On the entropy of symbolic image of a dynamical system

Osipenko G. S., Ampilova N.B.

1 Symbolic image of a dynamical system

Let $f : M \to M$ be a homeomorphism of a compact manifold M generating a discrete dynamical system

$$x_{n+1} = f(x_n), \tag{1}$$

and $\rho(x, y)$ be a distance on M. In what follows we use the concept of symbolic image of a dynamical system [17], which brings together symbolic dynamics [3, 14] and numerical methods [10]. Let $C = \{M(1), ..., M(n)\}$ be a finite closed covering of a manifold M. The set $M(i)$ is called cell with index i.

Definition. 1 *[16] Symbolic image of the dynamical system (1) for a covering C is an oriented graph G with vertices $\{i\}$corresponding to cells $\{M(i)\}$. The vertices i and j are connected by the edge $i \to j$ iff*

$$f(M(i)) \bigcap M(j) \neq \emptyset.$$

Symbolic image is a tool for a space discretization and graphic representation of the dynamic of a system under study, which allows the obtaining useful information about the global structure of the system dynamics. Symbolic image depends on a covering C. The existence of an edge $i \to j$ guaranties the existence of a point $x \in M(i)$ such that $f(x) \in M(j)$. In other words, an edge $i \to j$ is the trace of the mapping $x \to f(x)$, where $x \in M(i),\ f(x) \in M(j)$. If there isn't an edge $i \to j$ on G then there are not the points $x \in M(i)$ such that $f(x) \in M(j)$.

We do not place special restrictions on a covering C, but basing on the theorem about the triangulation of a compact manifold [19] we may without loss of generality assume that cells $M(i)$ are polyhedrons intesecting on their boundary disks. In practice M is a compact in R^d, and $M(i)$ are cubes or parallelepipeds. Let C be a covering of M by polyhedrons intersecting on their boundary disks. In what follows we also use a measurable partition C^*, such that a boundary disk belongs only one of adjoining cells. We assume that cells-polyhedrons are closures of their interiors.

Definition. 2 *A vertex of a symbolic image G is said to be recurrent if there is a periodic path passing through it. The set of recurrent vertices is denoted by RV. The recurrent vertices i and j are called equivalent if there exists a periodic path passing through i and j.*

Thus, the set of recurrent vertices RV is split into equivalence classes $\{H_k\}$. In the graph theory such classes are called strong connectivity components.

Let

$$diam\ M(i) = \max(\rho(x, y) : x, y \in M(i))$$

be the diameter of a cell $M(i)$ and $d = diam(C)$ be the maximum of the diameters of the cells. The number d is called the diameter of the covering C.

An oriented graph G is uniquely defined by its adjacency matrix (matrix of admissible transitions) Π. The matrix $\Pi = (\pi_{ij})$ has sizes $n \times n$, where n — is the number of vertices of G, and $\pi_{ij} = 1$ iff there exists the edge $i \to j$, else $\pi_{ij} = 0$. Hence an i-th row in Π corresponds to the vertex i (cell $M(i)$), and on the place j in this row there is 1 or 0 depending on the existence (or nonexistence) of nonempty intersection $f(M(i))$ and $M(j)$. Matrix of admissible transitions depends on the numbering of vertices (cells of the covering), so that a change of numeration leads to a change of matrix Π. Note that there exists a numeration transforming the matrix of admissible transitions to a canonical form.

Proposition. 1 *[1] Vertices of a symbolic image G may be numbered such that* *the adjacency matrix has the form*

$$\Pi = \begin{pmatrix} (\Pi_1) & \cdots & \cdots & \cdots & \cdots \\ & \ddots & & & \\ 0 & & (\Pi_k) & \cdots & \cdots \\ & \ddots & & \ddots & \\ 0 & & 0 & & (\Pi_s) \end{pmatrix},$$

where every diagonal block Π_k *corresponds either an equivalence class* H_k *of recurrent vertices or a nonrecurrent vertex and consists of one zero. Under diagonal blocks are only zeroes (upper triangular matrix)*

2 Entropy

In 1865 R. Clausius [5] introduced the most important in thermodynamics concept entropy. To explain the irreversibility of macroscopic states L. Boltzmann in 1872 [4] first introduced statistical approach in thermodynamics : he proposed to describe a state of a system by using its microstates. The Boltzmann entropy S is statistical entropy for the equiprobable distribution of a system over P states, it is defined as

$S = k \log(P)$, where k is the Boltzmann constant.

