Journal of Dynamical and Control Systems, Vol. 19, No. 2, April 2013, 301–308 (©2013)
DOI: 10.1007/s10883-013-9174-7

REMARKS ON THE ENTROPY OF SOFIC DYNAMICAL SYSTEM OF BLACKWELL'S TYPE

Z.I. BEZHAEVA and V.I. OSELEDETS

ABSTRACT. Consider a sofic dynamical system (X, T, μ), where $X = A^Z$ is the full symbolic compact set with the product topology, and $A = \{0, 1, \dots, d\}$. The shift is $T : \{x_n\} \to \{x_n^{'}\}, x_n^{'} = x_{n+1}$. The measure μ is a T-invariant sofic probability measure. For all words $a_1 \dots a_n$ the measure is $\mu(a_1 \dots a_n) = \mu(\{x : x_1 = a_1, \dots, x_n = a_n\}) = lm_{a_1} \dots m_{a_n} r$. Matrices $\{m_0, \dots, m_d\}$, $d \geq 1$, are nonzero substochastic matrices of order J. The matrix $P = m_0 + \dots + m_d$ is a stochastic matrix, the row l is a left P-invariant probability row and all entries of the column r are equal to 1.

We obtain an explicit formula for the entropy $h(T, \mu)$ of sofic dynamical system of Blackwell's type for which $rank(m_a) = 1, a \neq 0$.

1. INTRODUCTION

Consider a sofic dynamical system (X, T, μ), where $X = A^Z$ is the full symbolic compact set with the product topology, and $A = \{0, 1, \dots, d\}$. The shift is

$$T : \{x_n\} \to \{x_n^{'}\}, \qquad x_n^{'} = x_{n+1}.$$

The measure μ is a T-invariant sofic probability measure. The measure μ is called sofic measure if for all words $a_1 \dots a_n$

$$\mu(a_1 \dots a_n) = \mu\left(\{x : x_1 = a_1, \dots, x_n = a_n\}\right) = lm_{a_1} \dots m_{a_n} r,$$

where $\{m_0, \dots, m_d\}$, $d \geq 1$, are nonzero substochastic matrices of order J. The matrix $P = m_0 + \dots + m_d$ is a stochastic matrix, the row l is a left P-invariant probability row and all entries of the column r are equal to 1.

2000 *Mathematics Subject Classification.* 28D05, 28D20, 60G10.

Key words and phrases. Entropy, sofic measure, hidden Markov chain.
The work was supported by RFBR grant 11-01-00982.
This study was carried out within "The National Higher School of Economics" Academic Fund Program in 2013-2014, research grant No.12-01-0056.

Sofic measure coincides with some hidden Markov measure. There exists a stationary finite-state Markov chain $\{\xi_n\}$ and a stationary process η_n with states $a = 0, 1, \ldots, d$ such that $\eta_n = \Phi(\xi_n)$, and

$$\mu(a_1 \ldots a_n) = P(\eta_1 = a_1, \ldots, \eta_n = a_n).$$

In the paper [3] the authors used the title "manifestly algebraic measure" for sofic measure and in the paper [2] the authors used the name "rational measure."

In symbolic dynamics Markov measures have been thoroughly studied. Their entropy is given by an explicit formula. Computation of the entropy of sofic dynamical system is a difficult problem.

We obtain an explicit formula for the entropy $h(T, \mu)$ of sofic dynamical system of Blackwell's type for which $\mathrm{rank}(m_a) = 1, a \neq 0$. In Sec. 4 we show numerical computation of the entropy using our formula.

It is easy to obtain a similar formula for the entropy of sofic dynamical system, when $\mathrm{rank}(M_a) = 1$ for some a.

We show how Blackwell's method computation of the entropy [1] works for any sofic dynamical system and in particular for sofic dynamical system of Blackwell's type. It seems that our method of calculating the entropy in this case is significantly easier than elegant Blackwell's method.

2. The entropy of sofic dynamical system. Abramov's formula and countable Markov chains

Consider the set

$$X^{'} = \{x \in X : x_1 \neq 0, \{i : x_i \neq 0\} \text{ is an infinite set}\}.$$

We define the return function $F(x)$ on the space $X^{'}$ as

$$F(x) = \min\left\{n \in N : T^n x \in X^{'}\right\}.$$

The induced map $T^{'}$ on the space $X^{'}$ is given by the formula $T^{'}x = T^{F(x)}x$. The conditional measure $\mu^{'}$ for the measure μ on the space $X^{'}$ is a $T^{'}-$ invariant measure.

