Symbolic Dynamics

Zhengbo Qu
2000011611@stu.pku.edu.cn

Contents

We mainly discuss the Chapter 3 in [1], since the time is limited, we may omit some details and facts that are less important.

We first recall some notations and definitions that will be widely used in this chapter.

We identify every finite *alphabet* with $\{1, 2, ..., m\}$. And let $\Sigma_m = \{1, 2, ..., m\}^{\mathbb{Z}}$ and $\Sigma_m^+ = \{1, 2, ..., m\}^{\mathbb{N}_0}$.

The *cylinder* sets:

$$C_{j_1,\dots,j_k}^{n_1,\dots,n_k} = \left\{\omega = (\omega_l) : \omega_{n_i} = j_i, i = 1, ..., k\right\}$$

form a basis for the product topology of Σ_m and Σ_m^+.

The *metric*:

$$d(\omega, \omega') = 2^{-l},\ \ l = \min\{|i| : \omega_i \neq \omega_{i'}\}$$

generates this product topology.

1. Subshifts and Codes

Definition 1.1: A **subshift** is a closed subset $X \subset \Sigma_m$ invariant under the shift σ and its inverse. We refer to Σ_m as the **full m-shift**.

Definition 1.2: Let $X_i \subset \Sigma_{m_i}, i = 1, 2$ be two subshifts, a continuous map $c : X_1 \to X_2$ is a **code** if it commutes with the shifts.

A surjective code is a *factor map*. An injective code is called an *embedding*. A bijective code gives a topological conjugacy of the subshifts and is called an *isomorphism*.

Definition 1.3: For a subshift $X \subset \Sigma_m$, we define the $W_{n(X)}$ to be the set of words of length n that occur in X and denote its cardinality by $|W_{n(X)}|$.

Definition 1.4: Let X be a subshift, $k, l \in \mathbb{N}_0, n = k + l + 1$, and let α be a map from $W_{n(X)}$ to an alphabet $A_{m'}$. The (k, l) **block code** c_α from X toe the full shift $\Sigma_{m'}$ assigns to a sequence $x = (x_i) \in X$the sequence $c_{\alpha(x)}$ with $c_{\alpha(x)_i} = \alpha\left(x_{i-k}, ..., x_i, ..., x_{i+l}\right)$

The definition above is a little abstract, so we give an example here.

Example:

Let $X = \Sigma_2, k = l = 1$ and $n = 3$.

For $(ijk) \in W_3(X)$, we define α to be:

$$\alpha(ijk) = 1 \text{ if } i + j + k \text{ is odd;}$$

$$\alpha(ijk) = 2 \text{ if } i + j + k \text{ is even.}$$

Then compute (...1212121...) and (...111222111222...).

Any block code is a code since it is continuous and commutes with the shift. Conversely, we have the following proposition:

Proposition 1.1: Every code $c : X \to Y$ is a block code.

Proof: Let $\mathcal{A}$ be the symbol set of Y, and define $\overline{a} : X \to \mathcal{A}$ by $\overline{a}(x) = c(x)_0$.

Since X is compact, $\overline{a}$ is uniformly continuous, so there is a $\delta > 0$ such that $\overline{a}(x) = \overline{a}(x')$ whenever $d(x, x') < \delta$.

Choose $k \in \mathbb{N}$ so that $2^{-k} < \delta$, then $\overline{a}$ only depends on $x_{-k}, ..., x_0, ..., x_k$, and therefore we can define a map $\alpha : W_{2k+1} \to \mathcal{A}$ satisfying $c(x)_0 = \alpha(x_{-k}, ..., x_0, ..., x_k)$. Since c commutes with the shift, we conclude that $c = c_\alpha$. □

2. Subshifts of Finite Type

The complement of a subshift X is open and hence is a union of at most countably many cylinders. If C is a cylinder that is a subset of X, then for any $n \in \mathbb{Z}$, $\sigma^{n(C)}$.

That is to say, there is a countable list of forbidden words such that no sequence in X contains them, and each sequence in $\Sigma_m \setminus X$ contains at least one forbidden word.

Definition 2.1: If this list is finite, we call X a **subshift of finite type(SFT)**. Furthermore, X is a k-step SFT if it is defined by a set of words of length at most $k + 1$. A 1-step SFT is called a **topological Markov chain**.

We introduce a vertex shift Σ_A^v determined by an adjacency matrix A of zeros and ones. A *vertex shift* is an example of an SFT, since the forbidden words have length 2 and are precisely those that are not allowed by A. A sequence in Σ_A^v can be viewed as an infinite path in the directed graph Γ_A. Furthermore, we have the following proposition, which says that:

Proposition 2.1: Every SFT is isomorphic to a vertex shift.

Proof: Let X be a k-step SFT with $k > 0$.

