Mathematical Analysis & Convex Optimization

Vol. 1 (2020), No. 1, 49–55
https:\\ maco.lu.ac.ir
DOI: 10.29252/maco.1.1.6

Research Paper

MUTUAL ENTROPY MAP FOR CONTINUOUS MAPS ON COMPACT METRIC SPACES

MEHDI RAHIMI* AND AMIR ASSARI

Abstract. In this paper we introduce the concept of mutual entropy map for continuous maps on metric spaces. It is a non-negative extended real number which depends on two measures which are preserved by a system. Then we will extract the Kolmogorov entropy of ergodic systems from the mutual entropy as a special case when the two measures are equal.

MSC(2010): 37A35.
Keywords: Entropy, invariant, mutual entropy.

1. Introduction

Entropy, as a measure of uncertainty associated with a random variable, was introduced by Shannon [14] in information theory. Then it was introduced in ergodic theory by Kolmogorov [7] and Sinai [17], which measures the rate of increase in dynamical complexity as a system evolves with time. Adler, Konheim, and Mc Andrew [1] introduced the topological entropy as an invariant of topological conjugacy and also as an analogue of measure theoretic entropy. Later, Dinaburg and Bowen [3] gave a new, but equivalent, definition of topological entropy that led to the variational principle which connected the topological entropy and measure theoretic entropy.

The concept of entropy is studied from different viewpoints [2, 9]. For example, Shannon [14], McMillan [8] and Breiman [4] gave local approaches to entropy based on the theorem of Shannon-McMillan-Breiman. Another interesting topological version of the theorem of Shannon-McMillan-Breiman was given by Brin and Katok [5]. As another example, the entropy of a continuous map on a compact metric space is considered as a linear operator in [10] and a linear functional in [11], rather than a non-negative number.

There are also many generalized forms of entropy and information [6, 12, 13, 15, 16, 18] which are applied in other areas of science.

In information theory, the Shannon information, $H(X)$, is defined for a random variable X. This concept is generalized to the "mutual information", $I(X,Y)$, which depends on two random variables X and Y. The mutual information and Shannon information coincide when $X = Y$, i.e., $I(X,X) = H(X)$.

In this paper, motivating from the mutual information, we define the concept of *mutual entropy*, $\Gamma_T(\mu,\nu)$, for continuous systems on compact metric spaces. The definition of mutual entropy depends on two invariant measures, regardless to the definition of Kolmogorov

Date: Received: December 13, 2019, Accepted: April 4, 2020.
*Corresponding author.

entropy. We will show that the two concepts coincide when $\mu = \nu$. This also presents a local approach to the Kolmogorov entropy of dynamical systems.

In section 2, we first recall some preliminary facts on entropy. In section 3, we will introduce the concept of mutual entropy map and will prove some of its properties. Finally, in section 4, we will extract the Kolmogorov entropy from the mutual entropy map as a special case.

The family of all Borel probability measures on X is denoted by $M(X)$, and the family of all Borel probability measures on X that is preserved by T is denoted by $M(X,T)$. We also write $E(X,T)$ for the collection of ergodic measures of T.

2. Preliminary facts

Kolmogorov and Sinai introduced the measure theoretic entropy $h_\mu(T)$ for measure-preserving dynamical systems. Let $T : (X, \mathfrak{B}, \mu) \to (X, \mathfrak{B}, \mu)$ be a measure-preserving map on the probability space $(X, \mathfrak{B}, \mu)$. The measure theoretic entropy is defined as follows [19]:

$$h_\mu(T) = \sup_\xi \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} T^{-i}\xi).$$

Here the supremum is taken over all partitions and $\bigvee_{i=0}^{n-1} T^{-i}\xi$ is the partition generated by events of n successive observations.

A topological analogue of the measure theoretic entropy was introduced by Adler, Konheim and McAndrew [1] in 1965. The topological entropy of a continuous map $T : X \to X$ on the compact space X is defined as follows:

$$h_{top}(T) = \sup_\alpha \lim_{n\to\infty} \frac{1}{n} \log N(\bigvee_{i=0}^{n-1} T^{-i}\alpha)$$

where the supremum is taken over all open covers of X, and $N(\alpha)$ denotes the number of sets in a finite subcover of α with smallest cardinality.

