

Corolário 5.2.13. *Seja $(\mathcal{X},d)$ um espaço métrico compacto de diâmetro D, $T:\mathcal{X}\to\mathcal{X}$ uma transformação contínua e μ uma medida de probabilidade ergódica T-invariante. Então*

$$\tilde{h}_\mu(T,\delta) = \lim_{\epsilon\to 0} R_\mu(\epsilon) = h_\mu(T).$$

Demonstração. Pelo Teorema 5.2.9, sabemos que

$$\tilde{h}_\mu(4L\epsilon,T,1/L) \leqslant R_\mu(\epsilon) \leqslant \tilde{h}_\mu(\epsilon,T,\epsilon/2D).$$

Como $\tilde{N}_\mu(n,\epsilon,\delta) \leqslant N_\mu(n,\epsilon,\delta)$, pois $B_n(x,\epsilon) \subset \tilde{B}_n(x,\epsilon)$, obtemos

$$\tilde{h}_\mu(\epsilon/2,T,\epsilon/2D) \leqslant h_\mu(\epsilon/2,T,\epsilon/2D).$$

Logo,

$$\tilde{h}_\mu(4L\epsilon,T,1/L) \leqslant R_\mu(\epsilon) \leqslant h_\mu(\epsilon,T,\epsilon/2D).$$

Tomando o limite com ϵ tendendo a zero, temos

$$\lim_{\epsilon\to 0} \tilde{h}_\mu(4L\epsilon,T,1/L) \leqslant \lim_{\epsilon\to 0} R_\mu(\epsilon) \leqslant \lim_{\epsilon\to 0} h_\mu(\epsilon,T,\epsilon/2D),$$

e, consequentemente

$$\tilde{h}_\mu(T,\delta) \leqslant \lim_{\epsilon\to 0} R_\mu(\epsilon) \leqslant h_\mu(T).$$

Pela proposição anterior, sabemos que $\tilde{h}_\mu(T,\delta) \geqslant h_\mu(T)$ e, portanto, concluímos que

$$\tilde{h}_\mu(T,\delta) = \lim_{\epsilon\to 0} R_\mu(\epsilon) = h_\mu(T).$$

□

Definição 5.2.14. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ uma aplicação contínua. Então*

$$\tilde{S}(\mathcal{X}, d, \epsilon) = \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_d(n, \epsilon),$$

onde $\tilde{N}_d(n, \epsilon)$ é definido como sendo a cardinalidade máxima de um conjunto (n, ϵ)-separado com respeito a métrica $\tilde{d}_n$.

Frequentemente denotaremos $\tilde{S}(\epsilon) = \tilde{S}(\mathcal{X}, d, \epsilon)$ quando a dinâmica está bem especificada.

Lema 5.2.15. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ uma aplicação injetiva. Dado $\epsilon > 0$ e $\delta \in (0, 1/4)$, existe uma medida de probabilidade T-invariante μ_ϵ tal que*

$$\tilde{h}_\mu(\epsilon, T, \delta) \geqslant \tilde{S}\left(\frac{2\epsilon}{1-4\delta}\right) - 3\delta\tilde{S}\left(\frac{\epsilon}{1-4\delta}\right).$$

Demonstração. Fixado $\epsilon > 0$ e $\delta \in (0, 1/4)$. Seja

$$E_n = \left\{x_1, x_2, \ldots, x_{\tilde{N}_d(n,\epsilon)}\right\}$$

uma coleção de pontos (n, ϵ)-separados em relação a métrica $\tilde{d}_n$ com cardinalidade $\tilde{N}_d(n, \epsilon)$. Definimos também a medida

$$\sigma_n = \frac{1}{|E_n|} \sum_{x\in E_n} \delta_x,$$

onde δ_x é a medida de probabilidade suportada no ponto x e

$$\overline{\sigma}_n = \frac{1}{n} \sum_{k=0}^{n-1} T_*^k \sigma_n.$$

Consideramos uma sequência $(n_k)_{k\in\mathbb{N}}$ tal que

$$\tilde{S}(\mathcal{X}, d, \epsilon) = \lim_{k\to\infty} \frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon).$$

Conforme argumentado no Lema 4.0.4, podemos obter uma subsequência de (n_k) tal que $(\overline{\sigma}_{n_k})_{k\in\mathbb{N}}$ converge a uma medida de probabilidade T-invariante

μ. Além disso, podemos considerar a sequência de tal modo que $\lim_{k\to\infty} \frac{n_k}{k} = \infty$.

Consideramos $K \subset \mathcal{X}$ um conjunto aberto com $\mu(K) > 1-\delta$ e

$$\tilde{N}\left(K, n, (1-4\delta)\epsilon/2\right) = \tilde{N}_\mu\left(n, (1-4\delta)\epsilon/2, \delta\right),$$

onde $\tilde{N}(K,n,r)$ é definido como sendo o número mínimo de (n,r)-bolas dinâmicas médias necessárias para cobrir K. Além disso, conforme argumentado no Lema 4.0.4, existe $k_0 \in \mathbb{N}$ tal que $\overline{\sigma}_{n_k}(K) > 1-\delta, \forall k \geqslant k_0$. Então, definimos o conjunto

$$L_n = \left\{(i,j) \in \mathbb{N}^2 : 0 \leqslant i \leqslant n-1, 1 \leqslant j \leqslant \tilde{N}_d(n,\epsilon)\right\},$$

onde $(i,j) = 1$ se $T^i(x_j) \in K$ e $(i,j) = 0$ caso contrário.