Let Γ be the directed graph whose set of vertices is $W_{k(X)}$, a vertex $x_1...x_k$ is connected to a vertex $x_{1'}...x_{k'}$ by a directed edge if $x_1...x_k x_{k'} = x_1 x_{1'}...x_{k'} \in W_{k+1}(X)$.

Let A be the adjacency matrix of Γ, the code $c(x)_i = x_i...x_{i+k-1}$ gives an isomorphism from X to Σ_A^v □

An alternative to describe an infinite path in the graph Γ_A is *edge shift*:

Definition 2.2: A finite directed graph Γ, possibly with multiple directed edges connecting pairs of vertices, corresponds to an adjacency matrix B whose i, j-th entry is a non-negative integer that is the number of directed edges in Γ from the i-th vertex to the j-th vertex.

Note that vertex shift and edge shift are isomorphic.

3. The Perron-Frobenius Theorem

In this section, we return to linear algebra and focus on a kind of matrix that is useful in the dynamical system.

Definition 3.1:
1. A vector or matrix all of whose coordinates are positive (non-negative) is called **positive (non-negative)**
2. Let A be a square non-negative matrix. If for any i, j there is $n \in \mathbb{N}$ such that $(A^n)_{ij} > 0$, then A is called **irreducible**; otherwise A is called **reducible**.
3. If some power of A is positive, A is called **primitive**.

We now record some basic properties of primitive matrices.

1. An integer non-negative square matrix A is primitive, if and only if, the directed graph Γ_A has the property that there is $n \in \mathbb{N}$ such that, for every pair of vertices u and v, there is a directed path from u to v of length n.
2. An integer non-negative square matrix A is irreducible if and only if the directed graph Γ_A has the property that, for every pair of vertices u and v, there is a directed path from u to v.

We now try to prove the main theorem of this chapter, I have searched for various proofs, and I find that the proof in this book[1] is one of the simplest.

Theorem 3.1: Let A be a primitive $m \times m$ matrix. Then A has a positive eigenvalue λ with the following properties:
1. λ is a simple root of the characteristic polynomial of A,
2. λ has a positive eigenvector v,
3. Any other eigenvalue of A has a modulus strictly less than λ,
4. Any non-negative eigenvector of A is a positive multiple of v.

To prove this, we need to introduce a lemma first.

Lemma 3.1: Denote by $\operatorname{int}(W)$ the interior of a set W. Let $L : \mathbb{R}^k \to \mathbb{R}^k$ be a linear operator. and assume that there is a non-empty compact set P such that $0 \in \operatorname{int}(P)$ and $L^{i(P)} \subset \operatorname{int}(P)$ for some $i > 0$. Then the modulus of any eigenvalue of L is strictly less than 1.

Proof: We may assume that $L(P) \subset \operatorname{int}(P)$, otherwise for any $i > 0$, $L^{i(P)} \subset \operatorname{int}(P)$ is impossible. Therefore, $L^{n(P)} \subset \operatorname{int}(P)$ for any $n \in \mathbb{N}$. The matrix L cannot have an eigenvalue of modulus greater than 1, otherwise, the iterates of L would move some vector in the open set $\operatorname{int}(P)$ to ∞.

Now suppose that there is an eigenvalue σ with $|\sigma| = 1$. If $\sigma^j = 1$, then there exists a point on ∂P, impossible. If σ is not a root of unity, there is a 2-dimensional subspace U on which L acts as an irrational rotation and any point $p \in \partial P \cap U$ is a limit point of $\cup_{n>0}$ $L^{n(P)} = L(P)$, a contradiction. □

Now we can prove the Perron-Frobenius Theorem. Recall that a real non-negative $m \times m$ matrix is **stochastic** if the sum of the entries in each row is 1.

Proof: Since A is non-negative, it induces a continuous map f from the unit simplex $S = \{x \in \mathbb{R}^m : \Sigma x_j = 1, x_j \geq 0, j = 1, ..., m\}$ into itself; $f(x)$ is the radial projection of Ax onto S. By the Brouwer fixed point theorem, there is a fixed point $v \in S$ of f , which is a non-negative eigenvector of A with eigenvalue $\lambda > 0$. Since some power of A is positive, all coordinates of v are positive.

Let V be the diagonal matrix that has the entries of v on the diagonal. The matrix $M = \lambda^{-1} V^{-1} A V$ is primitive, and the column vector 1 with all entries 1 is an eigenvector of M with eigenvalue 1, i.e., M is a stochastic matrix.

