

5

Entropia

Neste capítulo, introduziremos a entropia e desenvolveremos suas propriedades básicas. Em seguida, descreveremos a teoria de Perron-Frobenius de matrizes não-negativas. Finalmente, usando essa teoria, calcularemos a entropia de shifts de tipo finito irredutíveis. Veremos que a entropia é dada pelo logaritmo do maior autovalor da matriz adjacente do shift.

5.1

Definição e Propriedades Básicas

Inicialmente, iremos rever algumas notações e conceitos de Álgebra Linear. O *polinômio característico* de uma matriz quadrada A é definido por

$$\mathcal{X}_A(t) = \det(t \text{Id} - A),$$

onde Id é a matriz identidade. Os *autovalores* de A são as raízes do polinômio $\mathcal{X}_A(t)$. Um *autovetor* de A correspondente ao autovalor λ é um vetor não nulo, tal que $Av = \lambda v$.

Uma matriz A é *positiva* se cada uma de suas entradas é positiva (estritamente maior do que 0), nesse caso escrevemos $A > 0$. De maneira semelhante, dizemos que A é *não negativa* se cada uma de suas entradas é não negativa. Neste caso escreveremos $A \geq 0$.

Definição 5.1 (Entropia) *Seja X um espaço de seqüências. A entropia de X é definida por*

$$h(X) = \lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_n(X)|,$$

onde $|\mathfrak{B}_n(X)|$ é o número de n -blocos que aparecem em pontos de X (ver a Definição 2.5).

A seguinte proposição implica que a entropia de um espaço de seqüências está bem definida e é finita.

Proposição 5.2 *Seja X um espaço de seqüências, então*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_n(X)| = \inf_{n \geq 1} \frac{1}{n} \log |\mathfrak{B}_n(X)|.$$

Adiaremos a prova desta proposição e antes calcularemos, na próxima seção, a entropia em alguns exemplos.

5.1.1

Exemplos

Exemplo 19 *Seja $X = X_{[r]}$ a r -seqüência completa. A entropia de X é $\log(r)$.*

Prova: A r -seqüência completa é a seqüência completa sobre o alfabeto $\mathcal{A} = \{0, 1, \dots, r-1\}$. Então, $|\mathfrak{B}_n(X)|$ é o número de n -blocos de tamanho r que aparecem em pontos de X , isto é,

$$|\mathfrak{B}_n(X)| = \underbrace{r \cdot r \cdots r}_{n \text{ vezes}} = r^n.$$

Logo, a entropia de X é dada por :

$$\begin{aligned} h(X) &= \lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_n(X)| = \lim_{n \rightarrow +\infty} \frac{1}{n} \log (r)^n = \lim_{n \rightarrow +\infty} \frac{1}{n} n \log(r) = \\ &= \lim_{n \rightarrow +\infty} \log(r) = \log(r). \end{aligned}$$

Provamos assim a afirmação. $\square$

Exemplo 20 *Considere o alfabeto $\mathcal{A} = \{0, 1\}$, o grafo G cuja matriz adjacente é*

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

e o shift vértice $X = \hat{X}_G$.

Considere o bloco $\mathfrak{P} = \{11\}$ e observe que $X = X_{\mathfrak{P}}$ (note que $\mathfrak{P}$ representa o bloco proibido). Observe também que X é o shift de ouro.

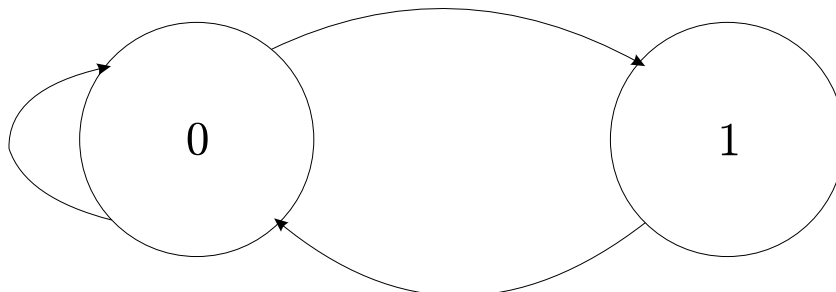


Figura 5.1: Grafo do shift de ouro

A entropia de X é

$$h(X) = \log \left(\frac{1 + \sqrt{5}}{2} \right),$$

onde $\frac{1+\sqrt{5}}{2}$ é o maior autovalor de A .

Observação 5.3 Para $n \geq 2$, existe uma correspondência injetiva entre $\mathfrak{B}_n(X) = \mathfrak{B}_n(\hat{X}_G)$ e $\mathfrak{B}_{n-1}(X_G)$.

De fato, como $X = \hat{X}_G$, temos que o shift de vértices $\hat{X}_G$ é dado sobre o alfabeto $\mathcal{A} = \{0, 1\}$. Logo, sua linguagem $\mathcal{L} = \bigcup_{n=2}^{\infty} \mathfrak{B}_n(\hat{X}_G)$ é dada por

$$\{00, 01, 10, \dots\}.$$

Por outro lado, o shift de arestas X_G é dado sobre o alfabeto $\mathcal{A} = \{00, 01, 10\}$ e sua linguagem $\mathcal{L} = \bigcup_{n=2}^{\infty} \mathfrak{B}_{n-1}(X_G)$ é dada por

$$\{00, 01, 10, \dots\}.$$

Assim, concluimos a observação e iremos a seguir provar o exemplo 20.

