Master 2 Mathematics and Computer Science Symbolic Dynamics. Lecture 2

MARIE-PIERRE BÉAL

University Gustave Eiffel
Laboratoire d'informatique Gaspard-Monge UMR 8049

Université Gustave Eiffel

- Conjugacy and state-splitting
- Perron-Frobenius theorem
- Conjugacy invariants.

Conjugacy and state-splitting

Block maps

Let X be a shift space on the alphabet A, and let B be another alphabet. Given integers m, a with $m + a \geq 0$, a *block map* is a map $f : \mathcal{B}_{m+a+1}(X) \to B$. The *sliding block code* defined by f and (m, a) is the map $\varphi : X \to B^{\mathbb{Z}}$ defined by $\varphi(x) = y$ if for every $i \in \mathbb{Z}$,

$$y_i = f(x_{[i-m,i+a]}).$$

Thus y is computed from x by sliding a window of length $m + a + 1$ on x. The integer m is the *memory* of φ and a is its *anticipation*.
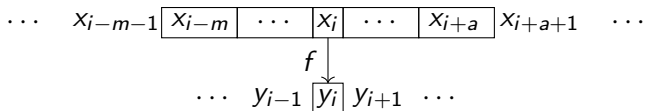
$$\cdots \quad x_{i-m-1} \boxed{\begin{array}{|c|c|c|c|c|} x_{i-m} & \cdots & x_i & \cdots & x_{i+a} \end{array}} x_{i+a+1} \quad \cdots$$

$$f \downarrow$$

$$\cdots \quad y_{i-1} \boxed{y_i} \; y_{i+1} \quad \cdots$$

Figure: The sliding block code.

Sliding block code

Proposition

Let X be a shift space on the alphabet A and let $\varphi \colon X \to B^{\mathbb{Z}}$ be a sliding block code. Then $\varphi(X)$ is a shift space.

Proof.

The map φ is clearly continuous and commutes with the shift, that is, $\varphi \circ S = S \circ \varphi$. Thus $Y = \varphi(X)$ is closed (because the continuous image of a compact space is compact and thus closed). It is shift-invariant because, if $y = \varphi(x)$ with $x \in X$, then $S(y) = S \circ \varphi(x) = \varphi(S(x))$ and thus $S(y) \in Y$. $\qquad\square$

Curtis-Hedlund-Lyndon theorem

Let X, Y be shift spaces. A map $\varphi \colon X \to Y$ is a *morphism* if φ is continuous and commutes with the shift map.

Theorem (Curtis, Hedlund, Lyndon)

Let X, Y be shift spaces. A map $\varphi \colon X \to Y$ is a morphism systems if and only if it is a sliding block code from X into Y.

Proof.

A sliding block code is clearly continuous and commutes with the shift.

Conversely, let $\varphi \colon X \to Y$ be a morphism . For every letter b from the alphabet B of Y, the set $[b]_Y$ is clopen and thus $\varphi^{-1}([b]_Y)$ is also clopen. Since a clopen set is a finite union of cylinders, there is an integer n such that $\varphi(x)_0$ depends only on $x_{[-n,n]}$. Set $f(x_{[-n,n]}) = \varphi(x)_0$. Then φ is the sliding block code associated with the block map f. $\qquad\square$

Let X, Y be shift spaces. A map $\varphi\colon X \to Y$ is a *conjugacy* if φ is a bijective morphism.

Its inverse is also a morphism.

It follows from the Curtis-Hedlund-Lyndon theorem that a conjugacy between shift spaces X and Y is the same as a sliding block code from X to Y which is invertible (the inverse is also a sliding block code since the inverse of a conjugacy is a morphism).

We say that $f\colon A^{m+a+1} \to B$ is a *k-block map*, where $k = m + a + 1$, and that the corresponding sliding block code is a *k-block code*.
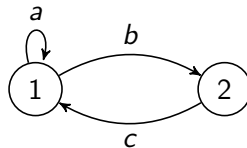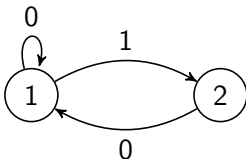
The simplest form of a sliding block code, called a 1-*block code*, occurs when $m = a = 0$. In this case, we use the same symbol to denote the 1-block map $\phi\colon A \to B$ and the 1-block code $\phi\colon A^{\mathbb{Z}} \to B^{\mathbb{Z}}$. In this way, we have for every $x \in X$ and $n \in \mathbb{Z}$, $\phi(x)_n = \phi(x_n)$.

Example

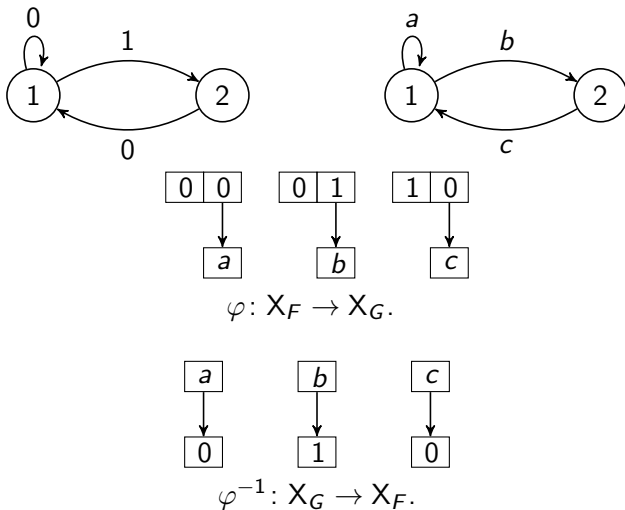X_F with $F = \{11\}$ is a shift of finite (the golden mean shift).
X_G with $G = \{ac, ba, bb, cc\}$ is a shift of finite.

$\varphi \colon X_F \to X_G.$

$\varphi^{-1} \colon X_G \to X_F.$

Edge shifts

An *edge shift* is the set of bi-infinite paths of a directed (multi)graph.