To prove parts 1 and 3, it suffices to show that 1 is a simple root of the characteristic polynomial of M and that all other eigenvalues of M have moduli strictly less than 1. Consider the action of M on row vectors. Since M is stochastic and non-negative, the row action preserves the unit simplex S. By the Brouwer fixed point theorem, there is a fixed row vector$w \in S$, all of whose coordinates are positive. Let $P = S - w$ be the translation of S by $-w$. Since for some $j > 0$ all entries of M_j are positive, $M^{j(P)} \in \operatorname{int}(P)$ and, by Lemma 3.1, the modulus of any eigenvalue of the row action of M in the $(m-1)$-dimensional invariant subspace spanned by P is strictly less than 1.

The last statement of the theorem follows from the fact that the codimension-one subspace spanned by P is M^t-invariant and its intersection with the cone of non-negative vectors in $\mathbb{R}^n$ is $\{0\}$. □

4. Topological Entropy and the Zeta Function of an SFT

Proposition 4.1: Let $X \subset \Sigma_m$ be a subshift. Then

$$h(\sigma\ |_X) = \lim_{n\to\infty} \frac{1}{n} \log|W_{n(X)}|$$

Proof: We first recall a proposition in Chapter 2, which says that:

Let (X, d) be a compact metric space, and $f : X \to X$ an expansive homeomorphism with expansiveness constant δ. Then $h(f) = h_{\varepsilon(f)}$ for any $\varepsilon < \delta$.

In this case, $\delta = 1$ is an expansiveness constant, and we let $\varepsilon = \frac{1}{3}$. Hence $h(\sigma\ |_X) = h_{\frac{1}{3}}(\sigma\ |_X) = h_{\frac{2}{3}}(\sigma\ |_X)$

We now prove the proposition.

For $W_{n+2}(X)$ with $n \in \mathbb{N}$, we define the corresponding $A_{n+2}(X)$ as following:

For $w_i \in W_{n(X)}, i = 1, 2, ..., |W_{n(X)}|$, let $x_i \in X$ be a sequence with

$$(x_i)_{-1}(x_i)_0(x_i)_1...(x_i)_n = w_i$$

and

$$A_{n+2}(X) = \left\{x_1, ..., x_{|W_{n(X)}|}\right\}.$$

We now prove that $A_{n+2}(X)$ is both $\left(n, \frac{1}{3}\right)$-spanning and $\left(n, \frac{1}{3}\right)$-separated.

1. $A_{n+2}(X)$ is $\left(n, \frac{1}{3}\right)$-spanning:

   For every $y \in X$, suppose that $y_{-1}y_0y_1...y_n = w_k$. Let $x_k \in A_{n+2}(X)$, thus $y_j = (x_k)_j$ for $-1 \le j \le n$.

   For $0 \le j \le n-1$, we have:

   $$\sigma_{-1}^{j(y)}\sigma_0^{j(y)}\sigma_1^{j(y)} = \sigma_{-1}^{j(x_k)}\sigma_0^{j(x_k)}\sigma_1^{j(x_k)}.$$

   Therefore, $d_{n(x_k,y)} \le \frac{1}{4} < \frac{1}{3} = \varepsilon$

2. $A_{n+2}(X)$ is $\left(n, \frac{1}{3}\right)$-separated:

   Let $x, y \in A_{n+2}(X)$ be two distinct sequences: then there exists $-1 \le j \le n$, such that $x_j \ne y_j$.

   If $j = -1$ or 0, $d_{n(x,y)} \ge \frac{1}{2}$ is obvious.

   If $1 \le j \le n$, $d\left(\delta^{j-1}(x), \delta^{j-1}(y)\right) = \frac{1}{2}$, also we have $d_{n(x,y)} \ge \frac{1}{2}$.

Hence we know that $A_{n+2}(X)$ is both $\left(n, \frac{1}{3}\right)$-spanning and $\left(n, \frac{1}{3}\right)$-separated. Since $|A_{n+2}(X)| = |W_{n+2}(X)|$, we have:

$$\text{span}\left(n, \frac{1}{3}, \sigma\right) \leq |W_{n+2}(X)| \leq \text{sep}\left(n, \frac{1}{3}, \sigma\right)$$

Since

$$\text{cov}(n, 2\varepsilon, f) \leq \text{span}(n, \varepsilon, f) \leq \text{sep}(n, \varepsilon, f) \leq \text{cov}(n, \varepsilon, f),$$

we have

$$\text{cov}\left(n, \frac{2}{3}, \sigma\right) \leq |W_{n+2}(X)| \leq \text{cov}\left(n, \frac{1}{3}, \sigma\right).$$

We now consider $\frac{1}{n}\log|W_{n(X)}|$ by taking limits of n, we have

$$\lim_{n\to\infty} \frac{1}{n}\log|W_{n(X)}| = h_{\frac{1}{3}}(\sigma \mid_X) = h_{\frac{2}{3}}(\sigma \mid_X) = h(\sigma \mid_X)$$

□

And in this section, we can compute the topological entropy of an edge shift and introduce the zeta function, an invariant that collects combinatorial information about the periodic points.