In 1948 Shannon, in a famous work [14] which originated information theory, introduced a notion of entropy applicable to shifts, which is essentially the same as Kolmogorov's. It led to an approach to the concept of entropy, based on the theorem of Shannon-McMillan-Breiman. An interesting topological version of the theorem of Shannon-McMillan-Briman was given by Brin and Katok [5]. Let (X, d) be a compact metric space, $\mathfrak{B}$ the Borel $\sigma-$algebra of X and μ a probability measure on $\mathfrak{B}$. Let $T : (X, d) \to (X, d)$ be a continuous map preserving μ. For $n \in \mathbb{N}$ the metric d_n is defined on X as follows:

$$d_n(x, y) := \max_{0\le j\le n} d(T^j(x), T^j(y)).$$

For $\epsilon > 0$ and $x \in X$, the $\epsilon-$ball of d_n centered at x is denoted by $B_n(x, \epsilon)$. In other words

$$B_n(x, \epsilon) = \{y \in X : \ d(T^j(y), T^j(x)) \le \epsilon, \ 0 \le j \le n\}.$$

Definition 2.1. Suppose that $T : (X, d) \to (X, d)$ is a continuous map on the compact metric space (X, d). Define

$$h^+_\mu(T, x, \epsilon) := -\limsup_{n\to\infty} \frac{1}{n} \log \mu(B_n(x, \epsilon)),$$

$$h^-_\mu(T, x, \epsilon) := -\liminf_{n\to\infty} \frac{1}{n} \log \mu(B_n(x, \epsilon)).$$

Brin and Katok [5] proved the following theorem and gave a local approach to measure theoretic entropy.

Theorem 2.2. *Suppose that $T : (X, d) \to (X, d)$ is a continuous map on the compact metric space (X, d) and $\mu \in M(X, T)$, then*

(i) $\lim_{\epsilon\to 0} h^+_\mu(T, x, \epsilon) = \lim_{\epsilon\to 0} h^-_\mu(T, x, \epsilon) := h_\mu(T, x)$ *for almost every x in X;;*
(ii) $h_\mu(T, x)$ *is $T-$invariant;*
(iii) $\int_X h_\mu(T, x)d\mu(x) = h_\mu(T)$, *where $h_\mu(T)$ is the measure theoretic (Kolmogorov) entropy.*

3. **Mutual entropy map**

In this section, we introduce the concept of *mutual entropy map* for continuous maps on metric spaces. Suppose that $T : (X, d) \to (X, d)$ is a continuous map on the metric space (X, d). Let, for $n \in \mathbb{N}$, $\epsilon > 0$ and $x \in X$, $B_n(x, \epsilon)$ be as in section 2.

Definition 3.1. Suppose that $T : (X, d) \to (X, d)$ is continuous. For $n \in \mathbb{N}$, $\epsilon > 0$ and $x, y \in X$ define

$$\pi_n(T, x, y; \epsilon) := \limsup_{p\to\infty} \frac{1}{p}\text{card}(\{0 \le j \le p-1 :\ T^j(y) \in B_n(x, \epsilon)\}).$$

$\pi_n(T, x, y, \epsilon)$ is called the $(n, \epsilon)-$tendency of y with respect to x.

Definition 3.2. Suppose that $T : (X, d) \to (X, d)$ is continuous. For $n \in \mathbb{N}$, $\epsilon > 0$ and $x, y \in X$ define

$$\varphi_n(T, x, y; \epsilon) := \begin{cases} -\frac{1}{n}\log \pi_n(T, x, y; \epsilon) & \text{if } \pi_n(T, x, y; \epsilon) \neq 0 \\ 0 & \text{if } \pi_n(T, x, y; \epsilon) = 0 \end{cases}$$

Definition 3.3. Suppose that $T : (X, d) \to (X, d)$ is continuous, $\epsilon > 0$, and $x, y \in X$. The $\epsilon-$relative entropy of y with respect to x is defined as follows:

$$\mathcal{J}(T, x, y; \epsilon) := \limsup_{n\to\infty} \varphi_n(T, x, y; \epsilon).$$

It is easy to see that, if $\epsilon_1 < \epsilon_2$ then $\mathcal{J}(T, x, y; \epsilon_1) \ge \mathcal{J}(T, x, y; \epsilon_2)$; therefore, the limit $\lim_{\epsilon\to 0} \mathcal{J}(T, x, y; \epsilon)$ exists as an extended real number.