Proposition

Every shift of finite type is conjugate to an edge shift.

Proof.

Let $X = X_F$ with F finite, and let n be the maximal size of words in F. We may assume that all words in F have size n.

Let $\mathcal{A} = (Q, E)$, where Q is the set of words of length $n - 1$ with edges $a_0 a_1 \ldots a_{n-1} \xrightarrow{a} a_1 \ldots a_{n-1} a$, where $a_0 a_1 \ldots a_{n-1} a \notin F$. We keep only the trim part of this automaton.

Then $\mathcal{A}$ is deterministic and local (all paths labeled by a word w of length n end in the same state q_w). $\square$

proof continued.

Let Y be the set of bi-infinite paths of $\mathcal{A}$.

Let $\phi\colon Y \to X$ defined by the 1-block map $f\colon E \to A$ with $f(e = (p, a, q)) = a$.

Then the sliding block code $\varphi\colon X \to Y$ with anticipation 0 and memory $n-1$ defined by the n-block map $g\colon A^n \to E$ with $g(a_0 a_1 \ldots a_{n-1}) = (p, a_{n-1}, q)$, where $p = a_0 a_1 \ldots a_{n-2}$ and $q = a_1 a_2 \ldots a_{n-1}$.

The map φ is the inverse of ϕ.

$\square$

Edge shifts

If we take a minimal deterministic presentation of an irreducible shift of finite type, then, up to a conjugacy, one may assume that all labels are distinct and the shift can be defined by the transition matrix of the graph G of the presentation.

Transition matrix of an automaton

$M = (m_{pq})_{p,q \in Q}$, where m_{pq} is the number of edges from p to q.
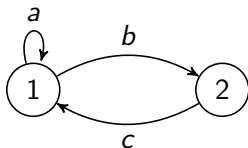
$$M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

If G is an automaton or a graph, the edge shift defined by G is denoted X_G.

State splitting of an automaton

An *out-splitting* of an automaton $\mathcal{A} = (Q, E)$ is a local transformation of $\mathcal{A}$ into an automaton $\mathcal{B} = (Q', E')$ obtained by selecting a state s and partitioning the set of edges going out of s into two non-empty sets E_1 and E_2.

- $Q' = Q \setminus \{s\} \cup \{s_1, s_2\}$,
- E' contains all edges of E neither starting at or ending in s,
- E' contains the edge (s_1, a, t) for each edge $(s, a, t) \in E_1$, and the edge (s_2, a, t) for each edge $(s, a, t) \in E_2$, if $t \neq s$.
- E' contains the edges (t, a, s_1) and (t, a, s_2) if (t, a, s) in E, when $t \neq s$,
- E' contains the edges (s_1, a, s_1) and (s_1, a, s_2) if (s, a, s) in E_1, and the edges (s_2, a, s_1) and (s_2, a, s_2) if $(s, a, s) \in E_2$.

State splitting of an automaton

An *input state splitting* is defined similarly.

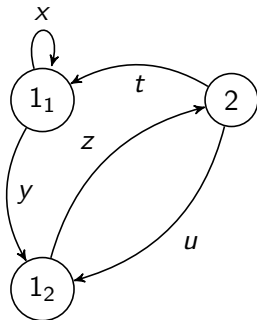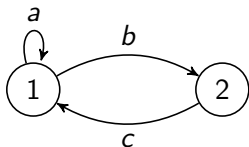The inverse operation is called an *output merging*, possible whenever s_1 and s_2 have the *same input edges*.
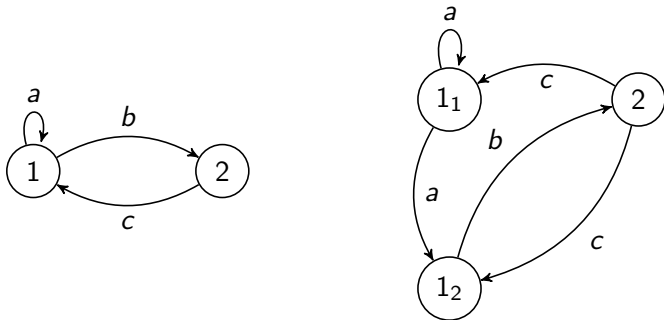
The state 1 is split into two states 1_1 and 1_2
with $E_1 = \{(1, a, 1)\}$ and $E_2 = \{(1, b, 2)\}$.

Output state splitting of a graph

State splitting

Proposition

Let G be a graph and H a split graph of G. Then X_G and X_H are conjugate.

Proof.

Let $G = (Q, E)$ (all labels are distinct).

Let $H = (Q', E')$ be an outsplit of G, obtained after splitting the state s into s_1, s_2 according to the partition E_1, E_2 of edges going out of s.

Let X_G be the edge shift defined by G and X_H be the edge shift defined by H.

Then X_G and X_H are conjugate. □

State splitting

Indeed, let $\varphi_{GH} \colon \; \colon E^{\mathbb{Z}} \to E'^{\mathbb{Z}}$ be the $(0, 1)$-sliding block code defined by the 2-block map $f \colon \mathcal{B}_2(X_G) \to E'$, where

$$
\begin{aligned}
f((t, a, u)(u, b, v)) &= (t, a, u) && \text{if } t, u \neq s, \\
f((t, a, s)(s, b, v)) &= (t, a, s_1) && \text{if } t \neq s \text{ and } (s, b, v) \in E_1, \\
f((t, a, s)(s, b, v)) &= (t, a, s_2) && \text{if } t \neq s \text{ and } (s, b, v) \in E_2, \\
f((s, a, t)(t, b, u)) &= (s_1, a, t) && \text{if } t \neq s \text{ and } (s, a, t) \in E_1, \\
f((s, a, t)(t, b, u)) &= (s_2, a, t) && \text{if } t \neq s \text{ and } (s, a, t) \in E_2, \\
f((s, a, s)(s, b, t)) &= (s_1, a, s_1) && \text{if } (s, b, t) \in E_1, \\
f((s, a, s)(s, b, t)) &= (s_1, a, s_2) && \text{if } (s, b, t) \in E_2.
\end{aligned}
$$

defines a conjugacy from X_G onto X_H. Its inverse is the sliding block code defined by the 1-block map $g \colon E' \to E$, where $g(t, a, u) = (\pi(t), \pi(a), \pi(u))$, with $\pi(t) = t$ if $t \neq s_1, s_2$, $\pi(s_1) = \pi(s_2) = s$, $\pi(a) = a$.