Prova: Lembre que, pela Observação 5.3, para todo $n \geq 2$ existe uma correspondência injetiva entre $\mathfrak{B}_n(X) = \mathfrak{B}_n(\hat{X}_G)$ e $\mathfrak{B}_{n-1}(X_G)$. Assim, aplicando a Proposição 3.14, temos que

$$|\mathfrak{B}_m(X_G)| = \sum_{i=1}^2 (A^m)_{IJ}.$$

Calcularemos as entradas $(A^m)_{IJ}$ de A^m diagonalizando a matriz A . Os autovalores de A são as raízes do polinômio característico

$$\chi_A(t) = \det \begin{bmatrix} t-1 & -1 \\ -1 & t \end{bmatrix} = t^2 - t - 1,$$

dados por:

$$\lambda = \frac{1 + \sqrt{5}}{2} \quad \text{e} \quad \mu = \frac{1 - \sqrt{5}}{2}.$$

Dois autovetores correspondentes são por:

$$\begin{bmatrix} \lambda \\ 1 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} \mu \\ 1 \end{bmatrix}$$

Consideramos agora as matrizes

$$P = \begin{bmatrix} \lambda & \mu \\ 1 & 1 \end{bmatrix}, \quad P^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -\mu \\ -1 & \lambda \end{bmatrix}.$$

Então

$$P^{-1} A P = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix},$$

portanto

$$P^{-1} A^m P = (P^{-1} A P)^m = \begin{bmatrix} \lambda^m & 0 \\ 0 & \mu^m \end{bmatrix}.$$

Assim, temos que

$$A^m = P \begin{bmatrix} \lambda^m & 0 \\ 0 & \mu^m \end{bmatrix} P^{-1}.$$

Usando o fato de que $\lambda\mu = \det A = -1$ e multiplicando as matrizes acima, obtemos

$$A^m = \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda^{m+1} - \mu^{m+1} & \lambda^m - \mu^m \\ \lambda^m - \mu^m & \lambda^{m-1} - \mu^{m-1} \end{bmatrix} = \begin{bmatrix} f_{m+1} & f_m \\ f_m & f_{m-1} \end{bmatrix},$$

onde

$$f_m = \frac{(\lambda^m - \mu^m)}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^m - \left(\frac{1 - \sqrt{5}}{2} \right)^m \right].$$

Observamos que

$$f_{m+1} + f_m = f_{m+2}. \quad (5-1)$$

Para ver isto, analisaremos o seguinte fato.

Como $\lambda = \frac{1+\sqrt{5}}{2}$, temos que $\lambda^2 = \lambda + 1$. Multiplicando a igualdade por $\frac{\lambda^m}{\sqrt{5}}$, chegaremos ao seguinte resultado.

$$\frac{\lambda^{m+2}}{\sqrt{5}} = \frac{(\lambda^{m+1} + \lambda^m)}{\sqrt{5}}. \quad (5-2)$$

Analogamente, como $\mu = \frac{1-\sqrt{5}}{2}$, temos que $\mu^2 = \mu + 1$. Fazendo o mesmo processo usado na equação 5-2, obtemos

$$\frac{\mu^{m+2}}{\sqrt{5}} = \frac{(\mu^{m+1} + \mu^m)}{\sqrt{5}}. \quad (5-3)$$

Subtraindo a equação 5-2 da equação 5-3, temos

$$\underbrace{\frac{(\lambda^{m+2} - \mu^{m+2})}{\sqrt{5}}}_{f_{m+2}} = \underbrace{\frac{(\lambda^{m+1} - \mu^{m+1})}{\sqrt{5}}}_{f_{m+1}} + \underbrace{\frac{(\lambda^m - \mu^m)}{\sqrt{5}}}_{f_m}.$$

Assim, acabamos de mostrar que equação 5-1 é válida.

Usando a fórmula para o número de m -blocos, obtemos

$$\begin{aligned} |\mathfrak{B}_m(X_G)| &= \sum_{i=1}^2 (A^m)_{IJ} = f_{m+1} + f_m + f_m + f_{m-1} = \\ &= f_{m+2} + f_{m+1} = f_{m+3}. \end{aligned}$$

Tal resultado é obtido através do seguinte cálculo,

$$\begin{aligned} f_{m+2} + f_{m+1} &= \frac{(\lambda^{m+2} - \mu^{m+2})}{\sqrt{5}} + \frac{(\lambda^{m+1} - \mu^{m+1})}{\sqrt{5}} = \frac{\lambda^{m+1}(\lambda+1) - \mu^{m+1}(\mu+1)}{\sqrt{5}} = \\ &= \frac{\lambda^{m+1}(\lambda^2) - \mu^{m+1}(\mu^2)}{\sqrt{5}} = \frac{\lambda^{m+3} - \mu^{m+3}}{\sqrt{5}} = f_{m+3}. \end{aligned}$$

Logo, usando a correspondência injetiva entre $\mathfrak{B}_n(X) = \mathfrak{B}_n(\hat{X}_G)$ e $\mathfrak{B}_{n-1}(X_G)$, obtemos

$$|\mathfrak{B}_n(X)| = |\mathfrak{B}_n(\hat{X}_G)| = |\mathfrak{B}_{n-1}(X_G)| = f_{n+2} = \frac{1}{\sqrt{5}} (\lambda^{n+2} - \mu^{n+2}).$$