Definition 3.4. Suppose that $T : (X, d) \to (X, d)$ is continuous, and $x, y \in X$. The relative entropy of y with respect to x (under the system T) is defined as follows:

$$\mathcal{J}_T(x, y) := \lim_{\epsilon\to 0} \mathcal{J}(T, x, y; \epsilon).$$

The map $X \times X \to [0, \infty]$ given by $(x, y) \to \mathcal{J}_T(x, y)$ is called the *relative entropy map* of T.

Remark: Since the map $(x, y) \to d_n(T^j(y), x)$ is continuous for all j and n, then one can easily see that the map $(x, y) \to \mathcal{J}_T(x, y)$ is Borel measurable and hence

$$\int_{X\times X} \mathcal{J}_T(x, y)d(\mu \times \nu)(x, y)$$

is well-defined.

Definition 3.5. Suppose that $T : X \to X$ is continuous. The map

$$\Gamma_T : M(X,T) \times M(X,T) \to [0,\infty]$$

defined by

$$\Gamma_T(\mu,\nu) := \int_{X\times X} \mathcal{J}_T d\mu \times \nu$$

is called the *mutual entropy map of* T.

It is obvious from the definition that:

(i) Γ_T is affine with respect to each variable, i.e.,

$$\Gamma_T(\sum_{i=1}^{n} \lambda_i \mu_i, \nu) = \sum_{i=1}^{n} \lambda_i \Gamma_T(\mu_i, \nu)$$

and

$$\Gamma_T(\mu, \sum_{i=1}^{n} \lambda_i \nu_i) = \sum_{i=1}^{n} \lambda_i \Gamma_T(\mu, \nu_i)$$

where $\lambda_i \geq 0$ and $\sum_{i=1}^{n} \lambda_i = 1$;

(ii) Γ_T is symmetric, i.e., $\Gamma_T(\mu,\nu) = \Gamma_T(\nu,\mu)$ for all $\mu,\nu \in M(X,T)$.

We recall that, two continuous maps $T_1 : X_1 \to X_2$ and $T_2 : X_2 \to X_2$ are said to be topologically conjugate, if there is a homeomorphism $\phi : X_1 \to X_2$ such that $\phi T_1 = T_2 \phi$. For example, one may easily see that, the maps $T_1, T_2 : [0,1] \to [0,1]$ defined by

$$T_1(x) = \begin{cases} 2x & \text{if } x < \frac{1}{2} \\ 2-2x & \text{if } x \geq \frac{1}{2} \end{cases}$$

and $T_2(x) = 4x(1-x)$ are topologically conjugate via $\phi(x) = \sin^2(\frac{\pi x}{2})$.

Theorem 3.6. *Suppose that $T_1 : X_1 \to X_1$ and $T_2 : X_2 \to X_2$ are topological conjugate continuous maps via the homeomorphism $\phi : X_1 \to X_2$. Then for $\mu,\nu \in M(X_1,T_1)$ we have*

$$\Gamma_{T_1}(\mu,\nu) = \Gamma_{T_2}(\mu\phi^{-1}, \nu\phi^{-1}).$$

Proof. Fix $x, y \in X_1$. Let $\epsilon > 0$ and $n \in \mathbb{N}$. Since ϕ is a homeomorphism, then there exists $0 < \delta < \epsilon$ such that

$$\phi(B_n(x,\delta)) \subset B_n(\phi(x),\epsilon).$$

Let $p \in \mathbb{N}$. For $0 \leq j \leq p-1$, if $T_1^j(y) \in B_n(x,\delta)$ then $\phi T_1^j(y) \in \phi(B_n(x,\delta)) \subset B_n(\phi(x),\epsilon)$. Since $\phi T_1 = T_2 \phi$ then $T_2^j \phi(y) \in B_n(\phi(x),\epsilon)$ and this is equivalent to