By Abramov's formula for the entropy of ergodic measure μ [8],

$$h(T, \mu) = h\left(T^{'}, \mu^{'}\right) \mu\left(X^{'}\right).$$

Let

$$Z_n = f\left(T^{n-1}x\right), \quad f(x) = \left(x_1, F(x) - 1, x_{F(x)+1}\right), \quad x \in X^{'}.$$

Proposition 1. *Z_n is the stationary countable-state Markov process with the states (i, n, j), $i, j = 1, \ldots, d$, $n = 0, 1, \ldots$.*

The Markovian property of the process Z_n is a consequence of our condition $\mathrm{rank}(m_a) = 1, a \neq 0$. We have $m_i = u_i v_i, i \neq 0$, where v_i is a probability row $v_i r = 1$, and u_i is a nonnegative column, so that

$$\mu(a_1 \ldots, i, \ldots a_n) = (l m_{a_1} \ldots u_i)(v_i \ldots m_{a_n} r), \qquad \text{if } i \neq 0.$$

Therefore

$$P(Z_1 = (k, m, i)) = \frac{l m_k m_0^m m_i r}{\mu(X')} = \frac{(l u_k)(v_k m_0^m u_i)}{\mu(X')},$$

$$\begin{aligned} P(Z_1 = (k, m, i),\ Z_2 = (i, n, j)) &= \frac{l m_k m_0^m m_i m_0^n m_j r}{\mu(X')} \\ &= \frac{(l u_k)(v_k m_0^m u_i)(v_i m_0^n u_j)}{\mu(X')}. \end{aligned}$$

We obtain the transition probability for Markov process Z_n:

$$p((k, m, i), (i, n, j)) = \frac{P(Z_1 = (k, m, i), Z_2 = (i, n, j))}{P(Z_1 = (k, m, i))} = v_i m_0^n u_j.$$

The entropy $h\left(T^{'}, \mu^{'}\right)$ is equal to the entropy of stationary countable - state Markov chain Z_n. Hence

$$h(T, \mu) = \mu\left(X^{'}\right) h\left(T^{'}, \mu^{'}\right) = - \sum_{k,m,i,n,j} \log(v_i m_0^n u_j)(v_i m_0^n u_j) l u_k v_k m_0^m u_i.$$

But

$$\begin{aligned} \sum_{k,m} l u_k v_k m_0^m u_i &= l \sum_{k \geq 1} m_k (I - m_0)^{-1} u_i = l(P - m_0)(I - m_0)^{-1} u_i \\ &= l(I - m_0)(I - m_0)^{-1} u_i = l u_i. \end{aligned}$$

Thus our main result is the following:

$$h(T, \mu) = - \sum_{i,j \geq 1, n \geq 0} \log(v_i m_0^n u_j) v_i m_0^n u_j l u_i.$$

3. Blackwell's entropy formula and Blackwell Markov chains

Our simple variant of Markov chain $\{\xi_n\}$ for sofic measure μ (see [5, 6]) is given by the transition matrix $\bar{P}$. The matrix $\bar{P}$ is the block matrix in which all block rows are equal to the block row $(m_0, \ldots, m_d)$, and left $\bar{P}$ - invariant probability block row is equal to $(l m_0, \ldots, l m_d)$.

The state space $\{1, \ldots, (d+1)J\}$ is the union of the blocks

$$B_0 = \{1, \ldots, J\}, \qquad \ldots, \qquad B_d = \{(d+1)(J-1), \ldots, (d+1)J\}$$

and $\Phi(i) = a, i \in B_a$.

Let d_a be the diagonal matrix with diagonal $(I_a(k), k = 1, \dots, (d+1)J)$, where $I_a(k) = 1$, $k \in B_a$, $a = 0, 1, \dots, d$. Let

$$\bar{m}_i = \bar{P} d_a, \quad a = 0, 1, \dots, d, \quad \bar{r} = \begin{pmatrix} r \\ \vdots \\ r \end{pmatrix}.$$

Then

$$P(\eta_1 = a_1, \dots, \eta_n = a_n) = \bar{l}\bar{m}_{a_1} \dots \bar{m}_{a_n}\bar{r} = l m_{a_1} \dots m_{a_n} r = \mu(a_1 \dots a_n).$$

We will suppose that there exists unique left P-invariant probability row. Then there exists unique left $\bar{P}$-invariant probability row. We give a short proof of the uniqueness of left $\bar{P}$-invariant probability row $(l_0, \dots, l_d)$.