Complete out-splitting

Let $G = (Q, E)$ be a graph. All edges have distinct labels; we may omit them.

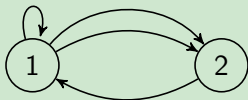The *complete out-splitting* of G is the graph $H = (Q', E')$ where
- $Q' = \{s_e \mid s \in Q, e = (s, t) \in E\}$,
- $E' = \{(s_e, t_f) \mid e = (s, t)\}$

Proposition

Let G be a graph and H the complete out-splitting of G. Then X_G and X_H are conjugate. There is a sequence of out-splittings from G to H.

Example

Compute the complete out-splitting of

Decomposition Theorem

An *trim graph* is a graph such that each state has at least one incoming edge and at least one outgoing edge.

> **Theorem (Decomposition Theorem, R. Williams 1973)**
>
> *Two edge shifts X_G and X_H defined by trim graphs G and H are conjugate if and only if there is a sequence of (input and output) state splittings and (input and output) state mergings from G to H.*

Higher block shift

Let X be a shift space on the alphabet A and let $k \geq 1$ be an integer.

The map $\gamma_k : X \to \mathcal{B}_k(X)^{\mathbb{Z}}$ defined for $x \in X$ by $y = \gamma_k(x)$ if, for every $n \in \mathbb{Z}$,

$$y_n = \langle x_n \cdots x_{n+k-1} \rangle, \tag{1}$$

is the k-th *higher block code*.
One also says that γ_k is a *coding by overlapping blocks* of length k.

The set $X^{(k)} = \gamma_k(X)$ is a shift space on $\mathcal{B}_k(X)$, called the k-th *higher block shift* of X.

Higher block code

The higher block code is an isomorphism of shift spaces, and the inverse of γ_k is the map $\pi_k \colon y \mapsto x$ such that, for all n, x_n is the first letter of the word u such that $y_n = \langle u \rangle$.
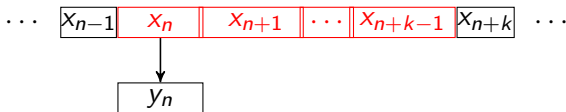

Figure: The k-th higher block code.

Thus, $X^{(k)}$ is conjugate to X.

Reduction to a 1-block code

Lemma

Let $\varphi \colon X \to Y$ be a sliding block code. Then there is a higher block shift $X^{(k)}$ of X, a conjugacy $\gamma \colon X \to X^{(k)}$ and a 1-block code $\phi \colon X^{(k)} \to Y$ such that $\phi \circ \gamma = \varphi$.

$$
\begin{array}{ccc}
X & \xrightarrow{\ \gamma\ } & X^{(k)} \\
 & {\scriptstyle\varphi}\searrow & \downarrow{\scriptstyle\phi} \\
 & & Y
\end{array}
$$

Proof.

Let $\varphi \colon X \to Y$ be a (m, a)-sliding block code.
Set $k = m + a + 1$ and $\gamma = \gamma_k \circ S^{-m}$.
Thus, $\gamma(x)_i = \langle x_{i-m} \cdots x_{i+a} \rangle$.
Put $\phi = \varphi \circ \gamma^{-1}$.
ϕ is a 1-block code.

Reduction to a 1-block code

Lemma

Let $\varphi\colon X_G \to X_H$ be a $(0, a)$-sliding block code with $a > 0$. Then there is an out-splitting G' of G such $X_{G'} = X_G^{(2)}$ and a conjugacy $\gamma\colon X_G \to X_G^{(2)}$ and a $(0, a-1)$-sliding block code $\varphi_2\colon X_G^{(2)} \to X_H$ such that $\varphi_2 \circ \gamma = \varphi$.

$$
\begin{array}{ccc}
X_G & \xrightarrow{\ \gamma\ } & X_G^{(2)} \\
& {\scriptstyle\varphi}\searrow & \downarrow {\scriptstyle\varphi_2} \\
& & X_H
\end{array}
$$

Reduction to a 1-block code

Proof.

Let $G' = (Q', E')$ be the complete out-splitting of G. Let $H = (R, F)$. Assume that φ is defined by the block map $f : E^{a+1} \to F$.

- $Q' = \{ s_e \mid s \in Q, e = (s, t) \in E \}$,
- $E' = \{ (s_e, t_f) \mid e = (s, t) \}$

We define $\varphi_2 : E' \to F$ by the $(0, a-1)$-block map $g : E'^a \to F$ with $g((s_{e_0}, s_{e_1})(s_{e_1}, s_{e_2}) \dots (s_{e_{a-1}}, s_{e_a})) = f(e_0 e_1 \dots e_a)$.

We get $\varphi_2 \circ \gamma = \varphi$.

We check that $X_{G'}$ is equal to $X_G^{(2)}$, up to a renaming of the states.

Let $\phi : E' \to \mathcal{B}_2(X_G)$ be the 1-block code defined by $\phi((s_e, t_f)) = (e, f)$. The inverse $\phi^{-1} : \mathcal{B}_2(X_G) \to E'$ is the 1-block code defined by $\phi^{-1}((e, f)) = (s_e, t_f)$, where $e = (s, t)$.