Definindo $\tilde{\tau}$ como sendo o número de uns em L_{n_k}, seguimos uma argumentação análoga ao Lema 4.0.4 e obtemos $\tilde{\tau} \geqslant n_k \tilde{N}_d(n_k,\epsilon)(1-\delta)$. Além disso, para $s > 1$ onde $s \in \mathbb{R}$, definimos $L_{n_k}(s)$ como sendo o conjunto dos elementos de L_{n_k} com as primeiras coordenadas no intervalo $\left[\left[\frac{n_k}{k}\right], n_k - \left[\frac{n_k}{s}\right] - 1\right]$ e concluímos também que

$$\tilde{\tau}(s) \geqslant n_k \tilde{N}(n_k,\epsilon)\left(1 - \delta - \frac{1}{s} - \frac{1}{k}\right),$$

onde $\tilde{\tau}(s)$ é o número de uns em $L_{n_k}(s)$.

Agora, ao assumirmos que

$$\frac{1}{1 - 2\delta - \dfrac{1}{k}} < s < \frac{1}{1-3\delta},$$

o que pode ser feito tomando k suficientemente grande de modo que $\delta > 1/k$, obtemos que o número de uns em $L_{n_k}(s)$ é pelo menos $n_k \tilde{N}_d(n_k,\epsilon)\delta$. Como $L_{n_k}(s)$ tem $\left(n_k - \left[\frac{n_k}{s}\right] - \left[\frac{n_k}{k}\right]\right)$ linhas, calculando o número médio de uns por linha concluímos que existe um índice m_k tal que a m_k-ésima linha possui pelo menos $\frac{n_k \tilde{N}_d(n_k,\epsilon)\delta}{n_k - \left[\frac{n_k}{s}\right] - \left[\frac{n_k}{k}\right]}$ uns e $\left[\frac{n_k}{k}\right] \leqslant m_k < n_k - \left[\frac{n_k}{s}\right]$.

Em paralelo a isto, consideramos B um conjunto $(m_k, \epsilon/2)$-separado com respeito a métrica $\tilde{d}_{m_k}$ com $|B| = \tilde{N}_d(m_k, \epsilon/2)$. Isto implica que podemos cobrir $\mathcal{X}$ com $\tilde{N}_d(m_k, \epsilon/2)$ $(m_k, \epsilon/2)$-bolas dinâmicas médias, ou seja,

$$\mathcal{X} = \bigcup_{y \in B} \tilde{B}_{m_k}(y, \epsilon/2).$$

Agora, observamos que se $i \neq j$ e

$$\tilde{d}_{m_k}(x_i, x_j) = \frac{1}{m_k} \sum_{l=0}^{m_k - 1} d(T^l(x_i), T^l(x_j)) \leqslant \epsilon,$$

então

$$\sum_{l=0}^{m_k - 1} d(T^l(x_i), T^l(x_j)) \leqslant m_k \epsilon.$$

Com isso, obtemos que

$$m_k \epsilon + \sum_{p=0}^{n_k - m_k - 1} d(T^{m_k + p}(x_i), T^{m_k + p}(x_j))$$

é maior ou igual à

$$\sum_{l=0}^{m_k - 1} d(T^l(x_i), T^l(x_j)) + \sum_{l=m_k}^{n_k - 1} d(T^l(x_i), T^l(x_j)) = n_k \cdot \tilde{d}_{n_k}(x_i, x_j),$$

que pela definição de E_{n_k}, implica

$$\sum_{p=0}^{n_k - m_k - 1} d(T^{m_k + p}(x_i), T^{m_k + p}(x_j)) \geqslant n_k \epsilon - m_k \epsilon = (n_k - m_k)\epsilon,$$

ou seja,

$$\tilde{d}_{n_k - m_k}(T^{m_k}(x_i), T^{m_k}(x_j)) \geqslant \epsilon.$$

Podemos concluir também que existe um um subconjunto

$$I \subset \left\{1, 2, \ldots, \tilde{N}_d(n_k, \epsilon)\right\}$$

tal que se $i \in I$, então $T^{m_k}(x_i) \in K$. Além disso, pelos argumentos anteriores, sabemos que

$$|I| \geqslant \frac{n_k \tilde{N}_d(n_k, \epsilon)\delta}{n_k - \left[\frac{n_k}{s}\right] - \left[\frac{n_k}{k}\right]} \geqslant \tilde{N}_d(n_k, \epsilon)\delta.$$

Então, pelo Princípio da Casa dos Pombos, existe um subconjunto $A \subset \{T^{m_k}(x_i)\}_{i \in I} \subset K$ tal que o diâmetro de A com respeito a $\tilde{d}_{m_k}$ é no máximo ϵ e

$$|A| \geqslant \frac{\tilde{N}_d(n_k, \epsilon)\delta}{\tilde{N}_d(m_k, \epsilon/2)}.$$

Logo, isto implica que se $a, b \in A$ e $a \neq b$, então $\tilde{d}_{n_k - m_k}(a, b) \geqslant \epsilon$ e, consequentemente, obtemos

$$\tilde{d}_{n_k}(a, b) \geqslant \tilde{d}_{n_k - m_k}(a, b) \geqslant \epsilon > \frac{\epsilon(n_k - m_k)}{n_k}.$$