Assim a entropia de X é dada por

$$\begin{aligned} h(X) &= \lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_n(X)| = \lim_{n \rightarrow +\infty} \frac{1}{n} \log \left[\frac{1}{\sqrt{5}} (\lambda^{n+2} - \mu^{n+2}) \right] = \\ &= \lim_{n \rightarrow +\infty} \frac{1}{n} \left[\log \frac{1}{\sqrt{5}} + (n+2) \log \lambda + \log \left(1 - \frac{\mu^{n+2}}{\lambda^{n+2}} \right) \right] = \\ &= \lim_{n \rightarrow +\infty} \left(1 + \frac{2}{n} \right) \log \lambda = \log \lambda \end{aligned}$$

Concluimos assim o cálculo da entropia. $\square$

5.1.2

Prova da Proposição 5.2

Já Definimos a entropia $h(X)$ de um espaço de seqüências como o limite $\frac{1}{n} \log |\mathfrak{B}_n(X)|$. Mostraremos que este limite sempre existe. O passo essencial é o seguinte lema:

Lema 5.4 *Sejam $(a_k)_k$ uma seqüência de números não negativos tal que*

$$a_{m+n} \leq a_m + a_n \quad \text{para todo } m, n \geq 1.$$

Então,

$$\lim_{n \rightarrow +\infty} \frac{a_n}{n} = \inf_{n \geq 1} \frac{a_n}{n} = \alpha.$$

Em particular, o limite é finito.

Assumindo o lema, provaremos a proposição. Fixados um par de números naturais $m, n > 1$, temos que se verifica

$$|\mathfrak{B}_{m+n}(X)| \leq |\mathfrak{B}_m(X)| \cdot |\mathfrak{B}_n(X)|.$$

Aplicando-se o logaritmo na desigualdade acima, obtemos:

$$\log \mathfrak{B}_{m+n}(X) \leq \log \mathfrak{B}_m(X) + \log \mathfrak{B}_n(X).$$

Portanto, a seqüência $a_n = \log \mathfrak{B}_n$ verifica as hipóteses do Lema 5.4. A proposição é agora consequência imediata do lema.

Prova do Lema: Por definição, temos que $\frac{a_n}{n} \geq \alpha$, para todo $n \geq 1$. Seja $\varepsilon > 0$. O nosso objetivo é mostrar que para n suficientemente grande,

$$\frac{a_n}{n} < \alpha + \frac{\varepsilon}{2}. \quad (5-4)$$

Como α é a maior das cotas inferiores de $\left\{ \frac{a_n}{n} \right\}_{n \geq 1}$, existe $k \in \mathbb{N}$ tal que

$$\alpha \leq \frac{a_k}{k} < \alpha + \frac{\varepsilon}{4}.$$

Por hipótese, para $0 \leq j < k$ e $m \geq 1$, se verifica

$$\frac{a_{mk+j}}{mk+j} \leq \frac{a_{mk}}{mk+j} + \frac{a_j}{mk+j} \leq \frac{a_{mk}}{mk} + \frac{a_j}{mk+j}.$$

Como, por hipótese, $a_j \leq j a_1$ e $a_{mk} \leq m a_k$ e pois $0 \leq \frac{j}{k} < 1$, temos

$$\frac{a_{mk+j}}{mk+j} \leq m \frac{a_k}{mk} + j \frac{a_1}{mk+j} \leq \frac{a_k}{k} + \frac{a_1}{m} < \alpha + \frac{\varepsilon}{4} + \frac{a_1}{m}.$$

Assim obtemos a equação (5-4) quando n é da forma $mk+j$ e m é grande.

Consideramos k fixo (como acima) e observamos que

$$\begin{aligned} \bigcup_{n \geq 1} \left(\frac{a_n}{n} \right) &= \left(\bigcup_{m \geq 1} \left(\frac{a_{mk}}{mk} \right) \right) \cup \cdots \cup \left(\bigcup_{m \geq 1} \left(\frac{a_{mk+(m-1)}}{mk+(m-1)} \right) \right) = \\ &= \bigcup_{j=0}^{k-1} \left(\bigcup_{m \geq 1} \left(\frac{a_{mk+j}}{mk+j} \right) \right). \end{aligned}$$

Como, pela primeira parte da prova, para $0 \leq j < k$ temos que $\frac{a_{mk+j}}{mk+j} \rightarrow \alpha$ obtemos $\frac{a_n}{n} \rightarrow \alpha$. $\square$

No capítulo 1 vimos que um espaço de seqüências Y é um fator do espaço de seqüências X , se existe um código ϕ de X em Y .

Proposição 5.5 *Se Y é um fator de X , então $h(Y) \leq h(X)$.*

Prova: Seja $\phi : X \rightarrow Y$ um código de X em Y . Então, por definição $\phi = \Phi_\infty^{[-m,k]}$ para uma aplicação bloco Φ e inteiros m, n . Como ϕ é sobrejetora, todo bloco em $\mathfrak{B}_n(Y)$ é a imagem por Φ de um bloco em $\mathfrak{B}_{n+m+k}(X)$. Assim, temos que

$$|\mathfrak{B}_n(Y)| \leq |\mathfrak{B}_{n+m+k}(X)|$$

Como pela proposição 5.2 os limites de $\frac{1}{n} \log |\mathfrak{B}_n(X)|$ e $\frac{1}{n} \log |\mathfrak{B}_{n+m+k}(X)|$ existem, se verifica

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_n(X)| &\leq \lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_{n+m+k}(X)| = \\ &= \lim_{n \rightarrow +\infty} \left(\frac{n+m+k}{n} \right) \frac{1}{n+m+k} \log |\mathfrak{B}_{n+m+k}(X)| \end{aligned}$$

Obtendo assim que $h(Y) \leq h(X)$ e finalmente alcançando o resultado desejado.