Reduction of the anticipation of ϕ^{-1}

Lemma

Let $\phi \colon X_G \to X_H$ be a 1-block conjugacy whose inverse has memory m' and anticipation $a' \geq 1$.
Then, there are out-splitting G' of G and H' of H and a 1-block conjugacy $\phi' \colon X_{G'} \to X_{H'}$ such that

$$\phi = \varphi_{H'H} \circ \phi' \circ \varphi_{GG'}.$$

$$X_G \xrightarrow{\ \varphi_{GG'}\ } X_{G'}$$
$$\phi \downarrow \qquad\qquad \downarrow \phi'$$
$$X_H \xleftarrow{\ \varphi_{H'H}\ } X_{H'}$$

and whose inverse has memory m' and anticipation $a' - 1$.

Reduction of the anticipation of ϕ^{-1}

Proof.

Let $\phi: X_G \to X_H$ be a 1-block conjugacy whose inverse ϕ^{-1} has memory m' and anticipation $a' \geq 1$.

Let $H' = (Q_{H'}, E_{H'})$ be the complete out-splitting of H. The edges of H' are (s_e, t_f) where $e = (s, t)$, $f = (t, u)$ are edges of H.

Let $G' = (Q_{G'}, E_{G'})$ be the out-splitting of G obtained by splitting each state v into states v_e, where e is an edge of H.
The edges going out of each state v of G are partitioned into the sets $\Delta(v)_e = \{(v, w) \text{ edge of } G \mid \phi((v, w)) = e\}$. The edges of G' are (v_e, w_f) where $\phi((v, w)) = e$.

We define $\phi': E_{G'} \to E_{H'}$ by $\phi'((v_e, w_f)) = (s_e, t_f)$, where $e = (s, t)$. Then ϕ' is a 1-block conjugacy.

$\square$

Reduction of the anticipation of ϕ^{-1}

Proof.

By hypothesis, ϕ^{-1} has memory m' and anticipation a'.
With $e_i = (s_i, s_{i+1})$, $f_i = (v_i, v_{i+1})$, we have

$$\ldots f_{-1} \cdot f_0 f_1 \ldots \xrightarrow{\varphi_{GG'}} \ldots ((v_{-1})_{e_{-1}}, (v_0)_{e_0}) \cdot ((v_0)_{e_0}, (v_1)_{e_1})((v_1)_{e_1}, (v_2)_{e_2}) \ldots$$

$$\left. \phi \right\downarrow \qquad\qquad\qquad\qquad\qquad\qquad\qquad \left\downarrow \phi' \right.$$

$$\ldots e_{-1} \cdot e_0 e_1 \ldots \xleftarrow{\varphi_{H'H}} \ldots ((s_{-1})_{e_{-1}}, (s_0)_{e_0}) \cdot ((s_0)_{e_0}, (s_1)_{e_1})((s_1)_{e_1}, (s_2)_{e_2}) \ldots$$

Hence ϕ'^{-1} has memory m' and anticipation $a' - 1$. $\qquad\square$

Proof of the decomposition theorem

Proof of the decomposition theorem.

Let us assume that there is a conjugacy $\varphi: X_G \to X_H$. We may assume that φ is a $(0, m+a)$-block code (with a composition with S^{-m}).

We reduce the anticipation of $\phi_i'^{-1}$ with out-splittings and the memory of $\phi_j'^{-1}$ with in-splittings, and get

$$
\begin{array}{ccccccccc}
X_G & \xrightarrow{\gamma} & (X_G)^{(m+a)} & \xrightarrow{\varphi_{GG_1'}} & X_{G_1'} & \xrightarrow{\varphi_{G_1'G_2'}} & X_{G_2'} & \longrightarrow \cdots \longrightarrow & X_{G_{m'+a'}'} \\
\downarrow{\varphi} & & \downarrow{\phi} & & \downarrow{\phi_1'} & & \downarrow{\phi_2'} & & \downarrow{\phi_{m'+a'}'} \\
X_H & \xleftarrow{\varphi_{H_1'H}} & X_{H_1'} & \xleftarrow{\varphi_{H_2'H_1'}} & X_{H_2'} & \longleftarrow \cdots \longleftarrow & X_{H_{m'+a'}'}
\end{array}
$$

Thus $\phi_{m'+a'}'$ is a 1-block conjugacy whose inverse is a 1-block conjugacy, implying that the graphs $G_{m'+a'}'$ and $H_{m'+a'}'$ are equal since they are trim.

Conversely, if there is a sequence of splittings and mergings from G to H, then X_G and X_H are conjugate. $\qquad\square$

Out-mergings and in-mergings do not commute

Let G be the graph with transition matrix

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

We can perform an out-merging of the states 2 and 3 of G (since the columns 2 and 3 are identical) and get G_1 with transition matrix

$$M_1 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}.$$

We can perform an in-merging of the states 2 and 3 of G (since the rows 2 and 3 are identical) and get G_2 with transition matrix

$$M_2 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}.$$

And no mergings are possible from G_1 or G_2.

Open problem

Strong shift equivalence

Two nonnegative integer matrices M, N are *elementary equivalent* if there are, possibly nonsquare, matrices R, S such that

$$M = RS, N = SR.$$

Two nonnegative integer matrices M, N are *strong shift equivalent* if there is a sequence of elementary equivalences from M to N:

$$M = R_0 S_0, S_0 R_0 = M_1,$$
$$M_1 = R_1 S_1, S_1 R_1 = M_2,$$
$$\vdots$$
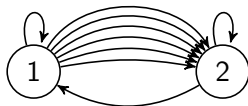$$M_\ell = R_\ell S_\ell, S_\ell R_\ell = N.$$

Theorem (Classification Theorem, R. Williams 1973)

Two edge shifts defined by matrices M and N are conjugate if and only if M and N are strong shift equivalent.