Let

$$(l_0, \dots, l_d)\bar{P} = (l_0, \dots, l_d).$$

Hence

$$(l_0 + \dots + l_d) m_a = l_a, \quad a \in \{0, 1, \dots, d\},$$
$$(l_0 + \dots + l_d) P = l_0 + \dots + l_d, \quad (l_0 + \dots + l_d) r = 1.$$

Therefore

$$l_0 + \dots + l_d = l, \quad l_a = l m_a.$$

The uniqueness of left $\bar{P}$-invariant probability row gives the ergodicity of Markov chain $\{\xi_n\}$ and implies the ergodicity of hidden Markov chain $\eta_n = \Phi(\xi_n)$. Hence, the measure μ is an ergodic measure.

Now we show how Blackwell's method [1] works for any sofic dynamical system and in particular for sofic dynamical system of Blackwell's type.

Blackwell introduced the stationary Markov chain

$$\alpha_n = (P(\xi_n = k \,|\, \eta_n, \eta_{n-1}, \eta_{n-2}, \dots), \quad k = 1, \dots, (d+1)J)$$

with states $\bar{w} = (0, \dots, w_a, \dots, 0)$, and $w_a \geq 0$, and $w_a r = 1, a = 0, 1, \dots, d$.

Let

$$g(a, \eta_0, \eta_{-1}, \dots) = P(\eta_1 = a \,|\, \eta_0, \eta_{-1}, \dots).$$

Then

$$h(T, \mu) = -E \log(g(\eta_1, \eta_0, \eta_{-1}, \dots)).$$

Blackwell proved that

$$g(\eta_1, \eta_0, \eta_{-1}, \dots) = \alpha_0 \bar{m}_{\eta_1} \bar{r}.$$

Blackwell's formula [1] for the entropy hidden Markov chain η_n is $-E \log(\alpha_0 \bar{m}_{\eta_1} \bar{r})$.

The transition probability of Markov chain α_n is given by $\bar{w}\bar{m}_a\bar{r}$ for the transition

$$\bar{w} \to \bar{w}\bar{m}_a / \bar{w}\bar{m}_a\bar{r},$$

if $\bar{w}\bar{m}_a\bar{r} > 0$.

Let the function φ be given by $\varphi((0,\dots,w_a,\dots,0)) = w_a$. Let $\beta_n = \varphi(\alpha_n)$. Then β_n is the stationary Markov process with states $w, w \geq 0, wr = 1$. The transition probability of Markov chain β_n is given by wm_ar for the transition

$$w \to wm_a/wm_ar,$$

if $wm_ar > 0$.

Hence, Blackwell's formula [1] for the entropy hidden Markov chain η_n is $-E\log(\beta_0 m_{\eta_1} r)$.

Let q be the distribution of β_0. Then by Blackwell's formula the entropy of sofic dynamical system is equal to

$$h(T,\mu) = -\int_W \sum_{a=0}^{d} \log(wm_ar)wm_arq(dw),$$

where $W = \{w : w \geq 0, wr = 1\}$.

We have

$$q\left(\{w : w = v_i m_0^n/v_i m_0^n r,\ i = 1,2,\dots d,\ n \geq 0\}\right) = 1,$$

because for $i \neq 0$

$$wm_i/wm_ir = wu_iv_i/wu_iv_ir = v_i.$$

Thus β_n is the countable-state Markov chain with states

$$v_i m_0^n/v_i m_0^n r, \qquad i = 1,\dots,d, \qquad n \geq 0.$$

For general case these states are different states.

The distribution q is given by the formula

$$q(v_i m_0^n/v_i m_0^n r) = lu_i v_i m_0^n r, \qquad i = 1,\dots,d, \qquad n \geq 0.$$

It is easy to prove this.

The transition probability is given by $v_i m_0^n m_j r/v_i m_0^n r$ for the transition

$$v_i m_0^n/v_i m_0^n r \to v_i m_0^n m_j/v_i m_0^n m_j r, \quad j \geq 0.$$

Finally

$$h(T,\mu) = -\sum_{i\geq 1, j\geq 0, n\geq 0} \log(v_i m_0^n m_j r/v_i m_0^n r) v_i m_0^n m_j r l u_i.$$

It is strange that Blackwell's formula and our formula coincide one with another, but it is easy to prove this.