Como $m_k < n_k - \left[\frac{n_k}{s}\right]$, então $-m_k > \left[\frac{n_k}{s}\right] - n_k$, implicando que

$$\frac{n_k - m_k}{n_k} > \frac{\left[\frac{n_k}{s}\right]}{n_k} > \frac{\dfrac{n_k}{s} - 1}{n_k} = \frac{1}{s} - \frac{1}{n_k} > 1 - 3\delta - \frac{1}{n_k},$$

devido a escolha de s. Com isso, concluímos que $\tilde{d}_{n_k}(a, b) > (1 - 3\delta - 1/n_k)\epsilon$. Como para k suficientemente grande $1/n_k < \delta$, isto implica que $\tilde{d}_{n_k}(a, b) > (1 - 4\delta)\,\epsilon$, mostramos que A é um conjunto $(n, (1-4\delta)\epsilon)$-separado em relação a $\tilde{d}_{n_k}$. Considerando P o número máximo de elementos de um conjunto $(n, (1 - 4\delta)\epsilon)$-separado em relação a $\tilde{d}_{n_k}$ contido em K, temos

$$\begin{aligned}
\tilde{N}_\mu\left(n, (1 - 4\delta)\epsilon/2, \delta\right) &= \tilde{N}_d\left(K, n, (1 - 4\delta)\epsilon/2\right) \\
&\geqslant P \\
&\geqslant |A| \\
&\geqslant \frac{\tilde{N}_d(n_k, \epsilon)\delta}{\tilde{N}_d(m_k, \epsilon/2)}
\end{aligned}$$

Definindo $\alpha = (1-4\delta)\epsilon/2$, obtemos

$$\begin{aligned}
\tilde{h}_\mu(\alpha, T, \delta) &= \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_\mu\left(n, \alpha, \delta\right) \\
&\geqslant \limsup_{k\to\infty} \frac{1}{n_k} \log \left(\frac{\tilde{N}_d(n_k, \epsilon)\delta}{\tilde{N}_d(m_k, \epsilon/2)} \right) \\
&= \limsup_{k\to\infty} \left(\frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon) - \frac{1}{n_k} \log \tilde{N}_d(m_k, \epsilon/2) \right) \\
&= \limsup_{k\to\infty} \left(\frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon) - \frac{m_k}{n_k} \cdot \frac{1}{m_k} \log \tilde{N}_d(m_k, \epsilon/2) \right).
\end{aligned}$$

Como $m_k < n_k - \left[\frac{n_k}{s}\right]$, então $\frac{m_k}{n_k} < 1 - \frac{1}{s} + \frac{1}{n_k}$. Logo,

$$\begin{aligned}
\tilde{h}_\mu(\alpha, T, \delta) &\geqslant \limsup_{k\to\infty} \left(\frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon) - \left(1 - \frac{1}{s} + \frac{1}{n_k}\right) \frac{1}{m_k} \log \tilde{N}_d(m_k, \epsilon/2) \right) \\
&\geqslant \limsup_{k\to\infty} \left(\frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon) - \left(3\delta + \frac{1}{n_k}\right) \frac{1}{m_k} \log \tilde{N}_d(m_k, \epsilon/2) \right) \\
&\geqslant \lim_{k\to\infty} \frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon) - \limsup_{k\to\infty} \left(\left(3\delta + \frac{1}{n_k}\right) \frac{1}{m_k} \log \tilde{N}_d(m_k, \epsilon/2) \right) \\
&= \lim_{k\to\infty} \frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon) - 3\delta \limsup_{k\to\infty} \frac{1}{m_k} \log \tilde{N}_d(m_k, \epsilon/2).
\end{aligned}$$

Como $\left[\frac{n_k}{k}\right] \leqslant m_k$, então $m_k \to \infty$ quando $k \to \infty$. Com isso, concluímos que

$$\tilde{h}_\mu(\alpha, T, \delta) \geqslant \tilde{S}(\epsilon) - 3\delta \tilde{S}(\epsilon/2),$$

pois, pela escolha da sequência (n_k), sabemos que

$$\tilde{S}(\epsilon) = \lim_{k\to\infty} \frac{1}{n_k} \log \tilde{N}_d(n_k, \epsilon)$$

e

$$
\begin{aligned}
\tilde{S}(\epsilon/2) &= \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_d(n, \epsilon/2) \\
&\geqslant \limsup_{k\to\infty} \frac{1}{m_k} \log \tilde{N}_d(m_k, \epsilon/2).
\end{aligned}
$$

Portanto,

$$\tilde{h}_\mu(\epsilon, T, \delta) \geqslant \tilde{S}\left(\frac{2\epsilon}{1-4\delta}\right) - 3\delta\tilde{S}\left(\frac{\epsilon}{1-4\delta}\right).$$

Em particular, isto implica que

$$\sup_{\mu\in\mathcal{M}_T(\mathcal{X})} \tilde{h}_\mu(\epsilon, T, \delta) \geqslant \tilde{S}\left(\frac{2\epsilon}{1-4\delta}\right) - 3\delta\tilde{S}\left(\frac{\epsilon}{1-4\delta}\right).$$

□

A partir de agora, precisamos de uma versão do Teorema 5.2.5 para medidas não ergódicas. Para isso, consideramos em $\mathcal{X}^n$ a métrica ρ_n dada por

$$\rho_n(x, y) = \frac{1}{n}\sum_{i=1}^{n} d(x_i, y_i),$$

onde $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$. Definimos também a medida μ^n como $\mu^n(\mathcal{A}) = \mu((T^{(n)})^{-1}(\mathcal{A}))$, para todo $\mathcal{A} \subset \mathcal{X}^n$ mensurável, onde $T^{(n)} : \mathcal{X} \to \mathcal{X}^n$ é definida como $T^{(n)}(x) = (x, T(x), \ldots, T^{n-1}(x))$. Além disso, para um subconjunto finito $C_n \subset \mathcal{X}$, definimos $\mathbb{E}_\mu(C_n)$ como

$$\mathbb{E}_\mu(C_n) = \int_{\mathcal{X}^n} \min_{c\in C_n} \rho_n(x, c)\, d\mu^n(x).$$

Definição 5.2.16. *Dado $R > 0$, definimos*

$$\delta_\mu(R) = \inf\left\{\mathbb{E}_\mu(C_n) : C_n \subset \mathcal{X}^n \text{ e } |C_n| \leqslant e^{nR}\right\}$$

e

$$D_\mu(R) = \inf\left\{\mathbb{E}_\mu\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^k X, Y_k \circ X)\right) : \frac{1}{n} I_\mu(X, (Y_0, \ldots, Y_{n-1})) \leqslant R\right\}.$$

As demonstrações dos próximos três resultados podem ser encontradas em [14].