In 1948 C. Shannon [20, 21] introduced the notion of capacity (C) for an information channel as follows

$$C = \lim_{T\to\infty} \frac{log_2 N(T)}{T},$$

where $N(T)$ is the number of admissible signals for the time T. He also defined information entropy as follows

$$h = -\sum_i p_i \log_2 p_i,$$

where p_i is a probability of i-th signal (message), $i \in 1, ...n$, and n is the number of signals. A. N. Kolmogorov in 1958 [12] introduced entropy in the theory of dynamical systems. Entropy is a fine invariant of a dynamical system, it may be interpreted as a measure of the system chaoticity. Comprehensive information on entropy in dynamical systems is given in [7, 11]. It turns out that entropy characteristics may be obtained both for a system described analytically and for its phase portraits. The application of such characteristics to digital image analysis is given in [2].

Motivation Consider a discrete dynamical system $x_{n+1} = f(x_n)$ on a compact manifold M, where $f : M \to M$ is a homeomorphism. Let $C = \{M(1), .., M(n)\}$ be a finite covering of M and the sequence $\{x_k = f^k(x), k = 0, \dots N-1\}$ be the N-length part of the trajectory of a point x. The covering C generates a coding of this part via a finite sequence $\xi(x) = \{i_k, k = 0, \dots N-1\}$, where $x_k \in M(i_k)$. In other words, i_k is the number of the cell from C which contains the point $x_k = f^k(x)$. Generally speaking, the mapping $x_k \to i_k$ is multivalued. The sequence $\xi = \{i_k\}$ is said to be (admissible) encoding of the trajectory $\{x_k = f^k(x)\}$ with respect to the covering C. Assume that we know all admissible N-length encodings, and there is a need to predict subsequent p-length encodings, i.e. to find admissible encodings of length $N + p$.

Let the number of admissible encodings $K(N)$ grows exponentially depending on N. We estimate the rate of growth of encodings by the number

$$h = \lim_{N \to +\infty} \frac{\log_b K(N)}{N}, \quad (2)$$

where b may be any real number greater than 1. The bases $b = 2$ (following to Shannon) or $b = e$ are in common use. The existence of the limit in (2) follows from the Polya lemma [1, 3, 14].

Lemma. 1 *[14], p. 103. If a sequence of non-negative numbers a_n satisfies the inequality*

$$a_{n+m} \le a_n + a_m,$$

then there exists $\lim_{n\to\infty} a_n/n$.

For the number of admissible encodings we have

$$K(n + m) \le K(n)K(m),$$

hence for the sequence $a_n = \log_b K(n)$ there exists the limit (2).
Thus, for the number of encodings $K(N)$ we obtain the estimation

$$K(N) \sim Bb^{hN},$$

where B is a constant. If $h \neq 0$ then

$$\frac{K(N+p)}{K(N)} \sim b^{hp}.$$

This relation means that for any N the uncertainty of future encodings grows with the exponent hp regardless the knowledge of previous encodings.
If the growth of the number of different encodings is not exponential (i. e. $h = 0$), for example as

$$K(N) \sim BN^A,$$

where A —- a positive number (may be large), then

$$\frac{K(N+p)}{K(N)} \sim (1 + \frac{p}{N})^A \to 1,$$

when $N \to \infty$. In other words, the uncertainty of the future decreases when the length N of known encodings increases.

Thus, if the growth of the number of different encodings is exponential, the uncertainty does not depend on N, in other case it decreases as N increases. Value h may be interpreted as a characteristics of uncertainty (chaoticity) of the system dynamic considered .

Topological entropy. Let f be a continuous mapping defined on a manifold M and $C = \{M(1), \dots, M(n)\}$ be an open covering of M. For an integer positive number N consider a sequence

$$\omega = \omega_1 \omega_2 \cdots \omega_N,$$

where ω_k is a number from 1 to n. Construct the intersection of the form

$$M(\omega) = M(\omega_1) \cap f^{-1}(M(\omega_2)) \cap \cdots \cap f^{-N+1}(M(\omega_N)), (*)$$

which is an open set. The admissible encoding ω corresponds to the nonempty intersection $M(\omega)$, i.e. there exists $x \in M(\omega_1)$ such that $f^k(x) \in M(\omega_{k+1})$.

The sequence ω codes the segment of the trajectory $\{f^k(x),\ k=1,\ 2,\ \cdots,\ N\}$. Consider all the admissible N-lentgh encodings $\{\omega\}$ and the collection of sets $C^N = \{M(\omega)\}$, which is an open covering. Choose in C^N a minimal by the number of elements (denoted further by $|C_N|$) finite subcovering C_N.

Then according to the Polya lemma there exists the limit

$$H(C) = \lim_{N\to\infty} \frac{\log |C_N|}{N}. \tag{3}$$

Definition. 3 *The number*

$$h(f) = \sup_C H(C),$$

where sup is taken over all open coverings C, is called topological entropy of the mapping $f : M \to M$.

