Lecture 2: An introduction to entropy via exercises

Nishant Chandgotia

Hebrew University of Jerusalem

September

Shifts of finite type

Recall that every subshift $X \subset \mathbb{A}^{\mathbb{Z}^d}$ is obtained as the set of configurations avoiding finitely many patterns from a forbidden list $\mathcal{F}$. $B_n := \{1, 2, \ldots, n\}^d$ is a box of size n. SFTs are those subshifts in which the forbidden list can be chosen to be finite. The language allowed on the shape B_n is given by

$$\mathcal{L}(X, B_n) := \{x|_{B_n} \ : \ x \in X\}.$$

Then the entropy is defined as

$$h_{\text{top}}(X) := \lim_{n \to \infty} \frac{1}{|B_n|} \log(|\mathcal{L}(X, B_n)|).$$

Nearest neighbour SFTs

Nearest neighbour SFTs are those SFTs in which the forbidden list can be are supported on shapes like $\{\vec{0}, \vec{e}_1\}$, $\{\vec{0}, \vec{e}_2\}$, etc.

Some Examples

The hard-core shift is the subshift with alphabet $\{0, 1\}$ where adjacent 1's are disallowed.

The even shift is the subshift with alphabet $\{0, 1\}$ where the gap between successive 1's is even.

The space of proper k-colourings is the subshift with alphabet $\{0, 1, \ldots, k-1\}$ where adjacent symbols are forced to be distinct.

The hard-core shift and the space of proper k-colourings are nearest neighbour SFTs while the even shift is not even an SFT.

Question 1

Fix $d = 1$. Compute the entropy of

1. The hard-core shift.
2. The even shift.
3. The space of proper k-colourings.

Factors and conjugacy

Recall that a embedding/conjugacy/factor map $\phi : X \to Y$ is a continuous bijective/surjective map which is equivariant.

A sliding block map is a map $\phi : X \to Y$ with the following structure: There exists $B \subset \mathbb{Z}^d$ and a map $\Phi : \mathcal{L}(X, B) \to \mathbb{A}$ such that $\phi(x)_{\vec{i}} := \Phi(\sigma^{\vec{i}}(x)|_B)$.

Theorem (Curtis-Hedlund-Lyndon)

A map $\phi : X \to Y$ is equivariant and continuous if and only if ϕ is a sliding block code.

Question 2

1. Prove that if Y can be embedded into X then $h_{\mathrm{top}}(Y) \leq h_{\mathrm{top}}(X)$.
2. Prove that if Y is a factor of X then $h_{\mathrm{top}}(Y) \leq h_{\mathrm{top}}(X)$.
3. Prove that if Y is conjugate to X then $h_{\mathrm{top}}(Y) = h_{\mathrm{top}}(X)$.

Question 3

1. Prove that shift spaces conjugate to SFTs are still SFTs.
2. Prove that every SFT is conjugate to a nearest neighbour SFT.

Shannon entropy

The *Shannon entropy* of probability measure μ on a countable set N is defined as

$$H(\mu) := \sum_{n \in N} -\mu(n) \log(\mu(n))$$

where $0.log(0) = 0$.

Theorem (Jensen's Inequality)

Let $(\Omega, \mu, \mathcal{B})$ be a probability space and $f \in L^1(\mu)$. Then for all convex functions $\phi : \mathbb{R} \to \mathbb{R}$,

$$\phi\left(\int_\Omega f(\omega) d\mu(\omega)\right) \leq \int_\Omega \phi \circ f(\omega) d\mu(\omega)$$

where equality occurs if ϕ is linear on the image of f or f is a constant function.

Question 4

1. $H(\mu) = 0$ if and only if μ is deterministic, that is, $\mu(x) = 1$ for some $x \in N$.
2. If μ is a probability measure on a finite set N then prove that the entropy is maximised for the uniform probability measure. (Hint: $x \to -\log(x)$ is a convex function).

Question 5

Let X be a nearest neighbour SFT. Prove that $H(\mu)$ is maximised (over possible probability measures μ on $\mathcal{L}(X, B_n)$ exactly when μ is a uniform Gibbs measure, meaning, that for all finite sets $B \subset B_n$ and $a \in supp(\mu)$ and $b \in \mathcal{L}(X, B_n)$,

$$\mu([b]_B \ : \ [a]_{B_n \setminus B}) = \mu([b]_B \ : \ [a]_{\partial B \cap B_n})$$

and is uniform on $b \in \mathcal{L}(X, B_n \cup \partial B_n)$ such that $b|_{\partial B_n} = a|_{\partial B_n}$.

Is the assumption of being a nearest neighbour SFT necessary?

Measure theoretic entropy

Given a shift space X, let $\mathcal{P}(X)$ denote the space of all shift-invariant probability measures on X. (Probability measures which are left-invariant under the shift action). Under the weak topology $\mathcal{P}(X)$ is a compact space (by the Banach-Alaoglu theorem but in our particular case it is just a diagonalisation argument).

The measure-theoretic entropy of μ is given by

$$h_\mu := \lim_{n\to\infty} \frac{1}{|B_n|} H(\mu_{B_n}).$$

Again by subadditivity it can be proven that the limit exists and

$$h_\mu = \inf \frac{1}{|B_n|} H(\mu_{B_n}).$$

Question 6

1. Prove that $\mu \to h_\mu$ is upper semi-continuous and that there is a measure $\mu \in \mathcal{P}(X)$ which maximises measure-theoretic entropy.
2. Exhibit an example to see that $\mu \to h_\mu$ need not be upper semi-continuous.

Measures of maximal entropy

Measures of maximal entropy of nearest neighbour SFTs are uniform Gibbs measures.

Thus for the full shift $\{0,1\}^{\mathbb{Z}^d}$, and $x, y \in \{0,1\}^{\mathbb{Z}^d}$ and $A \subset \mathbb{Z}^d$ a finite set we have

$$\mu([x]_A \mid [y]_{\mathbb{Z}^d \setminus A}) = \frac{1}{2^{|A|}}.$$

This is the same as choosing either 0 or 1 at each site in $\mathbb{Z}^d$ independently and with equal probability.

Variational Principle

If X is a subshift then

$$\sup_{\mu \in \mathbb{P}(X)} h(\mu) = h_{\mathsf{top}}(X).$$

Question 7

Prove that shift-invariant uniform Gibbs measures for the hard-core shift are measures of maximal entropy.

Can you identify a weaker assumption that will let the proof go through?