$\square$

Uma consequência da proposição anterior é a seguinte:

Proposição 5.6 *Seja $\mathcal{G} = (G, \mathcal{D})$ um grafo com denominações injetor à direita. Então, $h(X_{\mathcal{G}}) = h(X_G)$.*

Prova: Seja $\mathcal{G} = (G, \mathcal{D})$ um grafo com denominações injetor à direita. Conforme visto na Seção 1 do Capítulo 3, temos que $\mathcal{D}_\infty$ é um 1-código de X_G em $X_{\mathcal{G}}$. Sendo assim, pela proposição anterior obtemos que $h(X_{\mathcal{G}}) \leq h(X_G)$. Suponhamos que G tenha k denominações. Então, qualquer bloco em $\mathfrak{B}_n(X_{\mathcal{G}})$ tem no máximo k representações pois $\mathcal{G}$ é injetor à direita. Portanto,

$$|\mathfrak{B}_n(X_{\mathcal{G}})| \geq \frac{1}{k} |\mathfrak{B}_n(X_G)|.$$

Aplicando logaritmo, dividindo a equação por n e fazendo $n \rightarrow \infty$, obteremos $h(X_{\mathcal{G}}) \geq h(X_G)$ $\square$

Examinaremos a seguir a taxa de crescimento do número de pontos periódicos em um espaço de seqüências. Dado um espaço de seqüências X ,

denotamos por $p_n(X)$ o cardinal do sub-conjunto $\mathcal{P}_n(X)$ de X de pontos de período n .

A proposição a seguir relacionada a entropia de um grafo e o crescimento do número de pontos periódicos de X .

Proposição 5.7 *Se X é um espaço de seqüências, então*

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \log p_n(X) \leq h(X).$$

Notacao: $\mathcal{P}_n(X)$ é o conjunto de pontos de período n

Prova: Consideramos o conjunto $\mathcal{P}_n(X)$ de pontos de período n de X . Observamos que a cada ponto periódico x_0 de período n podemos lhe associar pontos $x_1, \dots, x_{n-1}, x_n = x_0$ tais que para todo i existe uma aresta do grafo ligando x_i e x_j . Considerando a família de pontos $x_0, x_1, \dots, x_{n-1}$, obtemos um ponto em $\mathfrak{B}_n(X)$. Obviamente, a princípio, podem existir pontos periódicos aos quais é possível associar vários blocos, em tal caso escolhemos um deles.

Obtemos assim uma aplicação de $p_n(X) \rightarrow \mathfrak{B}_n(X)$. Esta aplicação é injetora por construção. Portanto, $p_n(X) \leq |\mathfrak{B}_n(X)|$ e

$$\frac{1}{n} \log p_n(X) \leq \frac{1}{n} \log |\mathfrak{B}_n(X)|$$

Agora, considerando o limite superior quando $n \rightarrow \infty$ obtemos

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log p_n(X) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathfrak{B}_n(X)| = \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathfrak{B}_n(X)| = h(X),$$

onde as igualdades seguem da Proposição 5.2. $\square$

5.2

Teoria de Perron-Frobenius

Seja G um grafo com matriz adjacente $A_{r \times r}$. Já vimos na seção 2 do capítulo 2 que o número de n -blocos em um shift de arestas X_G é dado por

$$|\mathfrak{B}_n(X_G)| = \sum_{I, J=1}^2 (A^n)_{IJ},$$

onde $(A^n)_{IJ}$ é a seqüência de arestas e_i de G com $1 \leq i \leq n-1$.

Para calcularmos a entropia do shift de arestas X_G , precisaremos encontrar a taxa de crescimento das entradas $(A^n)_{IJ}$ da matriz A^n . A teoria de Perron-Frobenius é precisamente a ferramenta que nos permitirá fazer isso quando G for um grafo irredutível. Também mostrará que a taxa de cresci-

mento das entradas $(A^n)_{IJ}$ está controlada pelo maior autovalor da matriz A .

Seja $A = (a_{IJ})$ uma matriz quadrada $r \times r$ não-negativa, isto é, $a_{IJ} \geq 0$ para todo $1 \leq I, J \leq r$. Suponha que A não seja a matriz nula. Suponhamos agora, que A tenha um autovetor v da forma

$$v = (v_1, v_2, \dots, v_r), \quad v_I > 0.$$

Em tal caso, dizemos que v é um *autovetor positivo*.

Considere agora o autovalor λ associado ao autovetor positivo v . Como a matriz é positiva e o autovetor é positivo, o autovalor λ é necessariamente não nulo. Aplicando A em $A v = \lambda v$, obtemos

$$A^2 v = \lambda A v = \lambda^2 v.$$

Indutivamente,

$$A^n v = \lambda^n v,$$

para todo $n \geq 1$. Portanto, para todo I temos:

$$\sum_{J=1}^r (A^n)_{IJ} v_J = \lambda^n v_I.$$

De fato, usando o processo de indução sobre n , temos que para $n = 1$ o resultado é óbvio. Suponhamos o resultado válido para $n = k$. Mostraremos a seguir, que para $n = k + 1$ o resultado também é válido.