Let X and Y be the edge shifts defined by the graphs given by the matrices M and N respectively:

$$M = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 6 \\ 1 & 1 \end{bmatrix}$$

We may see these graphs as automata where all edge labels are distinct.

The shifts X and Y are conjugate (K. Baker, using computer research to prove strong shift equivalence).

It is not known whether X and Y defined by M_k and N_k are conjugate for $k \geq 4$.
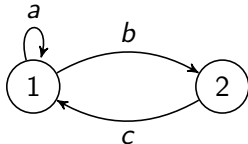
$$M_k = \begin{bmatrix} 1 & k \\ k-1 & 1 \end{bmatrix}, \quad N_k = \begin{bmatrix} 1 & k(k-1) \\ 1 & 1 \end{bmatrix}$$

Perron-Frobenius theorem

Let $G = (Q, E)$ be a graph. Its transition matrix is a nonnegative integer matrix M where

$M = (m_{pq})_{p,q \in Q}$, where m_{pq} is the number of edges from p to q.

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

Irreducible and primitive matrices

A nonnegative square matrix (with real coefficients) M is *irreducible* if for every pair s, t of indices, there is an integer $n \geq 1$ such that $M_{s,t}^n > 0$. Otherwise, M is *reducible*.

A matrix M is reducible if and only if, up to a permutation of the indices, it can be written

$$M = \begin{bmatrix} U & V \\ 0 & W \end{bmatrix}$$

for some matrices U, V, W with U, W being square matrices of dimension ≥ 1.

A nonnegative square matrix M is *primitive*, if there is some integer $n \geq 1$ such that all entries of M^n are positive.

The least such n is called the *exponent* of M, denoted $\exp(M)$.

A primitive matrix is irreducible but the converse is not necessarily true.

Irreducible matrix

Lemma

If M is a nonnegative $Q \times Q$ irreducible matrix, then $(I + M)^{n-1} > 0$, where $n = \operatorname{Card} Q$.

Proof.

Let G be the graph whose adjacency matrix is $I + M$.

Thus, $s \to t$ is an edge if and only if $(I + M)_{st} > 0$.

Since M is irreducible, there is a path of length at most $n - 1$ from s to t in G.

Since the state s has a self-loop, there is a path of length $n - 1$ from s to t in G.

Hence, $(I + M)_{st}^{n-1} > 0$ for all states $s, t \in Q$. $\qquad\square$

Periods

The *period* of an irreducible nonnegative square matrix $M \neq 0$ is the greatest common divisor of the integers n such that M^n has a positive diagonal coefficient. By convention, the period of $M = 0$ is 1. If M has period p, then M and M^p have, up to a permutation of indices, the forms:

$$
M = \begin{bmatrix} 0 & M_1 & 0 & \dots & 0 \\ 0 & 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & M_{p-1} \\ M_p & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad
M^p = \begin{bmatrix} D_1 & 0 & 0 & \dots & 0 \\ 0 & D_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & D_{p-1} & 0 \\ 0 & 0 & \dots & 0 & D_p \end{bmatrix}
$$

Thus M^p is block diagonal, with each diagonal block D_i primitive. An irreducible matrix is primitive if and only if it has period 1.

Periods

Example

The three matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

are nonnegative. The first one is reducible. The second one is irreducible but not primitive since it has period 2. The third one is primitive with exponent 2.

Theorem

Let M be a nonnegative real $Q \times Q$-matrix. Then

1. M has an eigenvalue λ_M such that $|\mu| \leq \lambda_M$ for every eigenvalue μ of M.

2. There corresponds to λ_M a nonnegative eigenvector v, and a positive one if M is irreducible. If M is irreducible, λ_M is the only eigenvalue with a nonnegative eigenvector.

3. If M is primitive, the sequence (M^n/λ_M^n) converges to the matrix yx where x, y are positive left and right eigenvectors relative to λ_M with $\sum_{s \in Q} y_s = 1$ and $\sum_{s \in Q} x_s y_s = 1$.

If M is irreducible, then λ_M is simple. The matrix M is primitive if and only if $|\mu| < \lambda_M$ for every other eigenvalue μ of M.

An *eigenvector* of a square real matrix M for the eigenvalue λ (a real or complex number) is a **non null** vector v (with real or complex coefficients) such that $Mv = \lambda v$.

The *spectral radius* of a square real matrix is the real number

$$\rho(M) = \max\{|\lambda| \mid \lambda \text{ eigenvalue of } M\}.$$

The theorem states in particular that if a matrix M is irreducible, $\rho(M)$ is an eigenvalue of M that is algebraically simple. Furthermore, if M is primitive, any eigenvalue of M other than $\rho(M)$ has modulus less than $\rho(M)$.

Proposition

Any nonnegative matrix M has a real eigenvalue λ_M such that $|\lambda| \leq \lambda_M$ for any eigenvalue λ of M, and there corresponds to λ_M a nonnegative eigenvector v.

If M is irreducible, there corresponds to λ_M a positive eigenvector v, and λ_M is the only eigenvalue with a nonnegative eigenvector.

Proof.

We first assume that M **is irreducible**.

For any nonnegative real vector $v \neq 0$, let

$$r_M(v) = \min\{(Mv)_s/v_s \mid v_s \neq 0\}.$$

Thus $r_M(v)$ is the largest real number r such that $Mv \geq rv$.

One has $r_M(\lambda v) = r_M(v)$ for any nonnegative nonzero real number λ.

Moreover, the mapping $v \rightarrow r_M(v)$ is continuous on the set of nonnegative nonzero vectors.

The set X of nonnegative vectors v such that $\|v\| = 1$ is compact.

Define $\lambda_M = \max\{r_M(v) \mid v \in X\}$.

Since a continuous function on a compact set reaches its maximum on this set, there is an $x \in X$ such that $r_M(x) = \lambda_M$.

Since $r_M(v) = r_M(\lambda v)$ for $\lambda \geq 0, \lambda \neq 0$, we have

$\lambda_M = \max\{r_M(v) \mid v \geq 0, v \neq 0\}$. $\square$

Proof.