It is easy to see that there is little point in using this definition for practical calculation of entropy. Consider some methods for entropy calculation.

A covering C_2 is said to be refined in a covering C_1, if any $A \in C_2$ lies in a set $B \in C_1$. A sequence of open coverings C_n is called exhaustive if for any open covering C there exists the number n^* such that the covering C_n is refined in C for $n \geq n^*$.

Proposition. 2 *[1] p. 122.*

1. *If C_n is a sequence of open coverings with diameters*

$$d_n = \max_{A\in C_n} diam A$$

   tending to zero, then C_n is an exhaustive sequence.

2. *Entropy of the mapping f is calculated as follows*

$$h(f) = \lim_{n\to\infty} H(C_n).$$

Consider coverings C_1 and C_2, and construct for each of them nonempty intersections (*). Denote the obtained collections of sets by C_1^N and C_2^N respectively. In each collection choose the minimal (by the number of elements) subcovering, and denote them C_{N1} and C_{N2}.

Proposition. 3 *If C_2 is refined in C_1, then*

$$|C_{N1}| \leq |C_{N2}|,$$

where $|C_{Ni}|$ denotes the number of elements in the set considered.

Proof. If $A_1 \subset B_1$ and $A_2 \subset B_2$, then $A_1 \cap f^{-1}(A_2)$ lies in $B_1 \cap f^{-1}(B_2)$. By the same way one can prove that elements of C_2^N are in corresponding elements of C_1^N. Consider C_{N1} and C_{N2}. Take from C_1^N all the elements which contain corresponding elements of C_{N2} and form from them the covering C_{N1}^*.

Then $|C_{N1}| \leq |C_{N1}^*| \leq |C_{N2}|$, because C_{N1} is a minimal subcovering for C_1^N. The proposition is proved.

Corollary. 1 *Assume that C_2 is refined in C_1, and the numbers $H(C_1)$, $H(C_2)$ are calculated by (3). Then*

$$H(C_1) \leq H(C_2).$$

3 Entropy of a symbolic image

Let G be a graph with the adjacency matrix Π. Denote by b_n the number of admissible n-length paths on G .

Definition. 4 *The number*

$$h(G) = \lim_{n\to\infty} \frac{\ln b_n}{n}$$

is called the entropy of the graph G.

Remember that the element $(\Pi^n)_{ij}$ is equal to the number of admissible n-length paths from i to j. Then

$$b_n = \sum_{ij} (\Pi^n)_{ij}$$

Theorem. 1 *Let C_n be a sequence of subcoverings of a closed covering, such that cells are polyhedrons intersecting on boundary disks, and diameters d_n of subcoverings tend to zero. Denote by G_n the symbolic images constructed for a* *mapping $f : M \to M$ in accordance with the sequence C_n. Then for the entropy of f the following inequality holds*

$$h(f) \leq \lim_{n\to\infty} h(G_n).$$

Proof. Let $C_n = \{M(i)\}$be a closed covering from the sequence described above and G_n be a symbolic image of the mapping f constructed according to C_n. Consider the set P_N of encodings of N-length segments of trajectories. It is obvious that P_N not greater than the number of admissible N-length paths on G_n,denoted by b_N.

Hence the number $|C^N|$ of nonempty intersections of the form

$$M(\omega) = M(\omega_1) \cap f^{-1}(M(\omega_2)) \cap \cdots \cap f^{-N+1}(M(\omega_N))$$

is not greater than b_N. Thus,

$$H(C_n) \leq h(G_n).$$

If C is an open covering then there exists the number n^* such that the covering C_n is refined in C for $n \geq n^*$. Then in accordance with Proposition 2 we have

$$H(C) \leq H(C_n) \leq h(G_n).$$

Now consider an exhausting sequence of open coverings $\{\widetilde{C}_m\}$. Let $n(m)$ be the number n^* constructed for the covering $\widetilde{C}_m$, then

$$H(\widetilde{C}_m) \leq H(C_n) \leq h(G_n),$$

where $n \geq n(m)$. If $m \to \infty$, then according to Proposition 2 we have

$$h(f) = \lim_{m\to\infty} H(\widetilde{C}_m) \leq \lim_{n\to\infty} h(G_n).$$

The theorem is proved.

Remember that a matrix $A(n \times n)$ is called decomposable if it admits an invariant subspace with dimension less than n, and a matrix A is called nonnegative (positive) if it has nonnegative (positive) elements. If for a nonnegative matrix A there is an integer $s > 0$ such that all the elements of A^s are positive, then A is called primitive. In particular the matrix of admissible transitions Π is nonnegative. It is nondecomposable if the symbolic image consists of one class of equivalent recurrent vertices.