$$\sum_{J=1}^r (A^{k+1})_{IJ} v_J = \sum_{J=1}^r (A^k A)_{IJ} v_J.$$

Sejam

$$A^k = B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{I1} & b_{I2} & \cdots & b_{In} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \quad \text{e} \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1J} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2J} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nJ} & \cdots & a_{nn} \end{bmatrix}.$$

Logo,

$$\begin{aligned}
 BA &= \sum_{J=1}^r \sum_{m=1}^r (B_{Im} A_{mJ}) v_J = \sum_{J=1}^r B_{Im} \sum_{m=1}^r A_{mJ} v_J = \\
 &= \sum_{J=1}^r B_{Im} \lambda v_J = \lambda \sum_{J=1}^r B_{Im} v_J = \lambda \sum_{J=1}^r A^k v_J = \\
 &= \lambda (\lambda^k) v_I = \lambda^{k+1} v_I.
 \end{aligned}$$

Assim, mostramos que o processo de indução sobre n é válido.

Sejam $c = \min\{v_1, v_2, \dots, v_r\} > 0$ e $d = \max\{v_1, v_2, \dots, v_r\} > 0$. Então,

$$c \sum_{J=1}^r (A^n)_{IJ} \leq \sum_{J=1}^r (A^n)_{IJ} v_J = \lambda^n v_I \leq d \lambda^n$$

Dividindo a desigualdade acima por c e somando sobre I , obtemos:

$$\sum_{I,J=1}^r (A^n)_{IJ} \leq \sum_{I=1}^r \frac{d}{c} \lambda^n = \frac{r d}{c} \lambda^n = d_0 \lambda^n; \text{ onde } d_0 = \frac{r d}{c} > 0.$$

Por outro lado, note que para cada I ,

$$c \lambda^n \leq \lambda^n v_I = \sum_{J=1}^r (A^n)_{IJ} v_J \leq d \sum_{J=1}^r (A^n)_{IJ} \leq d \sum_{I,J=1}^r (A^n)_{IJ}.$$

Tomando $c_0 = \frac{c}{d} > 0$ e dividindo a equação acima por d obtemos:

$$c_0 \lambda^n \leq \sum_{I,J=1}^r (A^n)_{IJ}.$$

Logo das equações acima, concluímos que:

$$c_0 \lambda^n \leq \sum_{I,J=1}^r (A^n)_{IJ} \leq d_0 \lambda^n.$$

A discussão acima pode ser resumida na seguinte proposição:

Proposição 5.8 *Seja $A \neq 0$ uma matriz não-negativa com um autovetor positivo v . Então o autovalor λ associado a v é estritamente positivo e existem constantes estritamente positivas c_0 e d_0 tais que:*

$$c_0 \lambda^n \leq \sum_{I,J=1}^r (A^n)_{IJ} \leq d_0 \lambda^n.$$

Os argumentos acima nos levam ao seguinte fato:

Corolário 5.9 *Sejam G um grafo cuja matriz adjacente é A e λ um autovalor de A associado a um autovetor positivo v . Então*

$$h(X_G) = \log \lambda.$$

Esse corolário nos garante que dado um shift de arestas X_G associado a uma matriz positiva A com autovetor positivo v , sua entropia será o logaritmo do autovalor associado ao autovetor v .

Prova: Em primeiro lugar, observamos que a matriz de adjacência de G é positiva. Portanto, podemos aplicar a Proposição 5.8 e obtermos que

$$c_0 \lambda^n \leq \sum_{I,J=1}^r (A^n)_{IJ} \leq d_0 \lambda^n.$$

Já vimos na Seção 2 do Capítulo 2 que

$$|\mathfrak{B}_n(X_G)| = \sum_{I,J=1}^r (A^n)_{IJ}.$$

Logo,

$$c_0 \lambda^n \leq |\mathfrak{B}_n(X_G)| \leq d_0 \lambda^n.$$

Aplicando logaritmo na última desigualdade obtemos:

$$\log c_0 + n \log \lambda \leq \log |\mathfrak{B}_n(X_G)| \leq \log d_0 + n \log \lambda.$$

Agora multiplicaremos por $\frac{1}{n}$ e obteremos:

$$\frac{\log c_0}{n} + \log \lambda \leq \frac{1}{n} \log |\mathfrak{B}_n(X_G)| \leq \frac{\log d_0}{n} + \log \lambda$$

Finalmente, considerando limites quando $n \rightarrow \infty$, temos que

$$\log \lambda \leq \underbrace{\lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_n(X_G)|}_{h(X_G)} \leq \log \lambda$$

Portanto, $h(X_G) = \log \lambda$, concluindo a prova do corolário. $\square$

A suposição de que a matriz A possui autovetor positivo implica em duas consequências adicionais.

Lema 5.10 *Seja A uma matriz quadrada positiva.*

- Todos os autovalores não nulos associados a autovetores positivos de A são iguais.
- Se λ é um autovalor não nulo associado a um autovetor positivo então $\lambda \geq |\mu|$ para todo autovalor de A .

Prova: Para ver a primeira parte, suponha que u é um autovetor positivo de A e que θ seja seu autovalor. Então $\theta = \lambda$. Isto é, todos os autovalores associados a autovetores positivos são iguais. Para ver isto é suficiente aplicar os argumentos acima ao autovetor u e ao autovalor θ , obtendo que $h(X_G) = \log \theta$, o que implica que $\theta = \lambda$.

