Entropy of dynamical systems

Siddhartha Bhattacharya

1 Drift

Let (X, d) be a metric space and let $T : X \to X$ be a an isometry. We fix two points $x_0, y_0 \in X$ and define

$$Drift(T) = \inf_n \frac{d(x_0, T^n y_0)}{n}.$$

Since T is an isometry, $d(x, T^n y) \leq d(x, x_0) + d(x_0, T^n y_0) + d(y_0, y)$ for all n. Dividing both sides by n and taking infimum we see that the value of $Drift(T)$ does not depend on the choice of x_o and y_0.

Examples : Suppose $X = \mathbb{R}$, equipped with the usual metric, and $Tx = x + 1$. Then it is easy to see that $Drift(T) = 1$. If $X = \mathbb{R}^2$ and T is a rotation, then $Drift(T) = 0$ as $\{T^n y\}$ is a bounded subset for all $y \in X$.

The following proposition shows that drift is a conjugacy invariant.

Proposition 1.1 *For $i = 1, 2$; let T_i be an isometry of a metric space (X_i, d_i) such that $T_1 = S^{-1} \circ T_2 \circ S$ for some bijective isometry $S : X_1 \to X_2$. Then $Drift(T_1) = Drift(T_2)$.*

Proof. Fix $x_0, y_0 \in X_1$. Since S is an isometry, it follows that

$$d(Sx_0, T_2^n(Sy_0)) = d(x_0, S^{-1} \circ T_2^n \circ S(y_0)) = d(x_0, T_1^n y_0)$$

for all n. Dividing both sides by n and taking infimum we see that $Drift(T_1) = Drift(T_2)$. $\square$

Note that the only property of d that is needed in the above definition is the triangle inequality. This observation allows us to define drift in more general situations. For example, suppose G is a finitely generated group and X_G is the collection of all finite generating subsets of G, i.e.,

$$X_G = \{A \subset G : |A| < \infty, \cup_1^\infty A^i = G\}.$$

We define a function $d : X_G \times X_G \to \mathbb{R}$ by

$$d(A, B) = log(\inf\{k : B \subset A^k\}).$$

Clearly, if $B \subset A^m$ and $C \subset B^n$ then $C \subset A^{mn}$. In particular, $d(A, C) \leq d(A, B) + d(B, C)$. If θ is an automorphism of G then $B \subset A^m$ if and only if $\theta(B) \subset \theta(A)^m$. This shows that $d(A, B) = d(\theta(A), \theta(B))$ for all $A, B \in X_G$. We define $Drift(\theta) \in \mathbb{R}$ by

$$Drift(\theta) = \inf_n \frac{d(A, \theta^n(B))}{n}.$$

As in the previous case, this number does not depend on the choice of A and B and it is a conjugacy invariant of θ.

More generally, let X be a set and let $||.||$ be a non-negative function from X to $\mathbb{R}$. For any map $T : X \to X$ we define a map $D_T : X \to \mathbb{R}$ and a real number $D(T) \in \mathbb{R}$ by

$$D_T(x) = \inf_n \frac{||T^n x||}{n}, \quad D(T) = \sup_x D_T(x).$$

Suppose (Y, d) is a metric space and $T : Y \to Y$ is an isometry. We set $X = Y \times Y$, $||(x, y)|| = d(x, y)$, $T^*(x, y) = (x, Ty)$; and note that $Drift(T) = D(T^*)$. In this special case, the function D_{T^*} is actually a constant.

2 Entropy of a semigroup isometry

An abelian semigroup S is said to be *idempotent* if $x + x = x$ for all $x \in S$. If S is an abelian idempotent semigroup and $x, y \in S$; we say $x \leq y$ if $x+y = y$. Note that this defines a partial order on S. A *norm* on an abelian idempotent semigroup S is a non-negative map $||.|| : S \to \mathbb{R}$ such that $||x|| \leq ||x + y|| \leq ||x|| + ||y||$ for all $x, y \in S$. The condition $||x|| \leq ||x + y||$ for all x and y is equivalent to the condition that $||.||$ is order preserving. If S and $S^{'}$ are normed idempotent abelian semigroups, a map $T : S \to S^{'}$ is said to be an *isometry* if $T(x + y) = Tx + Ty, ||Tx|| = ||x||$ for all $x, y \in S$. It is easy to see that the collection of normed idempotent abelian semigroups, together with isometries as morphisms forms a category.