Theorem. 2 *(Perron-Frobenius) [8, 14]*

- *If A is a decomposable nonnegative matrix then it has an eigenvector e with positive coordinates and the eigenvalue λ with multiplicity 1, and the other eigenvalues μ satisfy the inequality $|\mu| \leq \lambda$.*
- *If A is a decomposable nonnegative matrix and $|\mu| < \lambda$, then A is primitive.*
- *If A is a decomposable nonnegative matrix and it has $h > 1$ eigenvalues which are equal in modulus λ, then A is not primitive, and by an agreed renumeration of rows and columns it may be transformed to the form*

$$\begin{pmatrix} 0 & A_{12} & 0 & \cdots & 0 \\ 0 & 0 & A_{23} & \cdots & 0 \\ & \ddots & & \ddots & \\ 0 & 0 & 0 & \cdots & A_{h-1h} \\ A_{h1} & 0 & 0 & \cdots & 0 \end{pmatrix},$$

where A_{ij} are square blocks, and A^h consists of h primitive blocks.

Theorem. 3 *The entropy of the graph G is equal to the logarithm of the maximal eigenvalue of the adjacency matrix*

$$h(G) = \ln \lambda.$$

Proof. 1. Consider the case when G consists of one class of equivalent recurrent vertices. Let e be a positive eigenvector for the maximal eigenvalue λ,i.e.

$$\Pi e = \lambda e.$$

In the coordinate form we have

$$\sum_j (\Pi^n)_{ij} e_j = \lambda^n e_i. \tag{4}$$

Let $c = \min\{e_i\}$ and $d = \max\{e_i\}$. In accordance with the Perron-Frobenius theorem $c > 0$. Then the following inequalities hold

$$c \sum_j (\Pi^n)_{ij} \le \sum_j (\Pi^n)_{ij} e_j \le d\lambda^n.$$

It follows that

$$\sum_j (\Pi^n)_{ij} \le \frac{d}{c} \lambda^n$$

for any i. Summing by i we obtain

$$b_n = \sum_{ij} (\Pi^n)_{ij} \le \frac{dr}{c} \lambda^n,$$

where r is the number of rows in the matrix Π. It follows from (4) that

$$c\lambda^n \le \lambda^n e_i = \sum_j (\Pi^n)_{ij} e_j \le d \sum_j (\Pi^n)_{ij} \le d \sum_{ij} (\Pi^n)_{ij}.$$

Hence we obtain the estimation

$$\frac{c}{d} \lambda^n \le \sum_{ij} (\Pi^n)_{ij} = b_n,$$

and the inequalities

$$\frac{c}{d} \lambda^n \le b_n \le \frac{dr}{c} \lambda^n.$$

The required equality follows from it.

2. Consider the case when there are several classes of equivalent recurrent vertices, i.e the matrix Π is decomposable. According to Proposition 1, by a renumbering of vertices this matrix may be transformed to the form

$$\Pi = \begin{pmatrix} (\Pi_1) & \cdots & \cdots & \cdots & \cdots \\ & \ddots & & & \\ 0 & & (\Pi_k) & \cdots & \cdots \\ & \ddots & & \ddots & \\ 0 & & 0 & & (\Pi_s) \end{pmatrix},$$

where each diagonal block Π_k corresponds to either one of the classes of equivalent recurrent vertices H_k or some non-recurrent vertex, and consists of one zero. Under diagonal blocks are zeroes. Each class H_k has the entropy

$$h(H_k) = \ln \lambda_k,$$

where λ_k is the maximal eigenvalue of Π_k. By the definition the entropy of a symbolic image G equals

$$h(G) = \lim_{n\to\infty} \frac{\ln b_n}{n}.$$

Consider an admissible n-length path ω. Assume that on G there are s classes of equivalence H_k. The path ω passes both through the vertices from H_k and non-recurrent vertices which do not belong to these classes. Denote by ω_k the parts of ω which lie in the class H_k. If we delete from ω all the ω_k it will contain only different paths σ_l passing through non-recurrent vertices. Thus ω is the sum ω_k and σ_l. Combine all the paths σ_l into a sequence σ. Generally speaking it is not an admissible path. Let K be the number of non-recurrent vertices in G. The sequence σ contains nonrecurrent vertices without repetitions. Hence the number of the sequences σ is not greater than the number of permutations of K elements , and it equals $K!$.