O autovalor associado λ é positivo. Defina $d_0 = \frac{rd}{c} > 0$, onde $r > 0$, $c = \min\{v_1, v_2, \dots, v_r\} > 0$ e $d = \max\{v_1, v_2, \dots, v_r\} > 0$. Suponha que μ é um autovalor de A e w é um autovetor associado ao autovalor λ . Observe que μ e λ podem envolver números complexos. Defina a norma de w por $\|w\| = \sum_{I=1}^r |w_I|$. Para $n \geq 1$, temos que $A^n w = \mu^n w$. Temos a desigualdade seguinte:

$$\begin{aligned} |\mu|^n \|w\| &= \|\mu^n w\| = \|A^n w\| = \\ &= \left| \sum_{J=1}^r (A^n)_{1,J} w_J \right| + \dots + \left| \sum_{J=1}^r (A^n)_{r,J} w_J \right| \leq \\ &\leq \sum_{J=1}^r (A^n)_{1,J} |w_J| + \dots + \sum_{J=1}^r (A^n)_{r,J} |w_J| \leq \\ &\leq \left(\max_{1 \leq J \leq r} |w_J| \right) \sum_{I,J=1}^r (A^n)_{I,J} \leq \|w\| d_0 \lambda^n, \end{aligned}$$

Portanto, como $\|w\| \neq 0$ temos que

$$|\mu| \leq (d_0)^{\frac{1}{n}} \lambda.$$

Considerando o limite quando $n \rightarrow \infty$ obtemos $|\mu| \leq \lambda$. Portanto, λ é o maior autovalor de A . $\square$

Definição 5.11 (Matriz irredutível) Dizemos que uma matriz A não-negativa é irredutível se para cada par ordenado de índices I, J existe algum $n \geq 0$ tal que $A_{I,J}^n > 0$. Uma matriz A não-negativa é essencial, se nenhuma de suas linhas ou colunas é nula.

A teoria de Perron-Frobenius mostrará que uma matriz irredutível sempre terá um autovetor positivo.

Teorema 5.12 (Perron-Frobenius) *Seja $A \neq 0$ uma matriz irredutível. Então A possui um autovalor $\lambda_A > 0$ tal que $|\mu| \leq \lambda_A$ para qualquer outro autovalor μ de A .*

Para uma matriz irredutível A , chamaremos λ_A de autovalor de Perron de A e aos autovetores v_A associados a λ_A de autovetores de Perron de A .

De fato, o Teorema de Perron-Frobenius tem um enunciado um pouco mais geral que o enunciado por nós: o autovalor de Perron $\lambda_A > 0$ é geometricamente e algebricamente simples. Lembramos que um autovalor λ de A é *geometricamente simples*, se o seu autoespaço correspondente tem dimensão 1 e dizemos que λ é algebricamente simples se for raiz simples do polinômio característico de A .

Usaremos na demonstração do Teorema 5.12 dois lemas que serão enunciados a seguir e o Teorema do Ponto Fixo de Brouwer. Pelo Lema 5.10, é suficiente provar que a matriz A possui um autovetor v_A positivo, então o autovalor λ_A associado a v_A verificará o teorema.

Lema 5.13 *Se A é uma matriz quadrada não-negativa cuja i -ésima coluna é nula. Então todas as potências de A têm a i -ésima coluna nula.*

Prova: Sejam $C = (c_{ij})$ e $B = (b_{ij})$ duas matrizes quadradas $n \times n$ cuja k -ésima coluna é nula, isto é, $c_{jk} = b_{jk} = 0$, para todo $j = 1, \dots, n$. Temos que $CB = F = (f_{ij})$, onde

$$f_{jk} = \sum_{r=1}^n c_{jr} b_{rk} = \sum_{r=1}^n (c_{jr}) 0 = 0.$$

Como esta igualdade vale para todo j , a k -ésima coluna de F é nula.

Por indução, veremos que o resultado desejado é válido. Para $n = 2$, fazemos $B = A$ e $C = A$ e, pelo comentário acima, temos

$$A^2 = BC = 0.$$

Suponhamos o resultado válido para $n = k$. Mostraremos que para $n = k + 1$ o resultado também é válido. Observe que,

$$D = A^{k+1} = A \cdot A^k.$$

Fazendo $B = A^k$ e $C = A$ temos que

$$A^{k+1} = A^k A = B C = 0.$$

Concluindo assim a demonstração do Lema. $\square$

Lema 5.14 *Dada uma matriz $A = (a_{ij})$ irredutível $m \times m$ considere um vetor $v = (v_1, v_2, \dots, v_m)$ tal que $v_i \geq 0$ e $\sum v_i > 0$. Então, $Av \neq 0$.*

Prova: Escrevemos $Av = w$ e escolhemos $v_i > 0$ (tal v_i existe por hipótese). Como todas as entradas da matriz A são não-negativas e v é um vetor positivo, temos que, para todo $k = 1, \dots, m$,

$$w_k \geq a_{ki} v_i \geq 0.$$

Suponhamos, por absurdo, que $Av = 0$. Então $w_k = 0$ para todo k e portanto, $a_{ki} = 0$ para todo k . Assim a i -ésima coluna A deve ser nula.

Pelo Lema 5.13 segue que todas as potências de A tem i -ésima coluna nula, o que é um absurdo, pois A é irredutível. Assim, concluimos que $Av \neq 0$ e obtemos o resultado desejado. $\square$

Teorema 5.15 (Ponto fixo de Brouwer, (Gre)) *Seja*

$$\Delta = \{x \in \mathbb{R}^m; \quad x \geq 0, \quad \sum x_i = 1\}.$$

Então, qualquer aplicação contínua $f: \Delta \rightarrow \Delta$ tem um ponto fixo.

