Math 254B Lecture 16 Notes

Daniel Raban

May 6, 2019

1 Topological Entropy

1.1 Entropy of shift systems

Definition 1.1. Fix k, let $\Sigma_k = \{0, 1 \dots, k-1\}^{\mathbb{N}}$, and let σ be the shift operator. This called the **full shift**. A **subshift** is a closed subset $X \subseteq \Sigma_k$ such that $\sigma(X) = X$.

We will often write σ instead of $\sigma|_X$. If $x \in \Sigma_k \setminus X$ is open, there is $n \in \mathbb{N}$ and a word $w = w_0 \cdots w_{n-1} \in \{0, \dots, k-1\}^n$ such that $(x_0, \dots, x_{n-1} = w$ and $[w] = \{w : (x_0 \cdots x_{n-1}) = w\} \subseteq \Sigma_k \setminus X$. This is called a **forbidden word** of length n.

If w is forbidden of length n and $x \in \Sigma_k$ has $(x_i, x_{i+1}, \dots, x_{i+n-1}) = w$, then $x \in \Sigma_k \setminus X$.

Definition 1.2. X is a **subshift of finite type** (**SFT**) if there exist fiintely many forbidden words $w^{(1)}, \dots, w^{(n)}$ such that $X = \{x \in \Sigma_k : (x_i, \dots, x_{i+n^{(\ell)}}) \neq w^{(\ell)} \forall n \forall \ell = 1, \dots, m\}$.

Call words that are not forbidden **permitted**, and let $N_n(X)$ be the number of permitted words of length n. Observe that $\log(N_n(X))$ is a subadditive sequence:

$$N_{n+m}(X) \leq N_n(X) N_m(X) \qquad \forall n, m \geq 1.$$

This is because $w \in \{0, \dots, k-1\}^{n+m}$ is permitted, then so are $w_{[0,n-1)}$ and $w_{[n,n+m-1)}$. So we can look at

$$\lim_n \frac{1}{n} \log(N_n(X)),$$

the growth rate of the number of permitted words.

1.2 Packing numbers and covering numbers

What about on general topological dynamical systems? On Σ_k, let

$$\rho(x, y) = 2^{-\min\{n : x_n \neq y-n\}}.$$

This is a metric.[1] Observe that $\rho(x, y) \leq 2^{-n}$ iff $(x_0, \cdots, x_{n-1}) = (y_0, \cdots, y_{n-1})$.

[1] It's actually an ultrametric, even.

Definition 1.3. In any (X, ρ), given $\delta > 0$, define the **δ-packing number** as

$$\operatorname{pack}_\delta(X, \rho) := \max\{|F| : F \subseteq X, \rho(x, y) \geq \delta \, \forall x \neq y \in F\}.$$

In our setting, $N_n(X) = \operatorname{pack}_{2^{-n}}(X, \rho)$. Packing numbers are not so easy to work with, so let's introduce a few other notions.

Definition 1.4. In any (X, ρ), given $\delta > 0$, define the **δ-covering number** as

$$\operatorname{cov}_\delta(X, \rho) = \min\{|F| : \bigcup_{x \in F} B_\delta(x) = X\}.$$

Also define[2]

$$\operatorname{cov}'_\delta(X, \rho) = \min\{|\mathcal{E}| : \mathcal{E} \subseteq \mathscr{P}(X), \bigcup \mathcal{E} = X, \operatorname{diam}(E) < 2\delta \, \forall E \in \mathcal{E}\}.$$

Lemma 1.1. *Let $\delta > 0$. Then*

$$\operatorname{cov}_{2\delta} \leq \operatorname{cov}'_\delta \leq \operatorname{cov}_\delta,$$

$$\operatorname{pack}_{2\delta} \leq \operatorname{cov}_\delta \leq \operatorname{pack}_\delta.$$

Proof. If $\mathcal{E}$ as in the definition, pick $x_E \in E$ for all $E \in \mathcal{E}$. Then $\bigcup_{E \in \mathcal{E}} B_{2\delta}(x_i) = X$.

A best possible δ-packing gives a candidate for a δ-covering. □

1.3 Definition of topological entropy

Given a TDS (X, T), pick a compact metric ρ. The idea is to count the number of "effectively distinguished" orbits by time n.

Definition 1.5. The **Bowen metric**[3] at time n is

$$\rho_n(x, y) = \max_{1 \leq i \leq n-1} \rho(T^i x, T^i y).$$

This is the worst possible distance of the orbits of x and y by time n. We are taking a max of more times as n grows, so $\rho \leq \rho_1 \leq \rho_2 \leq \rho_3 \leq \cdots$.