Teorema 5.2.17. *Seja μ uma medida de probabilidade ergódica T-invariante em $\mathcal{X}$. Então*

$$\delta_\mu(R) = D_\mu(R).$$

Proposição 5.2.18. *A função $\mu \mapsto \delta_\mu(R)$ é afim e semicontínua superiormente. A função $R \mapsto \delta_\mu(R)$ é convexa e decrescente.*

Corolário 5.2.19. *Seja μ uma medida de probabilidade T-invariante em $\mathcal{X}$. Então*

$$\delta_\mu(R) = \int_{\mathcal{X}} \delta_{\mu_x}(R) d\mu(x),$$

onde $\mu = \int \mu_x d\mu(x)$ é a decomposição ergódica de μ.

Além disso, temos que

$$\inf_{R \geqslant 0} \delta_\mu(R) = 0.$$

De fato, como $\mathbb{E}_\mu(C_n) = \int_{\mathcal{X}^n} \min_{c \in C_n} \rho_n(x,c) d\mu^n(x)$, ρ_n é uma métrica e a integral é um funcional linear positivo, obtemos que $\mathbb{E}_\mu(C_n) \geqslant 0$ e, consequentemente, $\inf_{R \geqslant 0} \delta_\mu(R) \geqslant 0$.

Além disso, dado $\epsilon > 0$, seja B um conjunto ϵ-separado de cardinalidade máxima. Então, isto implica que para todo $x \in \mathcal{X}$, existe um elemento $b \in B$ tal que $d(x,b) < \epsilon$. Logo, definimos C_n como

$$C_n = B \times B \times \cdots \times B,$$

onde estamos tomando o produto cartesiano de B n vezes. Com isso, concluímos que $|C_n| = N(\epsilon)^n$ e $\rho_n(x,c) < \epsilon$, para algum $c \in C_n$.

De fato, para todo $x = (x_1, \ldots, x_n) \in \mathcal{X}^n, \exists c = (c_1, \ldots, c_n) \in C_n$ tal que $d(x_i, c_i) < \epsilon, \forall i \in \{1, \ldots, n\}$. Logo

$$\rho_n(x,c) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, c_i) < \frac{1}{n} \cdot n\epsilon < \epsilon.$$

Portanto,

$$\int_{\mathcal{X}^n} \min_{c \in C_n} \rho_n(x,c) d\mu^n(x) < \int_{\mathcal{X}^n} \epsilon \, d\mu^n(x) = \epsilon,$$

desde que $N(\epsilon)^n \leqslant e^{nR}$, o que é possível tomando R suficientemente grande. Logo, $\mathbb{E}_\mu(C_n)$ pode ser tomado tão próximo de zero quanto se queira, basta considerar C_n assim como definido anteriormente e R suficientemente grande. Com isso, concluímos que $\inf_{R \geqslant 0} \delta_\mu(R) = 0$.

Agora, como $\delta_\mu(R)$ é convexa e $\inf_{R\geqslant 0} \delta_\mu(R) = 0$, temos que $\delta_\mu(R)$ é estritamente decrescente em R se $\delta_\mu(R) > 0$. De fato, supondo por contradição que existem R_1 e R_2 tais que $R_1 < R_2$ e $\delta_\mu(R_1) = \delta_\mu(R_2) > 0$. Como $\inf_{R\geqslant 0} \delta_\mu(R) = 0$, exite R_3 tal que $R_1 < R_2 < R_3$ e $\delta_\mu(R_3) < \delta_\mu(R_1)$. Como $\delta_\mu(R)$ é convexa em R, isto implica que

$$\delta_\mu(R_1(1-\lambda) + \lambda R_3) \leqslant (1-\lambda)\delta_\mu(R_1) + \lambda\delta_\mu(R_3), \forall \lambda \in [0,1].$$

Considerando $\lambda' \in (0,1)$ tal que $(1-\lambda')R_1 + \lambda' R_3 = R_2$, obtemos

$$\begin{aligned}
\delta_\mu(R_2) &\leqslant (1-\lambda')\delta_\mu(R_1) + \lambda'\delta_\mu(R_3) \\
&< (1-\lambda')\delta_\mu(R_1) + \lambda'\delta_\mu(R_1) \\
&= \delta_\mu(R_1) \\
&= \delta_\mu(R_2).
\end{aligned}$$

Logo, obtemos que $\delta_\mu(R_2) < \delta_\mu(R_2)$, o que é uma contradição. Portanto, $\delta_\mu(R)$ é estritamente decrescente se $\delta_\mu(R) > 0$, pois, pela proposição 3.3.16 já sabemos que $R \mapsto \delta_\mu(R)$ é decrescente.

Definindo

$$D_\infty(R) = \sup_{x \in \mathcal{X}} D_{\mu_x}(R) = \sup_{x\in\mathcal{X}} \delta_{\mu_x}(R),$$

temos que $D_\infty(R)$ é uma função convexa pois é o supremo de funções convexas, consequentemente é também uma função contínua. Além disso, pela definição de $D_\infty(R)$, também temos que $\inf_{R\geqslant 0} D_\infty(R) = 0$ e $D_\infty(R)$ é estritamente decrescente em R se $D_\infty(R) > 0$.