Denote by $n(k)$ the length of the path ω_k from the class H_k. Then $n(1) + n(2) + \cdots + n(s) \le n$. According to item 1 for every class H_k there is a number d such that the number $b_k(n(k))$ of different $n(k)$-length paths ω_k is estimated as follows

$$b_k(n(k)) \le d\lambda_k^{n(k)} \le d\lambda^{n(k)},$$

где $\lambda = \max \lambda_k$. Then for the number of different paths ω_k which are in ω we have the estimation

$$\prod_k b_k(n(k)) \le d^s \lambda^{n(1)+n(2)+\cdots+n(s)} \le d^s\lambda^n.$$

Summing the above estimations obtain the following

$$b_n \le K! d^s \lambda^n.$$

Thus, we have the upper estimation for the entropy of G:

$$h(G) \le \lim_{n\to\infty} \frac{1}{n} \ln(K! d^s \lambda^n) = \ln \lambda.$$

Prove the opposite inequality. Note that the number of admissible paths on G is greater than the number of admissible paths in a class H_k. Then $h(G) \geq h(H_k) = \ln \lambda_k$ for any k, which gives the low estimation

$$h(G) \geq \ln \lambda.$$

Hence we have

$$h(G) = \ln \lambda,$$

and the proof is completed.

4 Flows on a symbolic image

Let $f : M \to M$ be a homeomorphism of a compact manifold M. A measure μ defined on M is said to be f-invariant, if for any measurable set $A \subset M$ the equality

$$\mu(f^{-1}(A)) = \mu(A) = \mu(f(A))$$

holds. In what follows we assume that all measures considered are the Borel ones. The Krylov-Bogoliubov theorem [13, 11] guaranties the existence of an invariant measure μ which is normed on M: $\mu(M) = 1$. Denote by $\mathcal{M}(f)$ the set of all f-invariant normed measures . This set is a convex closed compact in weak topology (see [15], p.511). The convergence $\mu_n \to \mu$ in this topology means that

$$\int_M \phi d\mu_n \to \int_M \phi d\mu$$

for any continuous function $\phi : M \to R$.

To understand how a distribution of a measure may appear on a symbolic image, consider the following observation. Assume that there exists a f-invariant normed measure μ on M, and the cells of a covering C are polyhedrons intersecting by boundary disks. Construct a measurable partition $C^* = \{M^*(i)\}$ such that a boundary disk belongs to one of adjoining cells. Then, to every edge $i \to j$ of a symbolic image G we can assign the measure

$$m_{ij} = \mu(M^*(i) \cap f^{-1}(M^*(j))) = \mu(f(M^*(i)) \cap M^*(j)), \tag{5}$$

where the last equality follows from the invariance of μ. Besides that, the invariance of μ leads to the equalities

$$\sum_k m_{ki} = \sum_k \mu(f(M^*(k)) \cap M^*(i))) = \mu(M^*(i)) =$$

$$\sum_j \mu(M^*(i) \cap f^{-1}(M^*(j))) = \sum_j m_{ij}.$$

The value $\sum_k m_{ki}$ is called the flow incoming in the vertex i, and the $\sum_j m_{ij}$ — the flow outcoming from i. The equality

$$\sum_k m_{ki} = \sum_j m_{ij} \tag{6}$$

may be interpreted as Kirchoff's law: for any vertex the incoming flow equals the outcoming one. Furthemore, we have

$$\sum_{ij} m_{ij} = \mu(M) = 1. \tag{7}$$

It means that the distribution m_{ij} is normed (probabilistic). Thus, a f-invariant measure μ generates on a symbolic image a distribution m_{ij} which satisfies the conditions (6) and (7). The above reasoning leads to the following definition.

Definition. 5 *Let G be an oriented graph. The distribution $\{m_{ij}\}$ on edges $\{i \to j\}$ such that*

- $m_{ij} \geq 0$;
- $\sum_{ij} m_{ij} = 1$;
- *for any vertex i*

$$\sum_k m_{ki} = \sum_j m_{ij}$$

is called flow on G.

The last property may be called the invariance of a flow. The norming condition may be written as $m(G) = 1$, where the measure of G is the sum of measures of all edges. Sometimes in the graph theory for such a distribution the term "closed flow" is used.

For the flow $\{m_{ij}\}$ on G we may define the measure of a vertex i as

$$m_i = \sum_k m_{ki} = \sum_j m_{ij}.$$

Then $\sum_i m_i = m(G) = 1$.

Thus, a f-invariant measure generates a flow on a symbolic image. Now we consider the inverse construction. Let on a symbolic image G a flow $m = \{m_{ij}\}$ be given, then we can construct the measure μ on M as follows

$$\mu(A) = \sum_i \frac{m_i(v(A \cap M(i))}{v(M(i))}. \tag{8}$$

Here v is a normed on M Lebesgue's measure, and on the assumption $v(M(i)) \neq 0$. In this case the measure of a cell $M(i)$ coincides with the measure of the vertex i: $\mu(M(i)) = m_i$. As v is the Lebesgue measure, the measure of boundary disks is equal to zero and the measure of a cell does not depend on the measure of its boundary. In general, the constructed measure μ is not f-invariant. But it is an approximation to an invariant measure in the sense that μ converges in weak topology to an invariant measure if the diameter of the covering tends to zero.

