3. Shifts of finite type

§3.1   Shifts of finite type

Let $S = \{1, 2, \ldots, k\}$ be a finite set of symbols. We will be interested in sets consisting of sequences of these symbols, subject to certain conditions. We will impose the following conditions: we assume that for each symbol i we allow certain symbols (depending only on i) to follow i and disallow all other symbols.

This information is best recorded in a $k \times k$ matrix A with entries in $\{0, 1\}$. That is, we allow the symbol j to follow the symbol i if and only if the corresponding (i, j)th entry of the matrix A (denoted by $A_{i,j}$) is equal to 1.

Definition.   Let A be a $k \times k$ matrix with entries in $\{0, 1\}$. Let

$$\Sigma_A^+ = \{(x_j)_{j=0}^\infty \mid A_{x_i, x_{i+1}} = 1, \text{ for } j \in \mathbb{N}\}$$

denote the set of all infinite sequences of symbols (x_j) where symbol j can follow symbol i precisely when $A_{i,j} = 1$. We call Σ_A^+ a *(one-sided) shift of finite type*.

Let

$$\Sigma_A = \{(x_j)_{j=-\infty}^\infty \mid A_{x_j, x_{j+1}} = 1, \text{ for } j \in \mathbb{Z}\}$$

denote the set of all bi-infinite sequences of symbols subject to the same conditions. We call Σ_A a *(two-sided) shift of finite type*.

Sometimes for brevity we refer to Σ_A^+ or Σ_A as a *shift space*.

An alternative description of Σ_A^+ and Σ_A can be given as follows. Consider a graph with vertex set $\{1, 2, \ldots, k\}$ and with a directed edge from vertex i to vertex j precisely when $A_{i,j} = 1$. Then Σ_A^+ and Σ_A correspond to the set of all infinite (respectively, bi-infinite) paths in this graph.

Define

$$\sigma^+ : \Sigma_A^+ \to \Sigma_A^+$$

by

$$(\sigma^+(x))_j = x_{j+1}.$$

Then σ^+ takes a sequence in Σ_A^+ and shifts it one place to the left (deleting the first term), We call σ^+ the *(one-sided, left) shift map*.

There is a corresponding shift map on the two-sided shift space. Define

$$\sigma : \Sigma_A \to \Sigma_A$$

by
$$(\sigma(x))_j = x_{j+1},$$

so that σ shifts sequences one place to the left. Notice that in this case, we do not need to delete any terms in the sequence, We call σ the *(two-sided, left) shift map.*

Notice that σ is invertible but σ^+ is not.

Examples.

1. Take A to be the $k \times k$ matrix with each entry equal to 1, Then any symbol can follow any other symbol. Hence Σ_A^+ is the space of all sequences of symbols $\{1, 2, \ldots, k\}$. In this case we write Σ_k^+ for Σ_A^+ and refer to it as the *full one-sided k-shift.* Similarly, we can define the full two-sided k-shift,

2. Take A to be the matrix
$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

   Then Σ_A^+ consists of all sequences of 1s and 2s subject to the condition that each 2 must be followed by a 1.

The following two exercises show that, for certain A, Σ_A^+ (or Σ_A) can be rather uninteresting.

Exercise 3.1
Let
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Show that Σ_A^+ is empty.

Exercise 3.2
Let
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Calculate Σ_A^+.

The following conditions on A guarantee that Σ_A^+ (or Σ_A) is more interesting than the examples in exercises 3.1 and 3.2.

Definition. Let A be a $k \times k$ matrix with entries in $\{0, 1\}$. We say that A is *irreducible* if for each $i, j \in \{1, 2, \ldots, k\}$ there exists $n = n(i, j) > 0$ such that $(A^n)_{i,j} > 0$. (Here, $(A^n)_{i,j}$ denotes the (i, j)th entry of the nth power of A.)

Definition.   Let A be a $k \times k$ matrix with entries in $\{0, 1\}$. We say that A is *aperiodic* if there exists $n > 0$ such that for all $i, j \in \{1, 2, \ldots, k\}$ we have $(A^n)_{i,j} > 0$.

In graph-theoretic terms, the matrix A is irreducible if there exists a path along edges from any vertex to any other vertex. The matrix A is aperiodic if this path can be chosen to have the same length (i.e. consist of the same number of edges), irrespective of the two vertices chosen.

Exercise 3.3

(i) *Consider the matrix*
$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

   Draw the corresponding graph. Is this matrix irreducible? Is it aperiodic?

(ii) *Consider the matrix*
$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

   Draw the corresponding graph. Is this matrix irreducible? Is it aperiodic?

Remark.   These shift spaces may seem very strange at first sight—it takes a long time to get used to them. However (as we shall see) they are particularly tractable examples of chaotic dynamical systems. Moreover, a wide class of dynamical systems (notably hyperbolic dynamical systems) can be modeled in terms of shifts of finite type. We have already seen a particularly simple example of this: the doubling map can be modeled by the full one-sided 2-shift.

§3.2   Periodic points

A sequence $x = (x_j)_{j=0}^{\infty} \in \Sigma_A^+$ is periodic for the shift σ^+ if there exists $n > 0$ such that $\sigma^n x = x$.

   One can easily check that this means that

$$x_j = x_{j+n} \text{ for all } j \in \mathbb{N}.$$

That is, the sequence x is determined by a finite block of symbols $x_0, \ldots, x_{n-1}$ and

$$x = (x_0, x_1, \ldots, x_{n-1}, x_0, x_1, \ldots, x_{n-1}, \ldots,).$$

Exercise 3.4

Consider the full one-sided k-shift. How many periodic points of period n are there?

