Lower bound on the entropy of two dimensional shifts of finite type

LA Xuan Hoang

Advisors: Nathalie AUBRUN and Michael RAO

LIP, team MC2

September 4, 2017

Outline

1. Shifts of finite type [Lind and Marcus, 1995]
2. Entropy [Lind and Marcus, 1995]
3. Two dimensional shifts of finite type [Chan and Rechnitzer, 2014]

Shifts of finite type

1. What is a shift space?
2. Shifts of finite type

What is a shift space?

Definitions

- An *alphabet* $\mathcal{A}$ is a set of letters (often $\{0,1\}$).
- A *word* w is a bi-infinite sequence of letters:

$$w = ...w_{-1}w_0.w_1w_2... \in \mathcal{A}^{\mathbb{Z}}.$$

- A *(n-)block* u is a finite sequence of letters (of length n):

$$u = u_1...u_n \in \mathcal{A}^n.$$

Examples:

- $w = ...1010.1010... = (10)^{\infty}. \in \{0,1\}^{\mathbb{Z}}$
  contains $u = 1010 \in \{0,1\}^4$.
- $\mathcal{A} = \{a, b, ..., z\}$, $w = (thisisaword)^{\infty}.(itcontainsablock)^{\infty} \in \mathcal{A}^{\mathbb{Z}}$
  contains $u = block \in \mathcal{A}^5$.

What is a shift space?

Definitions

- A *set of forbidden blocks* $\mathcal{F}$ is a set of blocks over $\mathcal{A}$.
- A *shift space* $X_{\mathcal{F}}$ is the set of words of $\mathcal{A}^{\mathbb{Z}}$ that do **not** contain any block from $\mathcal{F}$.

Examples of shift spaces over $\mathcal{A} = \{0, 1\}$:

- $\mathcal{F} = \varnothing$, $X_{\mathcal{F}} = \mathcal{A}^{\mathbb{Z}}$ is called the *full 2-shift*.
- $\mathcal{F} = \{11\}$, $X_{\mathcal{F}}$ is called the *golden mean shift*.
  $w = 0^{\infty}.101001000... \in X_{\mathcal{F}}$.
  $w = ..\underline{11}00.\underline{11}00\underline{11}00... \notin X_{\mathcal{F}}$.
- $\mathcal{F} = \{10^n1 |\ n \text{ is odd}\}$, $X_{\mathcal{F}}$ is called the *even shift*.
  $w = 0^{\infty},\ w = 1^{\infty} \in X_{\mathcal{F}}$.
  $w = ...1\underline{00.01}... \notin X_{\mathcal{F}}$.

What is a shift space?

Definition

The *language* $\mathcal{L}$ (of allowed blocks) of a shift space $\mathcal{Y}$ is the set of all blocks that appear in $\mathcal{Y}$:

$$\mathcal{L}(\mathcal{Y}) = \bigcup_{n \in \mathbb{N}} \mathcal{L}_n(\mathcal{Y}).$$

where $\mathcal{L}_n(\mathcal{Y})$ is the set of all n-blocks that appear in $\mathcal{Y}$.

Examples of languages of shift spaces over $\mathcal{A} = \{0, 1\}$:

- The full 2-shift has language $\{\varepsilon, 0, 1, 00, 01, 10, 11, ...\} = \mathcal{A}^*$.
- The golden mean shift has language $\{\varepsilon, 0, 1, 00, 01, 10, 000, 001, 010, 100, 101, ...\}$.
- The even shift has language $\{\varepsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 110, 111, ...\}$.

Shifts of finite type

Definition

A *shift of finite type (SFT)* is a shift space $\mathcal{Y}$ such that there exists a finite set $\mathcal{F}$ of forbidden blocks that verifies $\mathcal{Y} = X_{\mathcal{F}}$.

Examples:

- The full 2-shift is an SFT since $\mathcal{F} = \varnothing$.
- The golden mean shift is an SFT since $\mathcal{F} = \{11\}$.
- The even shift is not an SFT because there is no finite $\mathcal{F}$ that can "describe" it.