Theorem. 4 *[18] Let on a sequence of symbolic images $\{G_t\}$ of a homeomorphism f a sequence of flows $\{m^t\}$ be defined, and the maximal diameter d_t of partitions tends to zero when $t \to \infty$. Then*

- *there exists the sequence of measures μ_{t_k} (constructed according to (8)) which converges in weak topology to a f-invariant measure μ;*
- *if a subsequence of measures μ_{t_l} converges in weak topology to a measure μ^*, then μ^* is f-invariant.*

Theorem. 5 *[18] For any neighborhood (in weak topology) U of the set $\mathcal{M}(f)$ there is a positive number d_0 such that for any partition C with the diameter $d < d_0$ and any flow m on a symbolic image G with respect to C, the measure μ constructed according to (8) by m, lies in U.*

5 Metric entropy

Let μ be a normed invariant measure of a homeomorphism $f : M \to M$ and $C = \{M_1, M_2, \cdots, M_m\}$ a measurable partition of the manifold M.

Definition. 6 *The entropy of the partition C is defined as*

$$H(C) = -\sum_i \mu(M_i) \ln \mu(M_i).$$

Construct a covering C^N which consists of nonempty intersections of the form

$$A_{i_1} \cap f^{-1}(A_{i_2}) \cap \cdots \cap f^{-N+1}(M_{i_N}).$$

If such an intersection is nonempty then the sequence $i_1, i_2, \cdots i_N$ is admissible with respect to the covering C.

The metric entropy of f for the covering C is defined as

$$H(f, C) = \lim_{N\to\infty} \frac{1}{N} H(C^N).$$

The existence of the limit follows from the Polya lemma.

Definition. 7 *The entropy of f for an invariant measure μ is defined as*

$$h(f, \mu) = \sup_C H(f, C),$$

where sup is taken over all measurable finite partitions.

The connection between topological and metric entropy is given by the following theorem.

Theorem. 6 *[6, 9] The topological entropy of a homeomorphism f is the least upper bound of metric entropies*

$$h(f) = \sup_\mu h(f, \mu).$$

6 Stohastic Markov chains

Stohastic Markov chain [8, 14] is defined by a set of states of a system $\{i = 1,\ 2,\ \ldots n\}$ and the matrix of transition probabilities P_{ij} from a state i to state

j. Such a matrix is called stohastic if it satisfies the following conditions $P_{ij} \geq 0$ and $\sum_j P_{ij} = 1$ for every i. A probabilistic distribution $p = (p_1, p_2, \ldots, p_n)$, $\sum_i p_i = 1$ is said to be stationary if

$$(p_1, p_2, \ldots, p_n) \begin{pmatrix} P_{11} & P_{12} & \ldots & P_{1n} \\ P_{21} & P_{22} & \ldots & P_{2n} \\ \cdot & \cdot & \ldots & \cdot \\ P_{n1} & P_{n2} & \ldots & P_{nn} \end{pmatrix} = (p_1, p_2, \ldots, p_n),$$

i.e. p is a left eigenvector of P.

We show that there is a one-to-one correspondence between a Markov chain and a flow on a graph in which vertices correspond to the states with positive measure.

Let $m = \{m_{ij}\}$ be a flow on a graph G. The measure of a vertex i equals $m_i = \sum_j m_{ij} = \sum_k m_{ki}$. If $m_i \neq 0$ then the vertex $\{i\}$ is necessary recurrent. It is easy to verify that any flow $m = \{m_{ij}\}$ on G generates a stohastic Markov chain in which the states are vertices with nonzero measures, and the transition probabilities from i to j are calculated as

$$P_{ij} = \frac{m_{ij}}{m_i}.$$

In this case the stohastic matrix $P = (m_{ij}/m_i)$ has the stationary distribution coinciding with the distribution of the measure m over the vertices $(m_1, m_2, \ldots, m_n)$. This follows from the equality

$$(m_1, m_2, \ldots, m_n) \begin{pmatrix} \frac{m_{11}}{m_1} & \frac{m_{12}}{m_1} & \ldots & \frac{m_{1n}}{m_1} \\ \frac{m_{21}}{m_2} & \frac{m_{12}}{m_2} & \ldots & \frac{m_{2n}}{m_2} \\ \cdot & \cdot & \ldots & \cdot \\ \frac{m_{n1}}{m_n} & \frac{m_{n2}}{m_n} & \ldots & \frac{m_{nn}}{m_n} \end{pmatrix} = (m_1, m_2, \ldots, m_n).$$

Thus, any flow $m = \{m_{ij}\}$ on a graph G generates a stohastic Markov chain for which the distribution of the measure (m_i) on vertices is stationary.