Com estes resultados podemos agora demonstrar nossa versão do Teorema de Perron-Frobenius (Teorema 5.12).

Prova do Teorema: [Perron-Frobenius] Sejam A uma matriz irredutível. Pelo Lema 5.10, é suficiente provar que A possui um autovetor positivo.

Consideremos o seguinte conjunto,

$$\Delta = \{x \in \mathbb{R}^m; \quad x \geq 0, \quad \sum x_i = 1\}.$$

Dado um vetor $v = (v_1, v_2, \dots, v_m)$, $v_i \geq 0$ e $\sum v_i > 0$, segue pelo Lema 5.14 que $Av \neq 0$. Como A é não negativa e v é positivo isto implica que

$$\sum_{i=1}^m (Av)_i > 0.$$

Isto nos permite Considerar a seguinte aplicação contínua:

$$\begin{aligned}\Phi : \Delta &\longrightarrow \Delta \\ v &\longmapsto \frac{A v}{\sum (A v)_i}.\end{aligned}$$

Como Φ é contínua, pelo Teorema do ponto fixo de Brouwer existe $w \in \Delta$ tal que $\Phi(w) = w$. Assim, obtemos

$$\frac{A w}{\sum (A w)_i} = w.$$

Fazendo $\tilde{\lambda} = \sum (A w)_i > 0$, temos que

$$A w = \tilde{\lambda} w; \quad \text{onde} \quad \tilde{\lambda} > 0.$$

Assim, obtemos um autovetor positivo w . Pelo Lema 5.10, o autovalor $\tilde{\lambda}$ associado a w verifica a conclusões do Teorema de Perron-Frobenius. $\square$

5.3

Cálculo da Entropia

Seja X um shift de tipo finito, então existem um grafo G e uma matriz de adjacência A associada ao grafo tal que o shift X é conjugado ao shift de arestas de G , X_G , com matriz adjacente A . Usaremos este fato para provar que $h(X) = \log \lambda_A$, onde λ_A é o maior autovalor da matriz A , veja o Teorema 5.16. Calcularemos também a entropia de um shift sófico irredutível e veremos que a taxa de crescimento dos pontos periódicos do shift é igual a sua entropia. Ver os Exemplos 21 e 22.

Teorema 5.16 *Considere X um shift de tipo finito e um grafo G cuja matriz de adjacência é A .*

1. *Suponha que $X = X_G$ e que G é um grafo irredutível, então $h(X_G) = \log \lambda_A$, onde λ_A o maior autovalor da matriz A .*
2. *Se X é um shift de tipo finito irredutível com memória M e G é um grafo essencial tal que $X^{[M+1]} = X_G$, então $h(X) = \log \lambda_{A(G)}$, onde λ_A o maior autovalor da matriz A .*

Prova: Para provar o primeiro item observe que, por definição, a matriz de adjacência A do grafo G é irredutível. Pelo Teorema de Perron-Frobenius, a matriz A possui um autovalor positivo máximo λ_A . Assim, como G é um grafo

cujas matriz adjacente é A com maior autovalor λ_A , pelo Corolário 5.9 segue que:

$$h(X) = h(X_G) = \lambda_{A(G)}$$

Para provar a segunda parte observe que, por hipótese, temos que X tem memória M e $X = X_G$, logo pelo Teorema 3.17, $X_G = X \cong X^{[M+1]}$. Portanto, a entropia de X é dada por:

$$h(X) = h(X^{[M+1]}) = h(X_G)$$

Mas X é irredutível, logo A também o é. Aplicando o primeiro item do teorema concluímos que

$$h(X) = h(X^{[M+1]}) = h(X_G) = \lambda_{A(G)}.$$

A prova do teorema está concluída. $\square$

Exemplo 21 *Seja $\mathcal{A} = \{0, 1\}$, $\mathcal{P} = \{111\}$ e $X = X_{\mathcal{P}}$. Então X é um shift completo de dois símbolos. Observamos que não existem em X seqüências com três 1's consecutivos, portanto X tem memória dois. Por outro lado, o grafo G com $X^{[3]} = X_G$ tem posições 00, 01, 10 e 11 e sua matriz adjacente é*

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Observe que

$$A^2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

e que

$$A^3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 & 1 \\ 2 & 2 & 1 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

portanto, a matriz A é irredutível. Assim, podemos aplicar o Teorema 5.16. Assim, $h(X) = \log \lambda_A$. Onde $\mathcal{X}_A(t) = t^4 - t^3 - t^2 - t$, cujas raízes são $t = 0$, $t \approx -0.41964 \pm 0.60629i$ e $t \approx 1.83929$. Portanto

$$h(X) \approx \log 1.83929 \approx 0.87915.$$

Teorema 5.17 *Sejam X um shift sófico irreduzível, $\mathcal{G} = (G, \mathcal{D})$ uma apresentação injetora a direita de X e G um grafo irreduzível. Então $h(X) = \log \lambda_{A(G)}$.*

Prova: Pela Proposição 5.6, temos que $h(X) = h(X_{\mathcal{G}}) = h(X_G)$. Como G é irreduzível, podemos aplicar o teorema anterior, obtendo $h(X) = h(X_G) = \log \lambda_{A(G)}$. $\square$

Exemplo 22 *Seja X um espaço de seqüências. A figura abaixo nos mostra uma apresentação injetora a direita irreduzível de X . A matriz adjacente do grafo subjacente é dada por*

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Então, $\mathcal{X}_A(t) = t^4 - t^2 - t - 1$, cujo a maior raiz é $\lambda_A \approx 1.46557$. Sendo assim, $h(X) = \log \lambda_A \approx 0.55146$.