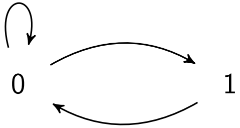
Shifts of finite type

Definition

A *graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a set of vertices $\mathcal{V}$ and a set of edges $\mathcal{E} \subseteq \mathcal{V}^2$. An *adjacency matrix* A is defined by $A_{i,j} = 1$ if and only if $(i, j) \in \mathcal{E}$ and 0 otherwise.

Example:

The graph represented by the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ is the following:



Shifts of finite type

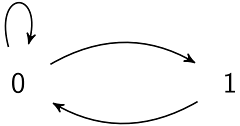
Definition

Let $\mathcal{G}$ be a graph, the *vertex shift* $\mathcal{X}_\mathcal{G}$ (or $\mathcal{X}_A$ where A is the adjacency matrix) over the alphabet $\mathcal{V}$ is:

$$\mathcal{X}_\mathcal{G} = \{x = (x_i)_{i\in\mathbb{Z}} \in \mathcal{V}^\mathbb{Z} | \forall i \in \mathbb{Z}, A_{x_i,x_{i+1}} = 1\}.$$

Example:

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $\mathcal{X}_A$ is the golden mean shift.



Entropy

1. What is the entropy of a shift space?
2. Computing the entropy

What is the entropy of a shift space?

Definition

The *entropy* h of a shift space $\mathcal{Y}$ is:

$$h(\mathcal{Y}) = \lim_{n\to\infty} \frac{1}{n} log_2 |\mathcal{L}_n(\mathcal{Y})|.$$

Examples:

- $h(\{0,1\}^{\mathbb{Z}}) = 1$.
- $h(\text{golden mean shift}) = log_2(\frac{1+\sqrt{5}}{2})$.

Proposition

Let $\mathcal{G}$ be a graph and A its adjacency matrix:

$$|\mathcal{L}_n(\mathcal{X}_{\mathcal{G}})| = \sum_{i,j} A^n_{i,j}.$$

Computing the entropy

The Perron-Frobenius theorem

Let A be a nonzero and irreducible matrix. Then A has a positive eigenvector v with corresponding positive eigenvalue $\lambda_P(A)$, called the Perron eigenvalue, that is geometrically simple and algebraically simple. If μ is another eigenvalue for A, then $|\mu| < \lambda_P(A)$. Any positive eigenvector for A is a positive multiple of v.

Remarks:

- Irreducible matrix A: $\forall (i,j), \exists n, A^n_{i,j} > 0$.
- $\exists (c,d) \in \mathbb{R}^2_+, c\lambda^n_P(A) \leq \sum_{i,j} A^n_{i,j} \leq d\lambda^n_P(A)$

Proposition

$$h(\mathcal{X}_A) = log_2(\lambda_P(A)).$$

Computing the entropy

Entropy of the golden mean shift:

- $X_{\mathcal{F}}$ where $\mathcal{F} = \{11\}$.
- $X_{\mathcal{F}} = \mathcal{X}_{\mathcal{G}}$ where $\mathcal{G} =$

$$\begin{tikzcd} 0 \arrow[loop above] \arrow[rr, bend left] & & 1 \arrow[ll, bend left] \end{tikzcd}$$

- $\mathcal{X}_{\mathcal{G}} = \mathcal{X}_A$ where $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.
- $\lambda_P(A) = \frac{1+\sqrt{5}}{2}$.
- $h(\text{golden mean shift}) = log_2(\frac{1+\sqrt{5}}{2})$.

Two dimensional shifts of finite type

1. What is a two dimensional shift of finite type?
2. Strips systems
3. Lower bound on the entropy

What is a two dimensional shift of finite type?

Definitions

- A *word* w is an infinite matrix of letters.
- A *block* u is a finite matrix.

$$w = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \dots & 1 & 0 & 1 & 0 & \dots \\ \dots & 0 & 1 & 0 & 1 & \dots \\ \dots & 1 & 0 & 1 & 0 & \dots \\ \dots & 0 & 1 & 0 & 1 & \dots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \in \mathcal{A}^{\mathbb{Z}^2}.$$

What is a two dimensional shift of finite type?