Afirmamos também que $R_\mu(\epsilon)$ e $D_\mu(R)$ são inversas uma da outra nos pontos onde $R_\mu(\epsilon)$ e $D_\mu(R)$ são estritamente decrescentes. De fato, pela definição de $D_\mu(R)$, temos

$$\begin{aligned} D_\mu(R_\mu(\epsilon)) &= \inf_{X,Y}\left\{\mathbb{E}_\mu\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^kX, Y_k\circ X)\right) : \frac{1}{n}I_\mu(X,Y) \leqslant R_\mu(\epsilon)\right\} \\ &= \inf_{X,Y}\left\{\mathbb{E}_\mu\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^kX, Y_k\circ X)\right) : \frac{1}{n}I_\mu(X,Y) \leqslant \inf\frac{1}{n}I_\mu(Z,W)\right\} \end{aligned}$$

onde Z, W satisfazem a condição $(*)_{n,\epsilon,\mu}$.

Como $R_\mu(\epsilon)$ é estritamente decrescente, para que $\frac{1}{n}I_\mu(X,Y) \leqslant R_\mu(\epsilon)$ é necessário que X, Y satisfaçam $(*)_{n,\epsilon_1,\mu}$ onde $\epsilon_1 \geqslant \epsilon$. Caso contrário, se $\epsilon_1 < \epsilon$, teríamos $R_\mu(\epsilon_1) > R_\mu(\epsilon)$, ou seja, $\inf \frac{1}{n}I_\mu(X,Y) > \inf \frac{1}{n}I_\mu(Z,W)$, o que seria uma contradição. Logo, obtemos $D_\mu(R_\mu(\epsilon)) = \epsilon$.

Por outro lado,

$$R_\mu(D_\mu(R)) = \inf_{Z,W} \frac{1}{n}I_\mu(Z,W),$$

tais que, Z, W satisfazem a condição $(*)_{n,D_\mu(R),\mu}$. Como $R_\mu(\epsilon)$ é estritamente decrescente, o ínfimo vai ocorrer quando

$$\mathbb{E}_\mu\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^kZ, W_k)\right) = D_\mu(R)$$

que é igual a

$$\inf_{X,Y}\left\{\mathbb{E}_\mu\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^kX, Y_k\circ X)\right) : \frac{1}{n}I_\mu(X,Y) \leqslant R\right\}.$$

Como $D_\mu(R)$ é estritamente decrescente, temos que se $R_1 < R$, isto implica que $D_\mu(R_1) > D_\mu(R)$. Então

$$\inf_{X,Y}\left\{\mathbb{E}\left(\frac{1}{n}\sum_{k=0}^{n-1} d(T^kX, Y_k\circ X)\right) : \frac{1}{n}I_\mu(X,Y) \leqslant R_1 < R\right\}$$

é maior do que

$$\inf_{X,Y} \left\{ \mathbb{E}\left(\frac{1}{n} \sum_{k=0}^{n-1} d(T^k X, Y_k \circ X) \right) : \frac{1}{n} I_\mu(X,Y) \leqslant R \right\}.$$

Logo, o ínfimo da esperança acima ocorre quando $\frac{1}{n} I_\mu(X,Y) = R$. Portanto, $R_\mu(D_\mu(R)) = R$.

Lema 5.2.20. *Seja* $D_{\mu_x} : (0,\infty) \to \mathbb{R}$ *uma função convexa decrescente para cada* $x \in \mathcal{X}$ *e definindo*

$$D_\infty(R) = \sup_x D_{\mu_x}(R).$$

Seja $I = \{t \in (0,\infty) : D_\infty(t) > 0\}$ *e* $J = Im(D_\infty)$. *Então* $D_\infty^{-1} : J \to I$ *está bem definida e para todo* $\epsilon \in J$, *temos*

$$D_\infty^{-1}(\epsilon) = \sup_x D_{\mu_x}^{-1}(\epsilon).$$

Demonstração. Pela convexidade de D_{μ_x} obtemos a convexidade de D_∞ que implica a continuidade de D_∞. Como D_∞ é estritamente decrescente em R quando $D_\infty(R) > 0$, temos que D_∞^{-1} está bem definida e que J é um intervalo aberto.

Afirmamos que

$$D_\infty^{-1}(\epsilon) \geqslant \sup_x D_{\mu_x}^{-1}(\epsilon).$$

De fato, supondo por contradição que $D_\infty(R) \geqslant D_{\mu_x}(R)$ e $D_\infty^{-1}(\epsilon) < \sup_x D_{\mu_x}^{-1}(\epsilon)$. Então, temos que

$$\begin{aligned} D_\infty^{-1}(D_\infty(R)) &< D_{\mu_x}^{-1}(D_\infty(R)) \\ &\leqslant D_{\mu_x}^{-1}(D_{\mu_x}(R)) \end{aligned}$$

pois $D_{\mu_x}^{-1}(\epsilon)$ é decrescente. Logo, obtemos que $R < R$, o que é uma contradição. Portanto, concluímos que

$$D_\infty^{-1}(\epsilon) \geqslant \sup_x D_{\mu_x}^{-1}(\epsilon).$$

Agora, afirmamos que

$$\sup_x D_{\mu_x}^{-1}(\epsilon) \geqslant D_{\infty}^{-1}(\epsilon).$$

Supondo novamente por contradição que existe $\epsilon_0 > 0$ e $\epsilon \in J$ tal que para todo $x \in \mathcal{X}$ temos $D_{\infty}^{-1}(\epsilon) - \epsilon_0 > D_{\mu_x}^{-1}(\epsilon)$ sempre que $D_{\infty}^{-1}(\epsilon)$ está bem definida. Seja $b = D_{\infty}^{-1}(\epsilon)$ e $a = D_{\infty}^{-1}(\epsilon) - \epsilon_0$. Como