4. Examples

Example 1. Let

$$P = \begin{pmatrix} p_{11} & 0 & p_{13} \\ 0 & p_{22} & p_{23} \\ p_{31} & p_{32} & 0 \end{pmatrix}.$$

Let Φ be the function defined by $\Phi(1) = \Phi(2) = 0, \Phi(3) = 1$. Let

$$d_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad d_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let $m_0 = Pd_0, m_1 = Pd_1$. Here $v_1 = (0,0,1), u_1^T = (p_{13},\ p_{23},\ 0)$. Then

$$p(n) = v_1 m_0^n u_1 = p_{31}p_{13}p_{11}^n + p_{32}p_{23}p_{22}^n,$$
$$lu_1 = \frac{p_{13}p_{23}}{p_{13}p_{23} + p_{23}p_{31} + p_{13}p_{12}},$$
$$h(T,\mu) = -\sum_{n\geq 0} \log(p(n))p(n)lu_1.$$

For Blackwell's example [1]

$$p_{13} = p_{32} = p_{22} = \frac{1}{3},$$
$$p_{11} = p_{23} = p_{31} = \frac{2}{3},$$
$$h(T,\mu) = 0.47817623354875\ldots.$$

Example 2. "A hidden Markov process is a discrete-time finite-state homogenous Markov chain observed through a memoryless invariant channel is a hidden Markov chain." [4]

Let P be the transition matrix of Markov chain with the states $1,\ldots,$ J. In this case $m_a = Pd_a$, where d_a is a nonnegative diagonal matrix and $(d_a)_{jj}$ is the transition probability of the transition $j \to a, a = 0,1,\ldots,d$.

The following example is given in [7]: $J = d+1$,

$$(d_a)_{jj} = \begin{cases} 1-\varepsilon_a, & j = a, \\ 0, & j \neq a, \end{cases}$$

if $a > 0$ and

$$(d_0)_{jj} = \begin{cases} 1, & j = 0, \\ \varepsilon_a, & j > 0. \end{cases}$$

Here

$$m_a = Pd_a, \qquad P = m_0 + \cdots + m_J,$$

and $\mathrm{rank}(m_a) = 1,\ a \neq 0$.

The authors of paper [7] obtained a formula for the entropy of hidden Markov chain. They used the matrices $e_a = d_a P$. But $\mu(a_1 \dots a_n) = l e_{a_1} \dots e_{a_n} r$.

We calculate the entropy for two concrete examples from [7].

Let $J = 2, d = 1, \varepsilon_1 = 0.01$,

$$P = \begin{pmatrix} 0.85 & 0.15 \\ 0.28 & 0.72 \end{pmatrix}.$$

Then $h = h(T, \mu) = 0.48545683971057\dots$ (In [7] $h = 0.7003661\dots$).

Let $J = 3, d = 2, \varepsilon_1 = 0.01, \varepsilon_2 = 0.02$ and

$$P = \begin{pmatrix} 0.4 & 0.25 & 0.25 \\ 0.25 & 0.45 & 0.3 \\ 0.2 & 0.55 & 0.25 \end{pmatrix}.$$

Then $h = h(T, \mu) = 1.05424072438249\dots$ (In [7] $h = 0.95961126164044\dots$).

References

1. D. Blackwell, The entropy of functions of finite state Markov chains. *Trans. First Praque Conf. Inf. Theory, Stat. Dec. Functions, Random Proc.* (1957), 13–20.
2. G. Hancel, D. I. Perrin, Rational probability measures. *Theor. Comp.Sci.* **65** (1989), 171–188.
3. M. Fannes, B. Nachtergaele and L. Slegers, Functions of Markov processes and algebraic measures. *Reviews in Mathematical Physics* **4** (1992), No. 1, 39–64.
4. Y. Ephraim, N. Merhav, Hidden Markov process. *IEEE Trans. Inform. Theory* **48** (2002), No. 6, 1518 – 1569.
5. Z. I. Bezhaeva, V. I. Oseledets, Erdös measures, sofic measures, and Markov chains. *Zap. Nauch. Sem. POMI* **326** (2005), 28–47.
6. Z. I. Bezhaeva, V. I. Oseledets, Erdös measures, sofic measures, and Markov chains. *Journal of Math. Science* **120** (2007), 28–47.
7. K. Marchand, J. Mulherkar and B. Nachtergaele, Entropy rate calculations of algebraic measures. *arXiv: 1105.2377v1* (2011), 1–8.
8. L. M. Abramov, Entropy of induced automorphism, *Dokl. Akad. Nauk SSSR* **128** (1959), 647–650.

(Received June 9 2012)

Authors' addresses:

Z.I. Bezhaeva

NRU HSE MIEM, Moscow, Russia;

National University, Moscow, Russia
E-mail: zbejaeva@hse.ru

V.I. Oseledets
Moscow State University, Moscow, Russia;
Financial State University, Moscow, Russia
E-mail: oseled@gmail.com