For any semigroup isometry $T : S \to S$ we define a map $I + T : S \to S$ by $(I + T)x = x + Tx$. The number $D(I + T)$ is called *the entropy* of T and is denoted by $h(T)$. Note that $h(T) = \sup\{||x||_T : x \in S\}$, where

$$||x||_T = \inf_n \frac{||(I + T)^n x||}{n} = \inf_n \frac{||x + \cdots + T^{n-1}x||}{n}.$$

It is easy to see that for any semigroup isometry $T : S \to S$, the function $||.||_T$ is another norm on S satisfying $||x||_T \leq ||x||$ for all $x \in S$. Furthermore, $||.||_T$ is invariant under T and $I + T$.

Example 1 : Let X be a set and let S_X be the collection of all finite subsets of X. For $A, B \in S_X$ set $A + B = A \cup B$, and $||A|| = |A|$, the cardinality of A. If $T : X \to Y$ is a bijection then define $T^* : S_X \to S_Y$ by $T^*A = T(A)$. It is easy to see that T^* is an isometry.

Example 2 : Let V be a vector space and let S_V be the collection of all finite dimensional subspaces of V. For $P, Q \in S_V$ define $P + Q$ to be the smallest subspace containing P and Q, and set $||P|| = \dim(P)$. Then for Any invertible linear map $T : V \to W$, the map $T^* : S_V \to S_W$ that takes Q to $T(Q)$ is an isometry.

Example 3 : Let X be a compact topological space. An open cover α of X is said to be *saturated* if it is closed under taking open subsets, i.e., if $U \in \alpha$ and $V \subset U$ implies $V \in \alpha$. Let S_X denote the collection of all saturated open covers of X. For $\alpha, \beta \in S_X$ we define $\alpha + \beta = \alpha \cap \beta$, and for any $\alpha \in S_X$ we set $||\alpha|| = log|\alpha|$, where $|\alpha|$ is the number of elements in the smallest subcover of α. If $T : X \to Y$ is a continuous map then T induces a map $T^* : S_Y \to S_X$ by

$$T^*(\alpha) = \{T^{-1}(U) : U \in \alpha\}.$$

It is easy to check that T^* is an isometry. The entropy of T^* is called the *topological entropy* of the continuous map T and it is denoted by $h_{top}(T)$.

Suppose (X, μ) is a measure space with $\mu(X) = 1$. A Partition $P = \{P_1, \ldots, P_m\}$ of X is a finite collection of pairwise disjoint, non-empty measurable subsets of X such that $\cup P_i = X$. Let S_X be the set of all partitions of X. For $P, Q \in S_X$ we define

$$P + Q = \{P_i \cap Q_j : P_i \cap Q_j \neq \emptyset\}.$$

It is easy to see that S_X becomes an abelian semigroup and $P + P = P$ for all P. For $P \in S_X$ we set

$$||P|| = -\sum_1^m \mu(P_i)\ \log_2(\mu(P_i)).$$

Proposition 2.1 *$||P|| \leq ||P + Q|| \leq ||P|| + ||Q||$ for all finite partitions P and Q.*

Proof. Choose $P = \{P_1, \ldots, P_m\}$ and $Q = \{Q_1, \ldots, Q_n\}$ in S_X. Set $p_i = \mu(P_i),\ q_j = \mu(Q_j)$ and $r_{ij} = \mu(P_i \cap Q_j)$. Now

$$||P + Q|| - ||P|| = \sum p_i\ \text{log}p_i - \sum r_{ij}\ \text{log}r_{ij}\ = \ -\sum r_{ij}(\text{log}r_{ij} - \text{log}p_i).$$

Since log is an increasing function, this shows that $||P + Q|| \geq ||P||$. We define $\phi : [0, 1] \to \mathbb{R}$ by $\phi(0) = 0$ and $\phi(x) = -x\text{log}x$ if $x > 0$. Since

$\phi''(x) \leq 0$ in $(0,1)$, it follows that ϕ is a concave function. Put $c_{ij} = r_{ij}/p_i$ if $p_i > 0$, and zero otherwise. Observe that $||P+Q|| - ||P|| = \sum p_i \phi(c_{ij})$. Since ϕ is concave, we deduce that

$$||P+Q|| - ||P|| \leq \sum_j \phi(\sum_i p_i c_{ij}) = \sum_j \phi(q_j) = ||Q||.$$