$$\{0 \leq j \leq p-1 :\ T_1^j(y) \in B_n(x,\delta)\} \subset \{0 \leq j \leq p-1 :\ T_2^j \phi(y) \in B_n(\phi(x),\epsilon)\}$$

hence

$$\pi_n(T_1,x,y;\delta) \leq \pi_n(T_2,\phi(x),\phi(y);\epsilon).$$

By taking the logarithm, dividing by n and letting $n \to +\infty$ we will have

$$\begin{aligned} \mathcal{J}_{T_1}(x,y;\delta) &= \limsup_{n\to\infty} \varphi_n(T_1,x,y;\delta) \\ &\geq \limsup_{n\to\infty} \varphi_n(T_2,\phi(x),\phi(y);\epsilon) \\ &= \mathcal{J}(T_2,\phi(x),\phi(y);\epsilon). \end{aligned}$$

By taking the limit in above inequality as $\epsilon \to 0$, we get

$$\mathcal{J}_{T_1}(x,y) \geq \mathcal{J}_{T_2}(\phi(x),\phi(y)).$$

Since ϕ is a homeomorphism then

$$\mathcal{J}_{T_1}(x,y) = \mathcal{J}_{T_2}(\phi(x),\phi(y)).$$

Finally, for $\mu,\nu \in M(X_1,T_1)$ we have

$$\begin{aligned}
\Gamma_{T_1}(\mu,\nu) &= \int_{X_1\times X_1} \mathcal{J}_{T_1}(x,y)d(\mu\times\nu)(x,y)\\
&= \int_{X_1}\left(\int_{X_1}\mathcal{J}_{T_1}(x,y)d\mu(x)\right)d\nu(y)\\
&= \int_{X_1}\left(\int_{X_1}\mathcal{J}_{T_2}(\phi(x),\phi(y))d\mu(x)\right)d\nu(y)\\
&= \int_{X_1}\left(\int_{X_2}\mathcal{J}_{T_2}(x,\phi(y))d\mu\phi^{-1}(x)\right)d\nu(y)\\
&= \int_{X_2}\left(\int_{X_2}\mathcal{J}_{T_2}(x,y)d\mu\phi^{-1}(x)\right)d\nu\phi^{-1}(y)\\
&= \int_{X_2\times X_2}\mathcal{J}_{T_2}(x,y)d(\mu\phi^{-1}\times\nu\phi^{-1})(x,y)\\
&= \Gamma_{T_2}(\mu\phi^{-1},\nu\phi^{-1}). \quad \square
\end{aligned}$$

Recall that, two metrics d_1 and d_2 on X are uniformly equivalent if $id:(X,d_1)\to(X,d_2)$ and $id:(X,d_2)\to(X,d_1)$ are both uniformly continuous. Note that, seemingly, the definition of $B_n(x,\epsilon)$, $\pi_n(T,x,y;\epsilon)$ and therefore $\mathcal{J}_T(x,y)$ depends on the metric. The following theorem states that, this is not the case.

Theorem 3.7. *Let d_1 and d_2 be uniformly equivalent metrics on X. Let also $\mathcal{J}_T^{(1)}$ and $\mathcal{J}_T^{(2)}$ be the corresponding relative entropy maps, then $\mathcal{J}_T^{(1)} = \mathcal{J}_T^{(2)}$.*

Proof. Suppose that

$$B_n^{(k)}(x,\epsilon) := \{y\in X : d_k(T^i(y),T^i(x)) < \epsilon, 1\leq i\leq n\}$$

and

$$\pi_n^{(k)}(x,y;\epsilon) := \limsup_{p\to\infty}\frac{1}{p}\mathrm{card}(\{0\leq j\leq p-1 : T^j(y)\in B_n^{(k)}(x,\epsilon)\}) \quad (k=1,2).$$