$$D_{\infty}(a) = D_{\infty}(D_{\infty}^{-1}(\epsilon) - \epsilon_0) > D_{\infty}(D_{\infty}^{-1}(\epsilon)) = \epsilon$$

e

$$D_{\infty}(a) = \sup_x D_{\mu_x}(a) > \epsilon,$$

existe $\delta > 0$ tal que $\sup_x D_{\mu_x}(a) - \delta > \epsilon$. Logo, isto implica que existe $x \in \mathcal{X}$ tal que

$$D_{\mu_x}(a) > \sup_x D_{\mu_x}(a) - \delta > \epsilon.$$

Portanto, obtemos

$$D_{\infty}(a) \geqslant D_{\mu_x}(a) > \epsilon = D_{\infty}(b) \geqslant D_{\mu_x}(b).$$

Pelo Teorema do Valor Intermediário, $\exists c \in (a, b]$ tal que $D_{\mu_x}(c) = \epsilon$, implicando que $c > a = D_{\infty}^{-1}(\epsilon) - \epsilon_0$, mas $D_{\infty}^{-1}(\epsilon) - \epsilon_0 > D_{\mu_x}^{-1}(\epsilon) = c$. Logo, obtemos uma contradição e, consequentemente, temos

$$D_{\infty}^{-1}(\epsilon) = \sup_x D_{\mu_x}^{-1}(\epsilon).$$

□

Proposição 5.2.21. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e μ uma medida de probabilidade T-invariante. Então*

$$\tilde{h}_{\mu}(L\epsilon, T, 1/L) \leqslant \sup_x R_{\mu_x}(\epsilon) \leqslant \sup_{\mu \text{ ergódicas}} R_{\mu}(\epsilon).$$

Demonstração. Afirmamos que dado $R > 0$ e $\epsilon_0 > 0$, existe $C_n \subset \mathcal{X}^n$ tal que $|C_n| \leqslant e^{n(R+\epsilon_0)}$ e

$$\mathbb{E}_\mu(C_n) \leqslant \int_{\mathcal{X}} D_{\mu_x}(R) d\mu(x) + \epsilon_0.$$

De fato, como $\delta_\mu(R)$ é decrescente, obtemos

$$\begin{aligned}
\delta_\mu(R+\epsilon_0) &\leqslant \delta_\mu(R) \\
&\leqslant \delta_\mu(R) + \epsilon_0 \\
&= \int_{\mathcal{X}} \delta_{\mu_x}(R) d\mu(x) + \epsilon_0 \\
&= \int_{\mathcal{X}} D_{\mu_x}(R) d\mu(x) + \epsilon_0 \\
&\leqslant \int_{\mathcal{X}} \sup_x D_{\mu_x}(R) d\mu(x) + \epsilon_0 \\
&= D_\infty(R) + \epsilon_0.
\end{aligned}$$

Logo,

$$\delta_\mu(R+\epsilon_0) = \inf\left\{\mathbb{E}_\mu(C_n) : C_n \subset \mathcal{X}^n \text{ e } |C_n| \leqslant e^{n(R+\epsilon_0)}\right\} \leqslant D_\infty(R) + \epsilon_0.$$

Em particular, para $R = R_\infty(\epsilon) = \sup_x R_{\mu_x}(\epsilon)$, obtemos

$$\mathbb{E}_\mu(C_n) \leqslant D_\infty(R_\infty(\epsilon)) + \epsilon_0 = \epsilon + \epsilon_0$$

para algum C_n. Como

$$\mathbb{E}_\mu(C_n) = \int_{\mathcal{X}^n} \min_{c \in C_n} \rho_n(x,c) d\mu^n(x) \leqslant \epsilon + \epsilon_0,$$

isto implica que $\forall x \in \mathcal{X}^n, \exists c \in C_n$ tal que $\rho_n(x,c) \leqslant \epsilon + \epsilon_0$. Então, concluímos que podemos cobrir $\mathcal{X}^n$ com bolas fechadas centradas nos pontos de C_n e raio $\epsilon + \epsilon_0$ com respeito a métrica ρ_n. Com isso, obtemos que

$$|C_n| \geqslant C_d(n, 2(\epsilon+\epsilon_0)) \geqslant N_d(n, 2(\epsilon+\epsilon_0)).$$

Seja $A \subset \mathcal{X}^n$ um conjunto $(n, 2(\epsilon+\epsilon_0))$-separado com cardinalidade igual a $N_d(n, 2(\epsilon+\epsilon_0))$ e $B \subset \mathcal{X} \times T(\mathcal{X}) \times \cdots \times T^{n-1}(\mathcal{X})$ um conjunto $(n, 2(\epsilon+\epsilon_0))$-separado com cardinalidade máxima tal que todo ponto $\overline{b} \in B$ é da forma $\overline{b} = (b, T(b), \ldots, T^{n-1}(b))$.

Logo,

$$\begin{aligned}|A| &\geqslant |B|\\ &\geqslant \min\left\{|C| : C \text{ é uma cobertura por bolas abertas de raio } 2(\epsilon+\epsilon_0)\right\},\end{aligned}$$

cujos centros das bolas são da forma $\overline{c} = (c, T(c), \ldots, T^{n-1}(c))$ e cobrem o conjunto K formado por todos os pontos $\overline{x} = (x, T(x), \ldots, T^{n-1}(x))$.