Proposition 4.2: Let A be a square non-negative integer matrix. Then the topological entropy of the edge shift Σ_A^e and the vertex shift Σ_A^v equals the logarithm of the largest eigenvalue of A

To prove this proposition, we first introduce a lemma.

Lemma 4.1: Let A be a non-negative, non-zero, square matrix, S_n the sum of entries of A^n, and λ the eigenvalue of A with the largest modulus, then:

$$\lim_{n\to\infty} \frac{\log S_n}{n} = \log \lambda$$

Proof: We just sketch the proof here. For convenience, we write $S_{n(M)}$ for the sum of entries of M^n.

We first suppose that $|\lambda| = 1$ and prove that $\lim_{n\to\infty} \frac{\log S_n}{n} = 0$, when $|\lambda| \neq 1$ we just consider the matrix $\frac{1}{|\lambda|}A$.

Next, we write its canonical form $J = P^{-1}AP$, the diagonal entries of J are A's eigenvalues. Furthermore, we take Q as the matrix transformed from J and all diagonal entries are 1.

Thus we can show that $|S_{n(J)}| \leq S_{n(Q)}$, and it is easy to see that $\lim_{n\to\infty} \frac{\log S_{n(Q)}}{n} = 0$. Hence $\lim_{n\to\infty} \frac{\log S_{n(J)}}{n} = 0$. Since $J^n = P^{-1}A^nP$, it is not difficult to show that $\lim_{n\to\infty} \frac{\log S_{n(J)}}{n} = 0$. □

We go back to the proposition.

Proof: We only consider the edge shift, the vertex shift is similar.

By proposition 4.1, we only need to show that $\lim_{n \to \infty} \frac{1}{n} \log |W_{n(\Sigma_A)}| = \log \lambda$.

It is not hard to prove that the number of allowed words of length n beginning with the symbol i and ending with j is the i, j-th entry of A^n, hence $|W_{n(\Sigma_A)}| = S_n$. By the lemma above, we complete the proof. □

For a discrete dynamical system f, we define $\mathrm{Fix}(f)$ to be the set of fixed points of f, *and we give a definition of the zeta function here.*

Definition 4.1: The zeta function $\zeta_{f(z)}$ of f is defined as:

$$\zeta_{f(z)} = \exp\left(\Sigma_{n=1}^{\infty} |\mathrm{Fix}(f^n)| \frac{z^n}{n}\right)$$

To be specific, this zeta function is called the *Artin Mazur zeta function.* The zeta function of the edge shift determined by an adjacency matrix A is denoted by ζ_A. And we can show that it is actually a rational function, and so are SFTs.

Proposition 4.3: $\zeta_{A(z)} = (\det(I - zA))^{-1}$

Proof: Since $\log\left(\frac{1}{1-x}\right) = \Sigma_{n=1}^{\infty} \frac{x^n}{n}$, we have $\exp\left(\Sigma_{n=1}^{\infty} \frac{x^n}{n}\right) = \frac{1}{1-x}$. Moreover, since the number of allowed words of length n beginning with the symbol i and ending with j is the i, j-th entry of A^n, we have

$$|\mathrm{Fix}(\sigma^n \mid \Sigma_A)| = \mathrm{tr}(A^n) = \Sigma_\lambda \lambda^n.$$

Suppose that A is a $N \times N$ matrix,

$$\zeta_{A(z)} = \exp\left(\Sigma_{n=1}^{\infty} \frac{\Sigma_{\lambda(\lambda z)}^{n}}{n}\right) = \Pi_\lambda \exp\left(\Sigma_{n=1}^{\infty} \frac{(\lambda z)^n}{n}\right) = \Pi_\lambda (1 - \lambda z)^{-1}$$

$$= \frac{1}{z^N} \Pi_{\lambda\left(\frac{1}{z} - \lambda\right)}^{-1} = \left(z^n \det\left(\frac{1}{z} I - A\right)\right)^{-1} = (\det(I - zA))^{-1}.$$

□

We have a generalized version of this proposition, which is too complicated to be proved here.

Theorem 4.1: The zeta function of a subshift $X \in \Sigma_m$ is rational if and only if there are matrices A and B such that $|\mathrm{Fix}(\sigma^n\ |_X)| = \mathrm{tr} A^n - \mathrm{tr} B^n$for all $n \in \mathbb{N}_0$.

Moreover, we have the theorem which says that the Artin–Mazur zeta function of an interval map f is the inverse of the kneading determinant of f, which is relevant to Milnor-Thurston kneading theory[2].

Bibliography

[1] M. Brin and G. Stuck, *Introduction to Dynamical Systems*. Cambridge University Press, 2002.

[2] J. Milnor and W. Thurston, "On iterated maps of the interval," in *Dynamical Systems*, J. C. Alexander, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 1988, pp. 465–563.