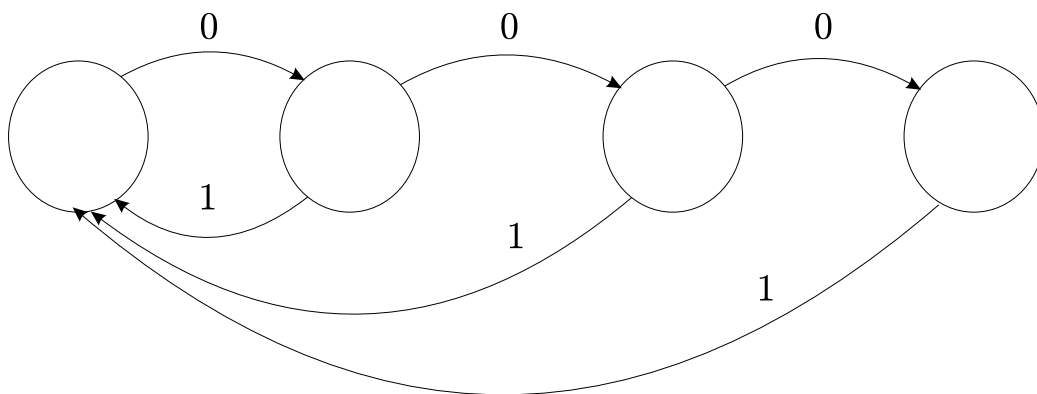


Figura 5.2: Injetor à direita

Teorema 5.18 *Sejam X um shift sófico irreduzível e $p_n(X)$ o cardinal do espaço de seqüências X de pontos de período n . Então:*

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \log p_n(X) = h(X).$$

Prova: Pela Proposição 5.7 a desigualdade

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \log p_n(X) \leq h(X)$$

é válida para qualquer espaço de seqüências X . Portanto, para provarmos o teorema é suficiente ver que a desigualdade em sentido inverso também é válida.

Em primeiro lugar, analisaremos o caso em que X é um shift irredutível de tipo finito. Neste caso, existe um grafo G irredutível tal que $X = X_G$. Obviamente, $p_n(X) = p_n(X_G)$ e $h(X) = h(X_G)$.

A irredutibilidade de G implica na existência de um número natural $N > 0$ tal que para todo par I, J de vértices de G , existe um caminho de I a J de tamanho menor ou igual a N . Assim, se $\pi \in \mathfrak{B}_n(X_G)$, então existe um caminho τ de tamanho menor ou igual a N de $t(\pi)$ (a posição final de π) para $i(\pi)$ (a posição inicial de π). Logo, $\pi\tau$ é um ciclo e $(\pi\tau)^\infty$ tem período $n + |\tau|$, que está entre n e $n + N$. Dessa maneira, cada bloco em $\mathfrak{B}_n(X_G)$ gera um ponto em X_G , cujo o período está entre n e $n + N$. Obviamente, blocos diferentes geram diferentes pontos periódicos. Portanto,

$$|\mathfrak{B}_n(X_G)| \leq p_n(X_G) + p_{n+1}(X_G) + \cdots + p_{n+N}(X_G).$$

Para cada n , existe $k(n) \in [0, N]$ tal que

$$p_{n+k(n)}(X_G) \geq p_{n+i}(X_G), \quad \text{para todo } i \in [0, N].$$

Portanto,

$$|\mathfrak{B}_n(X_G)| \leq (N + 1) p_{n+k(n)}(X_G).$$

Assim, pela definição de entropia,

$$\begin{aligned} h(X) = h(X_G) &= \lim_{n \rightarrow +\infty} \frac{1}{n} \log |\mathfrak{B}_n(X_G)| \leq \\ &\leq \limsup_{n \rightarrow +\infty} \left[\frac{1}{n} \log (N + 1) + \frac{1}{n} \log p_{n+k(n)}(X_G) \right] \leq \\ &\leq \limsup_{m \rightarrow +\infty} \frac{1}{m} \log p_m(X_G). \end{aligned}$$

Provamos assim o resultado para shifts irredutíveis de tipo finito.

Suponhamos agora que X é um shift sófico irredutível. Seja $\mathcal{G} = (G, \mathcal{D})$ uma apresentação irredutível injetora a direita do espaço de seqüências X . Da Proposição 5.6 temos que

$$h(X) = h(X_{\mathcal{G}}) = h(X_G).$$

Suponhamos que G tenha r vértices. Como $\mathcal{G}$ é injetor à direita, temos que cada ponto em $X_{\mathcal{G}}$ tem r representações em X_G . Como $\mathcal{D}_\infty$ leva pontos de

período n de X_G em pontos de período n de X_G , segue que:

$$p_n(X_G) \geq \frac{1}{r} p_n(X_G)$$

Assim, temos que

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \frac{1}{n} \log p_n(X_G) &\geq \limsup_{n \rightarrow +\infty} \left[\frac{1}{n} \log \frac{1}{r} + \frac{1}{n} \log p_n(X_G) \right] = \\ &= \limsup_{n \rightarrow +\infty} \frac{1}{n} \log p_n(X_G) = h(X_G) = h(X_G), \end{aligned}$$

concluindo a prova do teorema. □