§3.3   Cylinders

In future lectures we will need a particularly tractable class of subsets of shift spaces. These are the *cylinder sets* and are formed by fixing a finite set of co-ordinates. More precisely, in Σ_A we define

$$[y_{-m}, \ldots, y_{-1}, y_0, y_1, \ldots, y_n]_{-m,n} = \{x \in \Sigma_A \mid x_j = y_j, \ -m \leq j \leq n\},$$

and in Σ_A^+ we define

$$[y_0, y_1, \ldots, y_n]_{0,n} = \{x \in \Sigma_A^+ \mid x_j = y_j, \ 0 \leq j \leq n\}.$$

§3.4   A metric on Σ_A^+

What does it mean for two sequences in Σ_A^+ to be 'close'? Heuristically we will say that two sequences $(x_j)_{j=0}^\infty$ and $(y_j)_{j=0}^\infty$ are close if they agree for a large number of initial places.

More formally, for two sequences $x = (x_j)_{j=0}^\infty, y = (y_j)_{j=0}^\infty \in \Sigma_A^+$ we define $n(x,y)$ by setting $n(x,y) = n$ if $x_j = y_j$ for $j = 0, \ldots, n-1$ but $x_n \neq y_n$. Thus $n(x,y)$ is the first place in which the sequences x and y disagree. (We set $n(x,y) = \infty$ if $x = y$.)

We define a metric d on Σ_A^+ by

$$d((x_j)_{j=0}^\infty, (y_j)_{j=0}^\infty) = \left(\frac{1}{2}\right)^{n(x,y)} \quad \text{if } x \neq y$$

and $d((x_j)_{j=0}^\infty, (y_j)_{j=0}^\infty) = 0$ if $x = y$.

Exercise 3.5
Show that this is a metric.

In the two-sided case, we can define a metric in a similar way. Let $x = (x_j)_{j=-\infty}^\infty, y = (y_j)_{j=-\infty}^\infty \in \Sigma_A$. Define $n(x,y)$ by setting $n(x,y) = n$ if $x_j = y_j$ for $|j| \leq n-1$ and either $x_n \neq y_n$ or $x_{-n} \neq y_{-n}$. Thus $n(x,y)$ is the first place, going either forwards or backwards, in which the sequences x, y disagree. (We again set $n(x,y) = \infty$ if $x = y$.)

We define a metric d on Σ_A in the same way:

$$d((x_j)_{j=-\infty}^\infty, (y_j)_{j=-\infty}^\infty) = \left(\frac{1}{2}\right)^{n(x,y)} \quad \text{if } x \neq y$$

and $d((x_j)_{j=-\infty}^\infty, (y_j)_{j=-\infty}^\infty) = 0$ if $x = y$.

Theorem 3.1
Let Σ_A^+ be a shift of finite type.

(i) Σ_A^+ is a compact metric space.

(ii) The shift map σ^+ is continuous.

Remark.   The corresponding statements for the two-sided case also hold.

Proof.   (i) If $\Sigma_A^+ = \emptyset$ or if Σ_A^+ finite then trivially it is compact. Thus we may assume that Σ_A^+ is infinite.

Let $x^{(m)} \in \Sigma_A^+$ be a sequence (in reality, a sequence of sequences!). We need to show that $x^{(m)}$ has a convergent subsequence. Since $\Sigma_A^+ = \bigcup_{i=1}^k [i]$ at least one cylinder $[i]$ contains infinitely many elements of the sequence $x^{(m)}$; call this $[y_0]$. Thus there are infinitely m for which $x^{(m)} \in [y_0]$.

Since $[y_0] = \bigcup_{A_{y_0}, i=1} [y_0 i]$ we similarly obtain a cylinder of length 2, $[y_0 y_1]$ say, containing infinitely many elements of the sequence $x^{(m)}$.

Continue inductively in this way to obtain a nested family of cylinders $[y_0, \ldots, y_n]$, $n \geq 0$, each containing infinitely many elements of the sequence $x^{(m)}$.

Set $y = (y_n)_{n=0}^\infty \in X_A^+$. Then for each $n \geq 0$, there exist infinitely many m for which $d(y, x^{(m)}) \leq (1/2)^m$. Thus y is the limit of some subsequence of $x^{(m)}$.

(ii) We want to show the following: $\forall \varepsilon > 0 \; \exists \delta > 0$ s.t. $d(x, y) < \delta \Rightarrow d(\sigma(x), \sigma(y)) < \varepsilon$.

Let $\varepsilon > 0$. Choose n such that $1/2^n < \varepsilon$. Let $\delta = 1/2^{n+1}$. Suppose that $d(x, y) < \delta$. Then $n(x, y) > n+1$, so that x and y agree in the first $n + 1$ places. Hence $\sigma(x)$ and $\sigma(y)$ agree in the first n places, so that $n(\sigma(x), \sigma(y)) > n$. Hence $d(\sigma(x), \sigma(y)) = 1/2^{n(\sigma(x).\sigma(y))} < 1/2^n < \varepsilon$.

$\square$

Exercise 3.6
Let A be an irreducible $k \times k$ matrix with entries in $\{0, 1\}$. Show that the set of all periodic points for σ^+ is dense in Σ_A^+. (Recall that a subset Y of a set X is said to be dense if: for all $x \in X$ and for all $\varepsilon > 0$ there exists $y \in Y$ such that $d(x, y) < \varepsilon$, i.e. any point of X can be arbitrarily well approximated by a point of Y.)

Exercise 3.7
Let A be an irreducible $k \times k$ matrix with entries in $\{0, 1\}$. Show that there exists a point $x \in \Sigma_A^+$ with a dense orbit. (Hint: first show that if the orbit of a point visits each cylinder then it is dense. To construct such a point, mimic the argument used for the doubling map in Lecture 2. Use irreducibility to show that one can concatenate cylinders together by inserting finite strings of symbols between them.)