□

Example 4 : Let (X,μ) and (Y,ν) be a measure spaces with $\mu(X) = \nu(Y) = 1$. If $T : (X,\mu) \to (Y,\nu)$ is a measure preserving map then we define a map $T^* : S_Y \to S_X$ by

$$T^*(P) = \{T^{-1}(P_1), \ldots, T^{-1}(P_n)\}.$$

It is easy to see that T^* is an isometry. If T is a measure preserving map from (X,μ) to (X,μ) then the entropy of $T^* : S_X \to S_X$ is called the *entropy* of the measure preserving map T and it is denoted by $h(T)$.

Remark. Note that in all the above examples the association $X \mapsto S_X$ is functorial. Hence for any morphism $T : X \to X$ the entropy of $T^* : S_X \to S_X$ is a conjugacy invariant of the map T. In particular, the topological entropy of a continuous map is a topological conjugacy invariant and the measure theoretic entropy of a measure preserving map is a measurable conjugacy invariant.

3 Computations of entropy

Proposition 3.1 *Let (X,d) be a compact metric space and let $T : X \to X$ be an isometry. Then $h_{top}(T) = 0$.*

Proof. Let α be a saturated open cover of X. We choose $\epsilon > 0$ such that every open set with diameter less that ϵ is contained in an element of α. Let

β denote the open cover consisting of all open sets with diameter less than ϵ. Then $\alpha + \beta = \beta$ as α is saturated and $\beta \subset \alpha$. It follows that

$$||(I + T^*)^n(\alpha)|| \leq ||(I + T^*)^n(\alpha + \beta)|| = ||(I + T^*)^n(\beta)|| \; \forall n.$$

In particular, $||\alpha||_{T^*} \leq ||\beta||_{T^*}$. On the other hand $T^*\beta = \beta$. Since S_X is idempotent, this implies that $(I + T^*)^n\beta = \beta$ for all n. Taking norm and dividing by n we see that $||\beta||_{T^*} = 0$. □

For $m \geq 1$ let X_m denote the set $\{1, 2, \ldots, m\}^{\mathbb{N}}$, equipped with the product topology. Clearly X_m is a compact topological space, and the shift map $T_m : X_m \to X_m$ is a surjective continuous map. We define a metric d on X_m by $d(x, y) = 2^{-k}$ where $k = \inf\{i : x_i \neq y_i\}$. It is easy to see that the metric d induces the product topology on X_m.

Proposition 3.2 *$h_{top}(T_m) = log\ m$.*

Proof. Let $\pi : X_m \to \{1, \ldots, m\}$ denote the map defined by $\pi(x) = x(0)$, and let α denote the open cover consisting of all U such that π restricted to U is a constant. It is easy to see that $||\alpha + \cdots T^{*k-1}\alpha|| = log(m^k) = k\ log\ m$. Hence $||\alpha||_{T^*} = log\ m$. Now suppose β is any saturated open cover of X_m. We choose $\epsilon > 0$ such that any open set U with diameter less than ϵ is contained in an element of β. We also choose n such that $2^{-n} < \epsilon$. Then every element of $\alpha + \cdots T^{*n-1}\alpha$ has diameter less than ϵ. In particular $(I + T^*)^{n-1}(\alpha) \geq \beta$. This implies that

$$||\beta||_{T^*} \leq ||(I + T^*)^{n-1}(\alpha)||_{T^*} = ||\alpha||_{T^*} = log\ m.$$

□

Our next task is to compute measure theoretic entropy of Bernoulli shifts. We begin with the following approximation lemma.