We show that $Mx = \lambda_M x$.

By the definition of the function r_M, we have $Mx \geq \lambda_M x$.

Set $y = Mx - \lambda_M x$. Then $y \geq 0$.

Assume $Mx \neq \lambda_M x$. Then $y \neq 0$.

Since $(I + M)^n > 0$ for some $n \geq 1$, this implies that $(I + M)^n y > 0$.

But $(I + M)^n y = (I + M)^n (Mx - \lambda_M x) = M(I + M)^n x - \lambda_M (I + M)^n x = Mz - \lambda_M z$ with $z = (I + M)^n x$.

This shows that $Mz > \lambda_M z$, which implies that $r_M(z) > \lambda_M$, a contradiction.

Thus, λ_M is an eigenvalue with a nonnegative eigenvector x.

Since $(I + M)^n x = (1 + \lambda_M)^n x$ is positive, we get $x > 0$. $\qquad\square$

Proof.

Let us now show that $\lambda_M \geq |\lambda|$ for each real or complex eigenvalue λ of M.

Indeed, let v be an eigenvector corresponding to λ.

Then $Mv = \lambda v$.

Let $|v|$ be the nonnegative vector with coordinates $|v_s|$. Then $M|v| \geq |\lambda||v|$ by the triangular inequality.

By the definition of the function r_M, this implies $r_M(|v|) \geq |\lambda|$ and consequently $\lambda_M \geq |\lambda|$.

$\square$

Proof of Perron-Frobenius (If M is irreducible, then λ_M is simple)

Proof.

We show that if M is irreducible, λ_M is a simple eigenvalue.

Let v be an eigenvector of M for λ_M (with complex coefficients).
We have seen that $w = |v|$ is a nonnegative eigenvector of M for λ_M.
We have $w > 0$. Indeed, since $(I + M)^n > 0$ and $w \neq 0$, we have $(I + M)^n w > 0$, And $(I + M)^n w = (1 + \lambda_M)^n w$, implying $w > 0$.
Hence, v has no null coefficient and $|v|$ is a positive eigenvector of M for λ_M.

Now let w, w' be two eigenvectors of M for λ_M.
Let $z = w'_s w - w_s w'$. Then $Mz = \lambda_M z$ and $z_s = 0$, implying $z = 0$ and w is colinear to w'.

$\square$

Proof.

We show that if M is irreducible, λ_M is the only eigenvalue with a nonnegative eigenvector.

Let $Mv = \lambda v$ with $v \geq 0$, $v \neq 0$. Hence λ is a nonnegative real number.
Then $(I + M)^n v = (1 + \lambda)^n v > 0$ implying $v > 0$.

Let D be the diagonal matrix with coefficients v_s and $N = D^{-1}MD$.
We have $n_{st} = m_{st}v_t/v_s$ and $\sum_t n_{st} = \sum_t m_{st}v_t/v_s = \lambda$.
If w is a positive eigenvector of M for λ_M, then $D^{-1}w$ be a positive eigenvector of N for λ_M.
Let w be a positive eigenvector of N for λ_M, normalized in such a way that $w_s \leq 1$ for all s and $w_{s_0} = 1$ for some s_0.
Then $\lambda_M = \lambda_M w_{s_0} = \sum_t n_{s_0 t} w_t \leq \sum_t n_{s_0 t} = \lambda$.
Thus $\lambda = \lambda_M$.

Proof of Perron-Frobenius Points 1 and 2

Proof.

If M **is reducible**, we may consider a triangular decomposition:

$$M = \begin{bmatrix} U & V \\ 0 & W \end{bmatrix}.$$

Applying by induction the theorem to U and W, we obtain nonnegative eigenvectors u and w for the eigenvalues λ_U and λ_W of U and W. We prove that $\max(\lambda_U, \lambda_W)$ is an eigenvalue of M with some nonnegative eigenvector.

If $\lambda_U \geq \lambda_W$, then λ_U is an eigenvalue of M with nonnegative eigenvector $\begin{bmatrix} u \\ 0 \end{bmatrix}$.

$\square$

Proof of Perron-Frobenius Points 1 and 2

Proof.

If $\lambda_U < \lambda_W$, then λ_W is an eigenvalue of M with nonnegative eigenvector $\begin{bmatrix} u' \\ w \end{bmatrix}$, where $u' = \sum_{n \geq 0} (U/\lambda_W)^n \lambda_W^{-1} V w$.

Indeed, the spectral radius of U/λ_W is < 1, implying the convergence of $\sum_{n \geq 0} (U/\lambda_W)^n$ (Berstel Lemma).

Conversely, if λ is an eigenvalue of M with corresponding eigenvector $\begin{bmatrix} u \\ w \end{bmatrix}$, then λ is an eigenvalue of W if $w \neq 0$, and is an eigenvalue of U if $w = 0$.

We set $\lambda_M = \max(\lambda_U, \lambda_W)$.

Hence, $\max\{|\lambda| \mid \lambda \text{ eigenvalue of } M\} = \lambda_M$. $\qquad \square$

Example

Example

The following matrix M is primitive (hence irreducible).

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

The eigenvalues are the roots φ, φ' of $\det(zI - M) = z^2 - z - 1$, where

$$\varphi = \frac{1 + \sqrt{5}}{2}, \quad \varphi' = \frac{1 - \sqrt{5}}{2}.$$

The vector $\begin{bmatrix} \varphi \\ 1 \end{bmatrix}$ is a (right) positive eigenvector for φ.

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \varphi \\ 1 \end{bmatrix} = \begin{bmatrix} \varphi + 1 \\ \varphi \end{bmatrix} = \begin{bmatrix} \varphi^2 \\ \varphi \end{bmatrix} = \varphi \begin{bmatrix} \varphi \\ 1 \end{bmatrix}.$$

Conjugacy invariants

The (topological) entropy of a shift space X is

$$h(X) = \lim_{n \to \infty} \frac{1}{n} \log \mathrm{Card}(\mathcal{B}_n(X)).$$

The limit exists.

