15. Examples of measure-preserving transformations: shifts of finite type

§15.1 Shifts of finite type

Let A be a $k \times k$ matrix with entries equal to 0 or 1. Recall that we have defined the (two-sided) shift of finite type by

$$\Sigma_A = \{x = (x_n) \in \{1, \ldots, k\}^{\mathbb{Z}} \mid A(x_n, x_{n+1}) = 1 \ \forall n \in \mathbb{Z}\}$$

and the (one-sided) shift of finite type

$$\Sigma_A^+ = \{x = (x_n) \in \{1, \ldots, k\}^{\mathbb{Z}^+} \mid A(x_n, x_{n+1}) = 1 \ \forall n \in \mathbb{Z}^+\}.$$

The shift maps $\sigma : \Sigma_A \to \Sigma_A$, $\sigma : \Sigma_A^+ \to \Sigma_A^+$ are defined by

$$\sigma(\ldots, x_1, \underbrace{x_0}_{\text{0th place}}, x_1, x_2, \ldots) = (\ldots, x_0, \underbrace{x_1}_{\text{0th place}}, x_2, x_3, \ldots),$$

$$\sigma(x_0, x_1, x_2, x_3, \ldots) = (x_1, x_2, x_3, \ldots),$$

respectively, i.e., σ shifts sequences one place to the left.

As an analogue of intervals in this case, we have so-called ‘cylinder sets’, formed by fixing a finite set of co-ordinates. More precisely, in Σ_A we define

$$[y_{-m}, \ldots, y_{-1}, y_0, y_1, \ldots, y_n] = \{x \in \Sigma_A \mid x_i = y_i, \ -m \leq i \leq n\},$$

and in Σ_A^+ we define

$$[y_0, y_1, \ldots, y_n] = \{x \in \Sigma_A^+ \mid x_i = y_i, \ 0 \leq i \leq n\}.$$

In each case the cylinder sets form a semi-algebra which generates the Borel σ-algebra. (Cylinder sets are both open and closed.)

§15.2 The Perron-Frobenius theorem

The following standard result will be useful.

Theorem 15.1 (Perron-Frobenius)
Let B be a non-negative aperiodic $k \times k$ matrix (i.e. $B_{i,j} \geq 0$ for each $1 \leq i, j \leq k$ and there exists $n > 0$ such that $B_{i,j}^n > 0$ for all $0 \leq i, j \leq k$). Then:

(i) there exists a positive eigenvalue $\lambda > 0$ such that all other eigenvalues $\lambda_i \in \mathbb{C}$ satisfy $|\lambda_i| < \lambda$,

(ii) *the eigenvalue* λ *is simple (i.e. the corresponding eigenspace is one-dimensional),*

(iii) *there is a unique right-eigenvector* $v = (v_1, \ldots, v_k)^T$ *such that* $v_j > 0$, $\sum_{j=1}^n |v_j| = 1$ *and*

$$Bv = \lambda v,$$

(iv) *there is a unique positive left-eigenvector* $u = (u_1, \ldots, u_k)$ *such that* $u_j > 0$, $\sum_{j=1}^n |u_j| = 1$ *and*

$$uB = \lambda u,$$

(v) *eigenvectors corresponding to eigenvalues other than* λ *are not positive: i.e. at least one co-ordinate is positive and at least one co-ordinate is negative.*

§15.3 Markov measures

We will now see how to construct a large class of σ-invariant measures on shifts of finite type.

Definition. A $k \times k$ matrix P is said to be *stochastic* if:

(i) $P(i,j) \geq 0 \quad i,j = 1, \ldots, k$,

(ii) $\sum_{j=1}^k P(i,j) = 1, \quad i = 1, \ldots, k$.

Suppose that P is compatible with A, i.e.,

$$P(i,j) > 0 \iff A(i,j) = 1.$$

Suppose in addition that P, or equivalently A, is *aperiodic*, i.e., there exists n such that for each i,j we have $P^n(i,j) > 0$.

Applying the Perron-Frobenius theorem, we see that there exists a unique maximal eigenvalue λ. As P is stochastic, we must have that $\lambda = 1$ (why?). Moreover, by (ii) in the above definition, the right-eigenvector is $(1, \ldots, 1)$. Let $p = (p_1, \ldots, p_k)$ be the corresponding (strictly positive) left eigenvector, normalised so that $\sum_{i=1}^k p_i = 1$.

Now we define a probability measure $\mu = \mu_P$ on Σ_A, Σ_A^+ by

$$\mu_P[y_\ell, y_{\ell+1}, \ldots, y_n] = p_{y_\ell} P(y_\ell, y_{\ell+1}) \cdots P(y_{n-1}, y_n),$$

on cylinder sets. (By the Kolmogorov Extension Theorem, this uniquely defines a measure on the whole Borel σ-algebra.)

We shall show that the measure μ_P on Σ_A^+ is σ-invariant. By the Kolmogorov Extension Theorem, it is enough to show that μ_P and $\sigma_*\mu_P$ agree on cylinder sets. Now

$$\begin{aligned}
\sigma_*\mu_P[y_0, y_1, \ldots, y_n] &= \mu_P(\sigma^{-1}[y_0, y_1, \ldots, y_n]) \\
&= \mu_P\left(\bigcup_{j=1}^{k} [j, y_0, y_1, \ldots, y_n]\right) \\
&= \sum_{j=1}^{k} \mu_P[j, y_0, y_1, \ldots, y_n] \\
&= \sum_{j=1}^{k} p_j P(j, y_0) P(y_0, y_1) \cdots P(y_{n-1}, y_n) \\
&= p_{y_0} P(y_0, y_1) \cdots P(y_{n-1}, y_n) \\
&= \mu_P[y_0, y_1, \ldots, y_n],
\end{aligned}$$

as required (where to get the penultimate line we have used the fact that $pP = p$). Probability measures of this form are called *Markov measures*.

Given an aperiodic $k \times k$ matrix A there are of course many compatible stochastic matrices P. Each such stochastic matrix generates a different Markov measure. However, there are several naturally defined measures that turn out to be Markov, and we give two of them here.

§15.4 Full shifts

Recall that if $A(i, j) = 1$ for all i, j then

$$\Sigma_A = \{1, \ldots, k\}^{\mathbb{Z}}, \quad \Sigma_A^+ = \{1, \ldots, k\}^{\mathbb{Z}^+}$$

are the *full shifts* on k symbols. In this case we may define a (family of) measures by taking $p = (p_1, \ldots, p_k)$ to be any (positive) probability vector (i.e., $p_i > 0$, $\sum_{i=1}^{k} p_i$) and defining

$$\mu[y_l, \ldots, y_n] = p_{y_l} \cdots p_{y_n}.$$

Such a μ is called a *Bernoulli measure*.

Exercise 15.1
Show that Bernoulli measures are Markov measures (i.e. construct a matrix P for which $(p_1, \ldots, p_k)P = (p_1, \ldots, p_k)$.

§15.5 The Parry measure

As A is a non-negative aperiodic matrix, by the Perron-Frobenius theorem there exists a unique maximal eigenvalue λ with corresponding left and right eigenvalues $u = (u_1, \ldots, u_k)$ and $v = (v_1, \ldots, v_k)$, respectively. Define

$$\begin{aligned} P_{i,j} &= \frac{A_{i,j} v_j}{\lambda v_i} \\ p_i &= \frac{u_i v_i}{c} \end{aligned}$$

where $c = \sum_{i=1}^{k} u_i v_i$.

Exercise 15.2
Show that P is a stochastic matrix and that p is a normalised left-eigenvalue for P.

The corresponding Markov measure is called the *Parry measure*.