For given $\epsilon>0$, by uniform continuity of $id:(X,d_1)\to(X,d_2)$, choose $0<\delta<\epsilon$ such that

$$\forall x,y\in X: \quad d_1(x,y)<\delta \Rightarrow d_2(x,y)<\epsilon.$$

This easily indicates that $B_n^{(1)}(x,\delta)\subset B_n^{(2)}(x,\epsilon)$ for all $n\in\mathbb{N}$, so

$$\pi_n^{(1)}(x,y;\delta)\leq\pi_n^{(2)}(x,y;\epsilon)$$

and consequently

$$\varphi_n^{(1)}(x,y;\delta)\geq\varphi_n^{(2)}(x,y;\delta).$$

This easily results in $\mathcal{J}_T^{(1)}(x,y)\geq\mathcal{J}_T^{(2)}(x,y)$. Applying the uniform continuity of $id:(X,d_2)\to(X,d_1)$ and repeating the previous procedure, we will have the converse inequality.

□

As an example of equivalent metrics, let d_1 be the standard metric $d_1(x,y) = |x-y|$ on $[0,1]$. Let also, d_2 be the metric on $[0,1]$ defined by $d_2(x,y) = |e^x - e^y|$. Then d_1 and d_2 are uniformly equivalent metrics on $[0,1]$. So, for any continuous map $T : [0,1] \to [0,1]$, the relative entropy maps of T, corresponding to d_1 and d_2, coincide.

4. **Mutual entropy and Kolmogorov entropy**

In this section, we will extract the Kolmogorov entropy from the mutual entropy as a special case. It also gives a local approach to the Kolmogorov entropy. In the rest of this section $T : X \to X$ is a continuous map on the compact metric space X.

Now we are in a position to extract the Kolmogorov entropy from the mutual entropy for ergodic systems. The following theorem states that the mutual entropy coincides to the Kolmogorov entropy when $\mu = \nu$.

Theorem 4.1. *Suppose that $T : (X,d) \to (X,d)$ is a continuous map on the compact metric space (X,d). Then*

$$\Gamma_T(\mu,\mu) = h_\mu(T)$$

for all ergodic measures μ.

Proof. Let μ be an ergodic measure. Let $x, y \in X$ be given. Let also $\{\epsilon_n\}_{n\geq 1}$ be a sequence of positive numbers such that $\epsilon_n \to 0$. For $m, n \in \mathbb{N}$ we have

$$\begin{aligned}\pi_m(T,x,y;\epsilon_n) &= \limsup_{p\to\infty} \frac{1}{p}\mathrm{card}(\{0 \leq j \leq p-1 :\ T^j(y) \in B_m(x,\epsilon_n)\}) \\ &= \limsup_{p\to\infty} \frac{1}{p}\sum_{j=0}^{p-1} \chi_{B_m(x,\epsilon_n)}(T^j(y)).\end{aligned}$$

Since μ is ergodic then

$$\limsup_{p\to\infty} \frac{1}{p}\sum_{j=0}^{p-1} \chi_{B_m(x,\epsilon_n)}(T^j(y)) = \int_X \chi_{B_m(x,\epsilon_n)}(y)d\nu(y) = \mu(B_m(x,\epsilon_n))$$

for almost all $y \in X$, i.e., there is Borel measurable subset $A_{m,n,x}$ of X such that $\nu(A_{m,n,x}) = 1$ and

$$\pi_m(T,x,y;\epsilon_n) = \mu(B_m(x,\epsilon_n)) \tag{4.1}$$

for all $y \in A_{m,n,x}$. Let $A_x := \bigcap_{m,n=1}^{\infty} A_{m,n,x}$; then $\mu(A_x) = 1$, and (4.1) holds for all $y \in A_x$. Taking logarithm, dividing by m and letting $m \to +\infty$ and then $n \to +\infty$, in above equality, we conclude that:

$$\mathcal{J}_T(x,y) = h_\mu(T,x)$$

for all $y \in A_x$. Integrating both sides of the previous equality with respect to y, we have:

$$\int_X \mathcal{J}_T(x,y)d\mu(y) = \int_{A_x} \mathcal{J}_T(x,y)d\mu(y) = \int_X h_\mu(T,x)d\mu(y) = h_\mu(T,x). \tag{4.2}$$