Entropy

Lemma (Fekete Lemma)

Let $a_1, a_2, \ldots$ be a sequence of nonnegative numbers such that $a_{m+n} \leq a_m + a_n$ for $m, n \geq 1$, then $\lim_{n \to \infty} a_n/n$ exists and is equal to $\inf_{n \geq 1} a_n/n$.

Proof.

Let $\alpha = \inf_{n \geq 1} a_n/n$. We have $\inf_{n \geq 1} a_n/n \geq \alpha$ for all $n \geq 1$.
Fix $\varepsilon > 0$. There is $k \geq 1$ such that $a_k/k < \alpha + \varepsilon/2$.
Then for $0 \leq j < k$ and $m \geq 1$,

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

$$\leq \frac{ma_k}{mk} + \frac{ja_1}{mk} \leq \frac{a_k}{k} + \frac{a_1}{m} < \alpha + 1/2\varepsilon + \frac{a_1}{m}.$$

For $n = mk + j$ large enough, $a_1/m < \varepsilon/2$ and $a_n/n < \alpha + \varepsilon$. $\square$

Entropy

Lemma

If X is a shift space,

$$h(X) = \lim_{n \to \infty} \frac{1}{n} \log \operatorname{Card}(\mathcal{B}_n(X)).$$

exists and is equal to

$$\inf_{n \geq 1} \frac{1}{n} \log \operatorname{Card}(\mathcal{B}_n(X)).$$

Proof.

Let $m, n \geq 1$,

$$\operatorname{Card}(\mathcal{B}_{m+n}(X)) \leq \operatorname{Card}(\mathcal{B}_m(X)) \times \operatorname{Card}(\mathcal{B}_n(X)).$$

$\square$

Entropy

Proposition

If there is a sliding block map $\varphi \colon X \to Y$ which is onto, then $h(Y) \leq h(X)$.

Proof.

If φ is a (m, a)-sliding block code, every block in $\mathcal{B}_n(Y)$ is the image of a block in $\mathcal{B}_{n+m+a}(X)$.

Hence, $\mathrm{Card}(\mathcal{B}_n(Y)) \leq \mathrm{Card}(\mathcal{B}_{n+m+a}(X))$.

Thus,

$$
\begin{aligned}
h(Y) &= \lim_{n \to \infty} \frac{1}{n} \log \mathrm{Card}(\mathcal{B}_n(Y)) \\
&\leq \lim_{n \to \infty} \frac{1}{n} \log \mathrm{Card}(\mathcal{B}_{n+m+a}(X)) \\
&= \lim_{n \to \infty} \left(\frac{n + m + a}{n} \right) \frac{1}{n + m + a} \log \mathrm{Card}(\mathcal{B}_{n+m+a}(X)) \\
&= h(X).
\end{aligned}
$$

Corollary

If X and Y are conjugate, then $h(X) = h(Y)$.

Example

If X is the full shift $A^{\mathbb{Z}}$, then $h(X) = \log \mathrm{Card}(A)$.

Hence, the full shift on two letters is not conjugate to the full shift on three letters.

Proposition

Let $\mathcal{A} = (Q, E)$ be a trim deterministic automaton presenting a sofic shift X and $G = (Q, F)$ be the graph of $\mathcal{A}$. Then $h(X) = h(X_G)$.

Proof.

There is a 1-block sliding block code from X_G onto X (with no memory and no anticipation). Thus $h(X) \leq h(X_G)$.

If $\mathcal{A}$ has k states, and since $\mathcal{A}$ is deterministic, $\mathrm{Card}(\mathcal{B}_n(X_G)) \leq k\,\mathrm{Card}(\mathcal{B}_n(X))$. Thus $h(X) \geq h(X_G)$. $\qquad\square$

Computation of the entropy of a sofic shift

Let $G = (Q, E)$ be a graph with adjacency matrix M, we have

$$\mathcal{B}_n(X_G) = \sum_{s,t \in Q} (M^n)_{st}.$$

If G is strongly connected, i.e. if M is irreducible, there is a positive eigenvector v for λ_M.

Let $c = \min v_q$, $d = \max v_q$.

Since $\sum_{t \in Q}(M^n)_{st} v_t = \lambda_M^n v_s$,

$$\sum_{s,t \in Q} (M^n)_{st} v_t = \sum_{s \in Q} \lambda_M^n v_s.$$

$$c \sum_{s,t \in Q} (M^n)_{st} \leq \sum_{s,t \in Q} (M^n)_{st} v_t \leq d \operatorname{Card}(Q) \lambda_M^n.$$

$$\sum_{s,t \in Q} (M^n)_{st} \leq (d/c) \operatorname{Card}(Q) \lambda_M^n.$$

Similarly

$$(c/d)\lambda_M^n \leq \sum_{s,t \in Q} (M^n)_{st}.$$

Proposition

Let $\mathcal{A} = (Q, E)$ be an irreducible deterministic automaton presenting an irreducible sofic shift X and M its adjacency matrix. Then $h(X) = \log \lambda_M$.

The result holds for a trim deterministic automaton presenting a sofic shift X with a reduction to the irreducible components of M (exercise).

> **Example**
>
> For the even shift,
> $$M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$
>
> $$h(X) = \log \varphi, \text{ where } \varphi = \frac{1 + \sqrt{5}}{2}.$$

The even shift and the golden mean shift have the same entropy but are not conjugate.

Periodic points in a shift space

A point x of a shift space X is *periodic* if $S^n(x) = x$ for some $n \geq 1$ and we say that x has *period n*.