Lemma 3.3 *Suppose $(X, \mathcal{B}, \mu)$ is a probability space and suppose $\mathcal{A} \subset \mathcal{B}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{B}$. Then for any $P \in S$ and $\epsilon > 0$ there exists a partition $P_1 \subset \mathcal{A}$ such that $P \leq P_1 + Q$ for some $Q \in S$ with $||Q|| < \epsilon$.*

Proof. We first consider the case when P has only two elements, i.e., $P = \{B, B^c\}$ for some measurable set B. Note that $x\log x \mapsto 0$ as $x \mapsto 0$ or 1. Hence we can find $\delta > 0$ such that $\mu(E) < \delta$ implies $||\{E, E^c\}|| < \epsilon$. As $\sigma(\mathcal{A}) = \mathcal{B}$ we can find $A \in \mathcal{A}$ such that $\mu(F) < \delta$, where $F = A\Delta B$. Define $P_1 = \{A, A^c\}$ and $Q = \{F, F^c\}$. It is easy to see that P_1 and Q have the required properties. Now suppose $P = \{B_1, \ldots, B_n\}$. For $1 \leq i \leq n$, define $P^i = \{B_i, B_i^c\}$. Find P_1^i, Q^i as above with $||Q^i|| < \epsilon/n$, and set $P_1 = \sum P^i$, $Q = \sum Q^i$. □

We note the following consequence of the previous lemma.

Proposition 3.4 *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be a measure preserving map. Suppose $\mathcal{A}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{B}$. Then $h(T) = \sup\{||P||_{T^*} : P \subset \mathcal{A}\}$.*

Proof. Fix $\epsilon > 0$ and choose P' such that $h(T) \leq ||P'||_{T^*} + \epsilon$. Applying the previous lemma find P_1 and Q such that $P' \leq P_1 + Q$, $P_1 \subset \mathcal{A}$ and $||Q|| < \epsilon$. Since $||Q||_{T^*} \leq ||Q||$, it follows that

$$h(T) \leq ||P_1||_{T^*} + ||Q||_T + \epsilon = ||P_1||_{T^*} + 2\epsilon.$$

As ϵ is arbitrary, this proves the proposition. □

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be a measure preserving map. A partition P is said to be a *generator* if $\mathcal{B}$ is the smallest T-invariant σ-algebra containing $\{P_1, \ldots, P_n\}$.

Theorem 3.5 *If P is a generator then $h(T) = ||P||_{T^*}$.*

Proof. For any partition Q, let $A(Q)$ denote the collection of all subsets which can be expressed as union of elements of Q. It is easy to verify that $A(Q)$ is a finite algebra and $R \leq Q$ if and only if $R \subset A(Q)$. We define an algebra A_∞ by

$$A_n = A(P + T^*P + \cdots + T^{*n-1}P), \quad A_\infty = \bigcup_{n=1}^{\infty} A_n.$$

Note that A_∞ is the smallest T-invariant algebra containing P. Hence $\sigma(A_\infty) = \mathcal{B}$. If a partition Q is contained in A_∞ then $Q \subset A_n$ for some n. Hence

$$||Q||_{T^*} \leq ||\sum_{i=0}^{n-1} T^{*i}P||_{T^*} = ||(I+T^*)^{n-1}(P)||_{T^*} = ||P||_{T^*}.$$

From the previous lemma it follows that $h(T) = ||P||_{T^*}$. □

Let (X, μ) be a probability space and let $P, Q \in S$. Then P and Q are said to be *independent* if $\mu(P_i \cap Q_j) = \mu(P_i)\mu(Q_j)$ for all i and j. It is easy to see that if P and Q are independent then $||P+Q|| = ||P|| + ||Q||$.

Entropy of shifts : Let $Y = \{y_1, \ldots, y_n\}$ be a finite set and let ν be a probability measure on Y. Let $(X, \mathcal{B}, \mu) = (Y, \nu)^{\mathbb{Z}}$ and let $T : X \to X$ be the shift map. We define a partition $P = \{P_1, \ldots, P_n\}$ of X by

$$P_i = \{x \in X : x(0) = y_i\}.$$

Let $\mathcal{A}$ be the smallest T-invariant σ-algebra containing P. Since $P \subset \mathcal{A}$, the projection to 0'th coordinate is a measurable map with respect to $\mathcal{A}$. Since $\mathcal{A}$ is T-invariant, all coordinate projections are measurable with respect to $\mathcal{A}$. Hence $\mathcal{A} = \mathcal{B}$, i.e. P is a generator. We observe that for any k the partitions $P + \cdots + T^{*k-1}P$ and $T^{*k}P$ are independent. Applying induction on k we see that $||\sum_0^{k-1} T^{*i}P|| = k||P||$. Hence $h(T) = ||P||_{T^*} = ||P||$. We note that $h(T) = \log n$ in the special case when ν is the uniform measure.