Agora, observamos que se $\overline{x} = (x, \ldots, T^{n-1}(x)) \in B_{\rho_n}(\overline{c}, 2(\epsilon+\epsilon_0))$, então, pela definição de ρ_n, obtemos que

$$\tilde{d}_n(x,c) = \frac{1}{n}\sum_{i=0}^{n-1} d(T^i(x), T^i(c)) < 2\epsilon + 2\epsilon_0.$$

Com isso, concluímos que $x \in \tilde{B}_n(c, 2\epsilon+2\epsilon_0)$. Como $\overline{x}$ é arbitrário e a união das bolas centradas nos pontos de C cobrem K, ao considerarmos o conjunto C' dos pontos de $\mathcal{X}$ formado pela primeira coordenada dos pontos de C, formamos uma cobertura de $\mathcal{X}$ por bolas de raio $2\epsilon + 2\epsilon_0$ com respeito a métrica $\tilde{d}_n$. Logo,

$$|C| = |C'| \geqslant \tilde{N}_\mu(n, 2(\epsilon+\epsilon_0), \delta), \forall \delta \in (0,1).$$

Em particular, para $\delta = 1/L$, com $L > 1$, obtemos

$$e^{n(R_\infty(\epsilon)+\epsilon_0)} \geqslant |C_n| \geqslant \tilde{N}_\mu(n, 2(\epsilon+\epsilon_0), 1/L) \geqslant \tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L).$$

Tomando ϵ_0 suficientemente pequeno de modo que $\epsilon_0 < \epsilon$, obtemos

$$e^{n(R_\infty(\epsilon)+\epsilon_0)} \geqslant \tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L) \geqslant \tilde{N}_\mu(n, 4L\epsilon, 1/L).$$

Consequentemente,

$$\begin{aligned}R_\infty(\epsilon)+\epsilon_0 &\geqslant \limsup_{n\to\infty}\frac{1}{n}\log \tilde{N}_\mu(n, 2L(\epsilon+\epsilon_0), 1/L)\\ &\geqslant \limsup_{n\to\infty}\frac{1}{n}\log \tilde{N}_\mu(n, 4L\epsilon, 1/L)\end{aligned}$$

Fazendo $\epsilon_0 \to 0$, temos

$$R_\infty(\epsilon) \geqslant \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_\mu(n, 4L\epsilon, 1/L) = \tilde{h}_\mu(4L\epsilon, T, 1/L),$$

logo

$$\tilde{h}_\mu(4L\epsilon, T, 1/L) \leqslant R_\infty(\epsilon) = \sup_x R_{\mu_x}(\epsilon) \leqslant \sup_{\mu \text{ ergódicas}} R_\mu(\epsilon).$$

□

Definição 5.2.22. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto. $(\mathcal{X}, d)$ satisfaz a condição 1.2 se para todo $\delta > 0$, temos*

$$\lim_{\epsilon\to 0} \epsilon^\delta \log \#(\mathcal{X}, d, \epsilon) = 0,$$

onde $\#(\mathcal{X}, d, \epsilon)$ é o número mínimo de ϵ-bolas necessárias para cobrir $\mathcal{X}$.

Proposição 5.2.23. *Seja $(\mathcal{X}, d)$ um espaço métrico compacto e $T : \mathcal{X} \to \mathcal{X}$ uma aplicação contínua. Então*

$$\limsup_{\epsilon\to 0} \frac{\sup_\mu R_\mu(\epsilon)}{|\log(\epsilon)|} = \limsup_{\epsilon\to 0} \frac{\tilde{S}(\mathcal{X}, d, \epsilon)}{|\log(\epsilon)|}.$$

Se $(\mathcal{X}, d)$ satisfaz a condição 1.2, então

$$\overline{mdim}(\mathcal{X}, d, T) = \limsup_{\epsilon\to 0} \frac{\sup_\mu R_\mu(\epsilon)}{|\log(\epsilon)|}.$$

Demonstração. Seja $\left\{x_1, \ldots, x_{\tilde{N}_d(n,\epsilon/2)}\right\}$ um conjunto $(n, \epsilon/2)$-separado com respeito a métrica $\tilde{d}_n$, então $\left\{\tilde{B}_n(x_1, \epsilon/2), \ldots, \tilde{B}_n(x_{\tilde{N}_d(n,\epsilon/2)}, \epsilon/2)\right\}$ é uma cobertura de $\mathcal{X}$ com diâmetro no máximo ϵ e, com isso, temos que

$$\tilde{C}_d(n, \epsilon) \leqslant \tilde{N}_d(n, \epsilon/2),$$

onde $\tilde{C}_d(n, \epsilon)$ é definido analogamente a $C_d(n, \epsilon)$, porém usando $\tilde{d}_n$.