Now we prove the inverse fact: for any stohastic matrix$P = (P_{ij})$ and its stationary distribution $p = (p_i)$ there exists a flow $m = \{m_{ij}\}$ on a graph G for which the distribution of the measure on vertices coincides with the stationary distribution, i.e. $m_i = p_i$.

Actually, let P be a stohastic matrix and $pP = p$. Consider a graph G which has n vertices $\{i\}$, and the edge $i \to j$ there exists if $P_{ij} > 0$. Construct the distribution on edges $m_{ij} = P_{ij} p_i$ and show that the distribution is a flow on G. As P is stohastic then $\sum_j P_{ij} = 1$ for any i. Hence

$$\sum_j m_{ij} = \sum_j P_{ij} p_i = p_i \sum_j P_{ij} = p_i.$$

As $pP = p$ then $\sum_k p_k P_{ki} = p_i$, hence

$$\sum_k m_{ki} = \sum_k p_k P_{ki} = p_i = \sum_j m_{ij},$$

i.e for the distribution m_{ij} the Kirchoff law holds. Moreover, $\sum_{ij} m_{ij} = \sum_i p_i = 1$.

From the above it follows that the construction of a flow on a graph results in obtaining a Markov chain.

7 Flow entropy

The developed technics may be applied to estimate metric entropy. Let for a mapping f and a covering C a symbolic image G and a flow $m = \{m_{ij}\}$ be constructed. As it was proved above, any flow m may be considered as the approximation to an invariant measure μ, if the diameter of C is small enough. The flow m on G generates the Markov chain in which the states coincide with vertices of G, and transition probabilities are defined as

$$p_{ij} = \frac{m_{ij}}{m_i}.$$

The matrix $P = (p_{ij})$ has the stationary distribution $(m_1, m_2, \ldots, m_n)$ for which entropy is calculated by the formula (see [14], p. 443)

$$h_m = -\sum_i m_i \sum_j p_{ij} \ln p_{ij}.$$

Substituting $p_{ij} = m_{ij}/m_i$ we obtain

$$h_m = -\sum_i m_i \sum_j \frac{m_{ij}}{m_i} \ln(\frac{m_{ij}}{m_i}) = -\sum_{ij} m_{ij} \ln(\frac{m_{ij}}{m_i}) =$$

$$-\sum_{ij} m_{ij} \ln m_{ij} + \sum_{ij} m_{ij} \ln m_i = -\sum_{ij} m_{ij} \ln m_{ij} + \sum_i m_i \ln m_i.$$

By this means entropy can be calculated by the flow m_{ij} as

$$h_m = -\sum_{ij} m_{ij} \ln m_{ij} + \sum_i m_i \ln m_i. \tag{9}$$

The last equality allows estimating the entropy of f for the invariant measure μ , where the flow m is an approximation of μ.

8 Flow with maximal entropy

Let Π be the matrix of admissible transitions for a graph G. Our objective is to construct the flow which has maximal entropy among all the flows on G. As any flow is grouped on a component of recurrent vertices, it may be thought that G consists from one component.

Theorem. 7 *There is a flow m on G such that:*

$$h_m = h(G) = \ln \lambda.$$

Proof.

1. Eigenvalues of any real matrix $A = (a_{ij})$ coincide with the eigenvalues of the transposed (conjugate) matrix A^* . Really, as $\det A = \det A^*$, then

$$\det(A - \lambda E) = \det(A - \lambda E)^* = \det(A^* - \overline{\lambda} E).$$

Hence to an eigenvalue λ of A corresponds the conjunctive eigenvalue $\overline{\lambda}$ of A^*. The roots of a real characteristic polynomial are either real or complex-conjugate, hence the eigenvalues of the matrices A and A^* coincide.

2. Let A be the matrix of admissible transitions of a graph G and λ be the maximal eigenvalue from the Perron-Frobenius theorem. Then for A there

exists a left eigenvector e with nonnegative coordinates e_i,$\sum_i e_i = 1$, such that

$$eA = \lambda e,\ \ A^* e = \lambda e.$$

Hence for every i we have

$$\sum_j a_{ji} e_j = \lambda e_i, \tag{10}$$

which leads to the equality

$$\sum_j \frac{a_{ji} e_j}{\lambda e_i} = 1$$

for every i. Hence a matrix of the form

$$P = \left(p_{ij} = \frac{a_{ji} e_j}{\lambda e_i} \right)$$

is the stohastic matrix for which vector e is a stationary distribution:

$$eP = e.$$

The distribution on edges $i \to j$ defined by

$$m_{ij} = p_{ij} e_i = \frac{a_{ji} e_j}{\lambda}$$

is the flow m on the graph G such that the measure m_i of the vertex i equals e_i. The entropy of m is calculated by the formula

$$h_m = -\sum_{ij} m_{ij} \ln m_{ij} + \sum_i m_i \ln m_i.$$

Hence

$$h_m = -\sum_{ij} \frac{a_{ji} e_j}{\lambda} \ln \frac{a_{ji} e_j}{\lambda} + \sum_i e_i \ln e_i.$$