Examples of 2D-SFTs (also called "constraint") over $\mathcal{A} = \{0, 1\}$:

- $\mathcal{F} = \{\begin{bmatrix}1 & 1\end{bmatrix}, \begin{bmatrix}1 \\ 1\end{bmatrix}\}$, $X_{\mathcal{F}}$ is called the hard square (HS) constraint.
- $\mathcal{F} = \{\begin{bmatrix}1 & 1\end{bmatrix}, \begin{bmatrix}1 & * \\ * & 1\end{bmatrix}, \begin{bmatrix}* & 1 \\ 1 & *\end{bmatrix}\}$, $X_{\mathcal{F}}$ is called the read-write isolated memory (RWIM) constraint.
- $\mathcal{F} = \{\begin{bmatrix}1 & 1\end{bmatrix}, \begin{bmatrix}1 \\ 1\end{bmatrix}, \begin{bmatrix}1 & * \\ * & 1\end{bmatrix}, \begin{bmatrix}* & 1 \\ 1 & *\end{bmatrix}\}$, $X_{\mathcal{F}}$ is called the non attacking kings (NAK) constraint.

What is a two dimensional shift of finite type?

Definition

The entropy h of a constraint $\mathcal{Y}$ is defined as follows:

$$h(\mathcal{Y}) = \lim_{n\to\infty} \lim_{m\to\infty} \frac{1}{nm} |\mathcal{L}_{nm}(\mathcal{Y})|$$

where $\mathcal{L}_{nm}(\mathcal{Y})$ is the collection of different allowed matrices of size $n \times m$.

Strip systems

Definition

The strip system S_n of a 2D-SFT $\mathcal{Y}$ is a 1D-SFT $S_n(\mathcal{Y})$ defined as the set of allowed matrices of $\mathcal{Y}$ of height n over the alphabet of all allowed columns of height n.

A word $w \in S_4(\mathcal{Y})$ has the following form:

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

Remark:

$$\lim_{n\to\infty} \frac{h(S_n(\mathcal{Y}))}{n} = h(\mathcal{Y}).$$

However, in general, the entropy of a 2D-SFT is non computable. [Hochman and Meyerovitch, 2010]

Lower bound on the entropy

Conditions:

- Constraint over $\mathcal{A} = \{0, 1\}$.
- Forbidden blocks of size 2×2.
- The symmetry across a vertical line of a forbidden block must still be forbidden.

Considered constraints: HS, RWIM, NAK.

HS: $\left\{ \begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

RWIM: $\left\{ \begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * \\ * & 1 \end{bmatrix}, \begin{bmatrix} * & 1 \\ 1 & * \end{bmatrix} \right\}$

NAK: $\left\{ \begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & * \\ * & 1 \end{bmatrix}, \begin{bmatrix} * & 1 \\ 1 & * \end{bmatrix} \right\}$

Lower bound on the entropy

Definitions

- $\omega \begin{bmatrix} a & b \\ c & d \end{bmatrix} = 1$ if the block is allowed and 0 otherwise.
- The column transfer matrix $V^{[m]}$ (of size $2^m \times 2^m$) of a strip of height m:
  $V^{[m]}_{\sigma,\tau} = \prod_{i=1}^{m} \omega \begin{bmatrix} \sigma_{i+1} & \tau_{i+1} \\ \sigma_i & \tau_i \end{bmatrix}$

  $= 1$ if σ and τ can be consecutive columns and 0 otherwise.

Proposition

$V^{[m]}$ is symmetric so:

$$\lambda_P(V^{[m]}) = max_{\psi} \frac{\psi^T V^{[m]} \psi}{\psi^T \psi}.$$

Lower bound on the entropy

Propositions

- $\exists R$ symmetric, $\psi^T \psi = Tr(R^m)$.
- $\exists S$ symmetric, $\psi^T V^{[m]} \psi = Tr(S^m)$.
- Let ξ and η be the Perron eigenvalue of R and S respectively:

$$\frac{Tr(S^m)}{Tr(R^m)} \leq \lambda_P(V^{[m]}) \Rightarrow \frac{\eta}{\xi} \leq \lim_{m\to\infty} \lambda_P(V^{[m]})^{1/m}.$$

Lower bound on the entropy

Definitions

- Let X be the dominant eigenvector (of size $2n^2$) of R, we define $X(a)$ the vector of size n^2 that verifies: $X(a)_{\alpha,\beta} = X_{(\alpha|a|\beta)}$.
- Let Y be the dominant eigenvector (of size $4n^2$) of S, we define $Y(a,b)$ the vector of size n^2 that verifies: $Y(a,b)_{\alpha,\beta} = Y_{(\alpha|a,b|\beta)}$.