Pelo lema 3.1 de [5], obtemos

$$R_\mu(\epsilon) \leqslant \lim_{n\to\infty} \frac{1}{n} \log \tilde{C}_d(n, \epsilon) \leqslant \limsup_{n\to\infty} \frac{1}{n} \log \tilde{N}_d(n, \epsilon/2) = \tilde{S}(\mathcal{X}, d, \epsilon/2)$$

e

$$\sup_{\mu} R_\mu(\epsilon) \leqslant \tilde{S}(\mathcal{X}, d, \epsilon/2).$$

Então,

$$\frac{\sup_\mu R_\mu(\epsilon)}{|\log(\epsilon)|} \leqslant \frac{\tilde{S}(\mathcal{X}, d, \epsilon/2)}{|\log(\epsilon)|}.$$

Portanto,

$$\limsup_{\epsilon\to 0} \frac{\sup_\mu R_\mu(\epsilon)}{|\log(\epsilon)|} \leqslant \limsup_{\epsilon\to 0} \frac{\tilde{S}(\mathcal{X}, d, \epsilon/2)}{|\log(\epsilon)|},$$

consequentemente

$$\limsup_{\epsilon\to 0} \frac{\sup_\mu R_\mu(\epsilon)}{|\log(\epsilon)|} \leqslant \limsup_{\epsilon\to 0} \frac{\tilde{S}(\mathcal{X}, d, \epsilon)}{|\log(\epsilon)|}.$$

Por outro lado, como

$$\tilde{h}_\mu(4L\epsilon, T, 1/L) \leqslant \sup_{\mu \text{ ergódicas}} R_\mu(\epsilon)$$

para qualquer medida de probabilidade T-invariante, em particular vale para

$$\sup_{\mu} \tilde{h}_\mu(4L\epsilon, T, 1/L).$$

Escolhendo $\delta = 1/L$, podemos escrever a desigualdade acima como

$$\sup_{\mu} \tilde{h}_\mu(4\epsilon/\delta, T, \delta) \leqslant \sup_{\mu \text{ ergódicas}} R_\mu(\epsilon).$$

Pelo Lema 5.2.15, obtemos

$$\sup_{\mu \text{ ergódicas}} R_\mu(\epsilon) \geqslant \sup_{\mu} \tilde{h}_\mu(4\epsilon/\delta, T, \delta) \geqslant \tilde{S}\left(\frac{8\epsilon}{\delta(1-4\delta)}\right) - 3\delta\tilde{S}\left(\frac{4\epsilon}{\delta(1-4\delta)}\right).$$

Logo

$$\limsup_{\epsilon\to 0}\frac{\sup_{\mu \text{ ergódicas}} R_\mu(\epsilon)}{|\log(\epsilon)|} \geqslant \limsup_{\epsilon\to 0}\left(\frac{\tilde{S}\left(8\epsilon/\delta(1-4\delta)\right)}{|\log(\epsilon)|} - 3\delta\frac{\tilde{S}\left(4\epsilon/\delta(1-4\delta)\right)}{|\log(\epsilon)|}\right)$$

Fazendo $\delta \to 0$, obtemos

$$\limsup_{\epsilon\to 0}\frac{\sup_{\mu \text{ ergódicas}} R_\mu(\epsilon)}{|\log(\epsilon)|} \geqslant \limsup_{\epsilon\to 0}\frac{\tilde{S}(\epsilon)}{|\log(\epsilon)|}.$$

Portanto, temos que

$$\limsup_{\epsilon\to 0}\frac{\sup_{\mu} R_\mu(\epsilon)}{|\log(\epsilon)|} = \limsup_{\epsilon\to 0}\frac{\tilde{S}(\mathcal{X},d,\epsilon)}{|\log(\epsilon)|}.$$

Se $(\mathcal{X}, d)$ satisfaz a condição 1.2, foi mostrado em [5] que

$$\overline{mdim}(\mathcal{X},d,T) = \limsup_{\epsilon\to 0}\frac{\tilde{S}(\mathcal{X},d,\epsilon)}{|\log(\epsilon)|}$$

Logo, concluímos que

$$\overline{mdim}(\mathcal{X},d,T) = \limsup_{\epsilon\to 0}\frac{\sup_{\mu} R_\mu(\epsilon)}{|\log(\epsilon)|},$$

quando $(\mathcal{X}, d)$ satisfaz a condição 1.2. $\square$

Referências Bibliográficas

[1] A. Velozo, R. Velozo, Rate distortion theory, metric mean dimension and measure theoretic entropy. arXiv:1707.05762 (2017).

[2] A. Katok, B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems. Cambrigde University Press (1996).

[3] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. Inst. Hautes Études Sci. Publ. Math. 51 (1980), 137-173.

[4] E. Lindenstrauss, Mean dimension, small entropy factors and an embedding theorem. Inst. Hautes Études Sci. Publ. Math. 89 (1999), 227-262.

[5] E. Lindenstrauss, M. Tsukamoto, From rate distortion theory to metric mean dimension: variational principle, arXiv:1702.05722 (2017).

[6] E. Lindenstrauss, B. Weiss, Mean topological dimension. Israel J. Math. 115 (2000), 1-24.

[7] E. L. Lima, Espaços métricos. Quinta Edição, IMPA, Rio de Janeiro (2015).

[8] H. L. Royden, P. M. Fitzpatrick, Real Analysis. Fourth Edition, Pearson (2010).

[9] K. Oliveira, M. Viana, Fundamentos da Teoria Ergódica. Primeira Edição. SBM (2014)

[10] M. Gromov, Topological invariants of dynamical systems and spaces of holomorphic maps. I. Math. Phys. Anal. Geom. 2 (1999), no. 4, 323-415.

[11] M. Brin, G. Stuck, Introduction to Dynamical Systems. Cambrigde University Press (2002).

[12] P. Walters, introduction to ergodic theory. Graduate Texts in Mathematics, 79. Springer Verlag, New York-Berlin, (1982).

[13] R. Mañé, Ergodic Theory and Differentiable Dynamics, Springer Verlag (2011)

[14] R. Gray, Entropy and information theory. Second edition. Springer, New York, (2011).

[15] R. Bartle, The elements of integration and Lebesgue measure. Wiley Classic Library (2014).

[16] X. Zhou, L. Zhou, E. Chen, Brin-Katok formula for the measure theoretic r-entropy. C. R. Math. Acad. Sci. Paris 352 (2014), no. 6, 473-477.