If x is periodic, the smallest positive integer n for which $S^n(x) = x$, called the least period of x, divides all periods of x. Let

$$p_n(X) = \mathrm{Card}\{x \in X \mid S^n(x) = x\}.$$

Proposition

Let $\varphi \colon X \to Y$ be a sliding block map. If x is a periodic point of X and has period n, then $\varphi(x)$ is periodic and has period n and the least period of $\varphi(x)$ divides the least period of x. If X and Y are conjugate, then $p_n(X) = p_n(Y)$ for each $n \geq 1$.

Proposition

Let G be a graph of transition matrix M, the number of cycles of length n in G is $\mathrm{tr}(M^n)$ and this equals the number of points in X_G with period n.

Zeta function

The zeta function of a shift space X is the formal series

$$\zeta_X(z) = \exp\left(\sum_{n=1}^{\infty} \frac{p_n(X)}{n} z^n\right).$$

Proposition

If X and Y are conjugate, then $\zeta_X = \zeta_Y$.

Zeta function

Example

Let X be the full shift on a 2-letter alphabet.
Then, $p_n(X) = 2^n$ for each $n \geq 1$.

$$\zeta_X(z) = \exp\left(\sum_{n=1}^{\infty} \frac{2^n}{n} z^n\right) = \exp\left(\sum_{n=1}^{\infty} \frac{(2z)^n}{n}\right)$$
$$= \exp(-\log(1 - 2z)) = \frac{1}{1 - 2z}.$$

Example

Let G be the graph

Its transition matrix is

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

We have $p_n(X_G) = \text{tr}(M^n) = \varphi^n + \varphi'^n$, where

$$\varphi = \frac{1 + \sqrt{5}}{2}, \quad \varphi' = \frac{1 - \sqrt{5}}{2}.$$

are the roots of $\det(zI - M) = z^2 - z - 1$.

Zeta function

Example

$$\zeta_{X_G}(z) = \exp\left(\sum_{n=1}^{\infty} \frac{\varphi^n + \varphi'^n}{n} z^n\right) = \exp\left(\sum_{n=1}^{\infty} \frac{\varphi^n}{n} + \frac{\varphi'^n}{n} z^n\right)$$

$$= \exp(-\log(1 - \varphi z) - \log(1 - \varphi' z)) = \frac{1}{(1 - \varphi z)(1 - \varphi' z)}$$

$$= \frac{1}{1 - z - z^2} = \frac{1}{\det(I - Mz)}.$$

Theorem

Let G be a graph with adjacency matrix M. Then

$$\zeta_{X_G}(z) = \frac{1}{\det(I - Mz)}.$$

Proof.

We have

$$p_n(X_G) = \text{tr}(M^n) = \lambda_1^n + \lambda_2^n + \cdots + \lambda_{|Q|}^n,$$

where $\lambda_1, \lambda_2, \ldots \lambda_{|Q|}$ are the roots of the characteristic polynomial of M listed with multiplicities.

$$
\begin{aligned}
\zeta_{X_G}(z) &= \exp\left(\sum_{n=1}^{\infty} \frac{\lambda_1^n + \lambda_2^n + \cdots + \lambda_{|Q|}^n}{n} z^n\right) \\
&= \exp\left(\sum_{n=1}^{\infty} \frac{(\lambda_1 z)^n}{n} + \cdots + \sum_{n=1}^{\infty} \frac{(\lambda_{|Q|} z)^n}{n}\right) \\
&= \prod_{k=1}^{|Q|} \frac{1}{1 - \lambda_k z} = \frac{1}{\det(I - Mz)}.
\end{aligned}
$$

Observe that the zero eigenvalues "do not count". $\square$

If two edge shifts have the same zeta function, then they have the same entropy.

We have

$$\zeta_{X_G}(z) = \prod_{k=1}^{|Q|} \frac{1}{1 - \lambda_k z} = \frac{1}{\det(I - Mz)}.$$

So the zeta function determines the list of nonzero eigenvalues of the M (with repeated eigenvalues listed according to their multiplicity), called the *nonzero spectrum of M*.

Thus, it determines the largest eigenvalue of M, and thus the entropy.

If G is strongly connected $h(X_G) = \log \lambda_M$. Otherwise $h(X_G) = \max \log \lambda_{M_i}$, where M_i is the adjacency matrix of all strongly connected components of G.

Example

Let G and H be the graphs
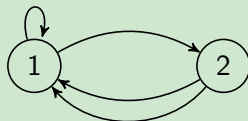
Show that the two edge shifts X_G and X_H are not conjugate although they have the same entropy.

Applications

Let M be the adjacency matrix of G and N the adjacency matrix of N.

$$\zeta_{X_G}(z) = \frac{1}{\det(I - Mz)}, \zeta_{X_H}(z) = \frac{1}{\det(I - Nz)}.$$

We have

$$\det(I - Mz) = \det \begin{bmatrix} 1 - z & -z \\ -z & 1 - z \end{bmatrix} = 1 - 2z.$$

$$\det(I - Nz) = \det \begin{bmatrix} 1 - z & -z \\ -2z & 1 \end{bmatrix} = 1 - z - 2z^2.$$

Hence X_G and X_H have not the same zeta function, and thus are not conjugate.

We can also see that $p_2(X_G) = 4$ and $p_2(X_H) = 5$.

We have

$$\det(zI - M) = \det \begin{bmatrix} z-1 & -1 \\ -1 & z-1 \end{bmatrix} = z^2 - 2z = z(z-2).$$

Also obtained directly from $\det(I - Mz)$ by changing z into $1/\lambda$ and mutiplying by λ^2: $\lambda^2(1 - 2/\lambda) = \lambda(\lambda - 2)$.

$$\det(zI - N) = \det \begin{bmatrix} z-1 & -1 \\ -2 & z \end{bmatrix} = z^2 - z - 2 = (z-2)(z+1).$$

Hence, M and N have the same largest eigenvalue. Thus, X_G and X_H have the same entropy.