Propositions

- $\xi X = RX \Rightarrow \xi X(a) = \sum_b F(a,b)X(b)F(b,a)$.
- $\eta Y = SY \Rightarrow \eta Y(a,b) = \sum_{c,d} \omega \begin{bmatrix} a & b \\ c & d \end{bmatrix} F(a,c)Y(c,d)F(d,b)$.

Lower bound on the entropy

Corner transfer matrix renormalisation group

There exist a set of $n \times n$ matrices $A(a)$, called the corner transfer matrices, and a set of matrices $F(a, b)$, called the half column/row transfer matrices, that satisfy:

$$X(a) = A^2(a) \text{ and } Y(a, b) = A(a)F(a, b)A(b).$$

Remark: The constraint must be invariant under rotation by $\pi/2$. Only HS and NAK verify this condition.

HS: $\left\{ \begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

RWIM: $\left\{ \begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * \\ * & 1 \end{bmatrix}, \begin{bmatrix} * & 1 \\ 1 & * \end{bmatrix} \right\}$

NAK: $\left\{ \begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & * \\ * & 1 \end{bmatrix}, \begin{bmatrix} * & 1 \\ 1 & * \end{bmatrix} \right\}$

Lower bound on the entropy

Propositions

Suppose we can construct A and F for a matrix of size $2^p \times 2^p$, we can define A_l and F_l of size $2^{p+1} \times 2^{p+1}$:

$$A_l(c)_{d,a} = \sum_b \omega \begin{bmatrix} a & b \\ c & d \end{bmatrix} F(d,b)A(b)F(b,a).$$

$$F_l(d,c)_{b,a} = \omega \begin{bmatrix} a & b \\ c & d \end{bmatrix} F(b,a).$$

These equations can be intuitively understood with the following diagram:

$$\begin{bmatrix} & & \nwarrow \\ & A_l & \\ & & \\ c & & \end{bmatrix} = \begin{bmatrix} & \leftarrow & & \nwarrow \\ & F & A & \\ a & b & & \\ c & d & F & \uparrow \end{bmatrix}$$

Lower bound on the entropy

The algorithm [Chan and Rechnitzer, 2014]

1. Start with $A(a) = F(a, b) = [1]$ and $n = 1$.
2. Expand A and F into A_l and F_l.
3. Increase n by 1 under a certain condition.
4. Diagonalize A_l and let P be the matrix of the eigenvectors corresponding to the n largest eigenvalues.
5. Apply the similarity tranforms.
6. Normalize A and F so that the top-left elements of $A(0)$ and $F(0, 0)$ are both 1.
7. Go back to step 2.
8. Once n is sufficient, compute the initial X and Y.
9. Apply the power method until a desired precision is reached.

Lower bound on the entropy

Implementation: C++ with MPFR and Eigen libraries.

Condition for step 3

- Convergence rate of A and $F \rightarrow$ they do not always converge.
- Convergence rate of the eigenvalues of $A \rightarrow$ lack of precision.
- Convergence rate of the eigenvalues and the eigenvectors of $A \rightarrow$ only works for the constraints in the article.

Lower bound on the entropy

Some results:

- HS: 1.503 048 082 475 332 264 322 1
- RWIM: 1.448 957 4
- NAK: 1.342 643 951

References

Yao-ban Chan and Andrew Rechnitzer (2014)

Accurate lower bounds on two dimensional constraint capacities from corner transfer matrices.

IEEE Transactions on Information Theory 60, 3845-3858.

Mike Hochman and Tom Meyerovitch (2010)

A characterization of the entropies of multidimensional shifts of finite type.

Annals of Mathematics 171(3), 2011-2038.

Douglas Lind and Brian Marcus (1995)

An Introduction to Symbolic Dynamics and Coding.

Cambridge University Press.