Finally, the result follows by integrating the previous equality with respect to x. □

5. Concluding remarks

In this paper, we introduced the concept of mutual entropy $\Gamma_T(\mu, \nu)$ for a continuous map on a compact metric space. This is a non-negative quantity which is invariant under topological conjugacy. When $\mu = \nu$ we get the Kolmogorov entropy as a special case. The definition of the mutual entropy depends on the function $\mathcal{J}_T : X \times X \to [0, \infty]$ where the integral of $\mathcal{J}_T$ with respect to $\mu \times \mu$ equals the Kolmogorov entropy. It is also proved that, the introduced quantity is independent of the choice of metric as long as the topology does not change.

Acknowledgments

The authors would like to thank the referees for their comprehensive and useful comments which helped the improvement of this work to the present form.

References

[1] R. L. Adler, A. G. Konheim, and M. H. McAndrew, Topological entropy. *Trans. Amer. Math. Soc.*, **114**: 309–319, 1965.
[2] M. Amini, Entropy of infinite systems and transformations. *Int. J. Nonlinear Anal. Appl.*, **10**(1): 27–33, 2019.
[3] R. Bowen, Invariant measures for Markov maps of the interval. *Comm. Math. Physics*, **69**: 1-–17, 1979.
[4] L. Breiman, The individual theorem of information theory. *Ann. Math. Statist.*, **3**: 809-–811, 1957.
[5] M. Brin, A. Katok, *On local entropy in geometric dynamics*, volume 30 of *Lecture Notes in Mathematics*. New York, Springer-Verlag, 30–38, 1983.
[6] J. N. Kapur, Generalized Entropy of Order α and Type β. *The Math. Seminar*, **4**: 78–94, 1967.
[7] A. N. Kolmogorov, New metric invariant of transitive dynamical systems and endomorphisms of Lebesgue spaces. *Doklady of Russian Academy of Sciences*, **119**(5): 861–864, 1958.
[8] B. McMillan, The basic theorems of information theory. *Ann. of Math. Statistics*, **24**: 196–219, 1953.
[9] U. Mohammadi, Observational Modeling of the Kolmogorov-Sinai Entropy. *Sahand Communications in Mathematical Analysis*, **13**(1): 101-114, 2019.
[10] M. Rahimi, A. Riazi, Entropy operator for continuous dynamical systems of finite topological entropy. *Bulletin of the Iranian Mathematical Society*, **38**(4): 883-892, 2012.
[11] M. Rahimi, A. Riazi, Entropy functional for continuous systems of finite entropy. *Acta Mathematica Scientia*, **32B**(2): 775-782, 2012.
[12] P. N. Rathie, On a Generalized Entropy and a Coding Theorem. *J. Appl. Probl*, **7**: 124–133, 1970.
[13] A. Réyni, On Measures of Entropy and Information. *Proc. 4th Berk. Symp. Math Statist. and Probl., University of California Press*, **1**: 547–461, 1961.
[14] C. Shannon, A mathematical theory of communication. *Bell Syst. Tech. Journal*, **27** : 379–423, 1948.
[15] B. D. Sharma, and I. J. Taneja, Entropy of Type (α, β) and Other Generalized Additive Measures in Information Theory. *Metrika*, **22**: 205–215, 1975.
[16] B. D. Sharma, and I. J. Taneja, Three Generalized Additive Measures of Entropy. *Elec. Inform. Kybern*, **13**: 419–433, 1977.
[17] Ya. G. Sinai, On the notion of entropy of a dynamical system. *Doklady of Russian Academy of Sciences*, **124**: 768–771, 1959.
[18] R. S. Varma, Generalizations of Reyni's Entropy of Order α. *J. Math. Sci.*, **1**: 34–48, 1966.
[19] P. Walters, *An introduction to ergodic theory*, Springer-Verlag, 1982.

(Mehdi Rahimi) Department of Mathematics, Faculty of Science, University of Qom, Qom, Iran.
Email address: m10.rahimi@gmail.com

(Amir Assari) Department of Basic Science, Jundi-Shapur University of Technology, Dezful, Iran.
Email address: amirassari@jsu.ac.ir