Define

$$N_{n,\delta}(X, T, \rho) := \operatorname{cov}'_\delta(X, \rho_n).$$

Lemma 1.2. *The sequence* $\log(N_{n,\delta})$ *is subadditive:*

$$N_{n+m,\delta} \leq N_{n,\delta} N_{m,\delta}.$$

[2] This second notion is not so standard, but it is convenient for our purposes.

[3] Since T is not necessarily injective, if we do not take the max, this is only a pseudometric.

Proof. Let $\mathcal{E}$ be a minimal covering by sets of ρ_n-diameter $< 2\delta$. Let $\mathcal{F}$ be a minimal covering by sets of ρ_m-diameter $< 2\delta$.Define $\mathcal{F}' = \{T^{-n}(U) : U \in \mathcal{F}\}$. Check that if $x, y \in T^{-n}(U) \in \mathcal{F}'$, then $T^n x, T^n y \in U$, so $\rho_m(T^n x, T^n y) < 2\delta$; i.e.

$$\max_{n \leq i \leq n+m-1} \rho(T^{n+i}x, T^{n+i}y) < 2\delta.$$

Define $\mathcal{G} = \{E \cap F : E \in \mathcal{E}, F \in \mathcal{F}'\}$. If $x, y \in E \cap F$, then $\rho(n+m)(x, y) < 2\delta$; this is uniform in x, y. Now $|\mathcal{G}| \leq |\mathcal{E}| \cdot |\mathcal{F}'| = |\mathcal{E}| \cdot |\mathcal{F}|$. $\square$

Definition 1.6. Define $h_{\text{top},\delta}(X, T, \rho) = \lim_n \frac{1}{n} N_{n,\delta}$. The **topological entropy** of (X, T, ρ) is

$$h_{\text{top}}(X, T, \rho) := \lim_{\delta \to 0} \frac{1}{n} \log(h_{\text{top},\delta}(X, T, \rho)) = \sup_{\delta > 0} \frac{1}{n} \log(h_{\text{top},\delta}(X, T, \rho)).$$

Lemma 1.3. *It does not matter that we used* cov'*:*

$$h_{\text{top}} = \sup_{\delta > 0} \frac{1}{n} \log \text{cov}_\delta(X, \rho_n) = \sup_{\delta > 0} \frac{1}{n} \log \text{pack}_\delta(X, \rho_n).$$

1.4 Properties of topological entropy

Proposition 1.1. *Let* $\pi : (X, T, \rho_X) \to (Y, S, \rho_Y)$ *be a semiconjugacy. Then*

$$h_{\text{top}}(X, T, \rho_X) \geq h_{\text{top}}(Y, S, \rho_Y).$$

Proof. By compactness π is uniformly continuous. So for any $\varepsilon > 0$, there is a $\delta > 0$ such that $\rho_X(x, y) < \delta \implies \rho_Y(\pi x, \pi y) < \varepsilon$. Then

$$(\rho_X)_n(x, y) < \delta \implies (\rho_Y)_n(\pi x, \pi y) < \varepsilon.$$

So if $F \subseteq X$, then

$$\bigcup_{x \in F} B_\delta^{(\rho_{X,n})}(x) = X \implies \bigcup_{y \ in \pi(F)} B_\varepsilon^{(\rho_{Y,n})}(y) = Y. \qquad \square$$

Corollary 1.1. *Topological entropy is independent of the metric:* $h_{\text{top}}(X, T, \rho) = h_{\text{top}}(X, T)$.

Proof. With two competing metrics on X, $\pi = \text{id}_X$ is a semiconjugacy in both directions. $\square$

Lemma 1.4. *Entropy dilates with time steps:* $h_{\text{top}}(X, T^k) = k h_{\text{top}}(X, T)$.

Proof. Observe that $(\rho_k^T)_n^{T^k} = \rho_{kn}^T$. Take logs, normalize, and send $n \to \infty$. $\square$

Lemma 1.5. $h_{\text{top}}(X \times Y, T \times S) = h_{\text{top}}(X, T) + h_{\text{top}}(Y, S)$

Example 1.1. Let (Σ_k, σ) have metric $\rho(x, y) = 2^{-\max\{n: x_n \neq y_n\}}$. Then

$$\rho_n = \begin{cases} 1 & (x_1, \dots, x_{n-1}) \neq (y_0, \dots, y_{n-1}) \\ 2^{-k} & \text{first difference is at time } n + k. \end{cases}$$

So

$$\operatorname{pack}_\delta(X, \rho_n) = N_{n+\log_2(1/\delta)}(X).$$

So you can check that topological entropy agrees with our definition on subshifts.

What does this have to do with invariant measures? Next time, we will show that for any $\mu \in P^\sigma(X)$. $H(\mu_{[0,n)}) \leq \log(N_n(X))$, and

$$h_\mu(\sigma) \leq h_{\text{top}}(\sigma|_X).$$