Here we assume that $0 \ln 0 = 0$. That means that the sum is taken over i, j for which $a_{ij} = 1$. Thus we obtain

$$h_m = -\sum_{ij} \frac{a_{ji} e_j}{\lambda} (\ln a_{ij} + \ln e_i - \ln \lambda) + \sum_i e_i \ln e_i =$$

$$(\sum_i (\sum_j \frac{a_{ji} e_j}{\lambda}) \ln \lambda - \sum_i (\sum_j \frac{a_{ji} e_j}{\lambda}) \ln e_i + \sum_i e_i \ln e_i =$$

$$\ln \lambda \sum_i e_i - \sum_i e_i \ln e_i + \sum_i e_i \ln e_i = \ln \lambda.$$

The proof is completed.

References

[1] *Alekseev V.* Simvolicheskaia dynamika, 11 matematicheskaia shcola, Izd. instituta matematiki AN SSSR, Kiev, 1976. (in Russian)

[2] *Ampilova N., Soloviev I.* Entropies in investigation of dynamical systems and their application to digital image analysis. Journal of Measurements in Engineering, 6, No. 2, 107-118(2018) https://doi.org/10.21595/jme.2018.19891

[3] *Bowen R.* Symbolic Dynamics, Ann. Math. Soc., 8 (1982), Providence.

[4] *Boltzmann L.* Sitzungsber. Kaiserl. Akad. Wiss. 66 (2) 275 (1872)

[5] *Clausius R.* Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie. Annalen der Physik und Chemie,125,No.7, 353-400 (1865).

[6] *Dinaburg E.* Sviaz megdu razlichnimi entropiinimi characteristicami dynamicheskih system, Izv. AN SSSR, matem. 35, No.2,324-366(1971).(in Russian)

[7] *Downarowicz T.* Entropy in dynamical systems. Cambridge Univ. Press, 2011.

[8] *Gantmaher F.* Teoria matrits, M., Nauka, 1967.(in Russian)

[9] *Goodman T.* Relating topological entropy and measure theoretic entropy. Bulletin of London Math. Soc. 3, No. 2, 176-180 (1971.

[10] *Hsu C., S.* Cell-to-Cell Mapping, Springer-Verlag, N.Y., 1987.

[11] *Katok A., Hasselblat B.* Introduction to the Modern Theory of Dynamical Systems, Cambridge University Press, 1995.

[12] *Kolmogorov A.* Novii metricheskii invariant transitivnih dynamicheskih system i avtomorphismov prostranstv Lebega", Dokl. AN SSSR, 119, No.5,861–864 (1958) (in Russian)

[13] *Kryloff N., Bogollouboff N.* La theorie generale de la mesure dans son application a l'etude das systemes dynamiques de la mecanique non lineaire, Ann. Math., 38, No.1,65-113 (1937).

[14] *Lind D., Marcus B.* An introduction to symbolic dynamics and coding. Cambridge University Press, 1995.

[15] *Nemitsky V., Stepanov V.* Kachestvennaia teoria differentialnih uravnenii, Moskva-Leningrad: GITL, 1949.(in Russian)

[16] *Osipenko G.* O simvolicheskom obraze dinamicheskoi systemi, Kraevie zadachi, Sb.trudov, Perm, , 101-105(1983).(in Russian)

[17] *Osipenko G.* Dynamical systems, Graphs, and Algorithms. Lectures Notes in Mathematics, 1889, Berlin: Springer,2007.

[18] *Osipenko G.* Symbolic images and invariant measures of dynamical systems. Ergodic Theory and Dynamical Systems, 30, 1217–1237 (2010).

[19] *Prasolov V.* Elementi combinatornoi i and differentialnoi topologii, M.: MCNMO , 2004. (in Russian)

[20] *Shannon C.* A Mathematical Theory of Communication. Bell System Technical Journal. 27, No. (3), 379–423 (July 1948). doi:10.1002/j.1538-7305.1948.tb01338.x. hdl:11858/00-001M-0000-002C-4314-2.

[21] *Shannon C.* A Mathematical Theory of Communication. Bell System Technical Journal. 27, No.4, 623–666 (October 1948. doi:10.1002/j.1538-7305.1948.tb00917.x. hdl:11858/00-001M-0000-002C-4314-2.