Časopis pro pěstování matematiky

Monica E. Dumitrescu

Some informational properties of Markov pure-jump processes

Časopis pro pěstování matematiky, Vol. 113 (1988), No. 4, 429--434

Persistent URL: `http://dml.cz/dmlcz/118348`

SOME INFORMATIONAL PROPERTIES OF MARKOV PURE-JUMP PROCESSES

MONICA BAD DUMITRESCU, Bucharest

(Received July 7, 1986)

Summary. In this note Boltzmann's entropy for the imbedded process is constructed. Also, the generalized equipartition property is established for an ergodic Markov pure-jump information source.

Keywords: Markov pure-jump process, equipartition property, Boltzmann's entropy.

AMS Classification: 60J75.

1. INTRODUCTION

Let us consider a Markov pure-jump process with finite state-space and the imbedded associated process, which is a bivariate discrete-time Markov process.

We calculate Boltzmann's entropy for the interval of time $[0, T]$, denoted H_T, and establish the existence of the entropy of the process,

$$H = \lim_{T\to\infty} \frac{1}{T} H_T .$$

Then we show that Perez's condition for the validity of the abstract alphabet version of the Shannon-McMillan's limit theorem is satisfied for the information source corresponding to the imbedded process. Hence, an ergodic Markov pure-jump information source has the generalized equipartition property.

2. MODEL

Let $\{X_t, t \geqq 0\}$ be a separable continuous-time Markov process defined on $(E, \mathscr{K}, P)$, with a finite state-space $S = \{1, 2, \ldots, s\}$. We suppose that the stationary transition probabilities $p_{ij}(t)$ are continuous at $t = 0$:

$$\lim_{t\to 0} p_{ij}(t) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

(then they are continuous for all t).

Let $Q = \|q_{ij}\|_{i,j=1,\dots,s}$ be the infinitesimal generator of the process, where

$$q_{ij} = \lim_{t\to 0} \frac{p_{ij}(t)}{t}, \quad 0 \leqq q_{ij} < \infty, \quad i \neq j$$

$$q_{ii} = -q_i, \quad q_i = \lim_{t\to 0} \frac{1 - p_{ii}(t)}{t}, \quad q_i > 0$$

$$\sum_{j\neq i} q_{ij} = q_i .$$

It is well known that the matrix of transition probabilities $P(t) = \|p_{ij}(t)\|$ can be expressed in the form $P(t) = \exp(tQ)$. Then, the process will be completely determined by the initial probabilities and by its infinitesimal generator Q.

If we assume that all the states inter-communicate, the chain will be ergodic. Then $p_{ij}(t) > 0$ for all $i, j \in S$ and all $t > 0$. The limits $\lim_{t\to\infty} p_{ij}(t) = \pi_j$ exist and $\{\pi_j, j \in S\}$ is the stationary distribution of the process, with $\pi_j > 0$, $\sum_{j\in S} \pi_j = 1$, $\sum_{i\in S} \pi_i p_{ij}(t) = \pi_j$ for all t.

The chain has the following constructive development:

The process starts off at time 0 broadcasting the signal i with probability $p_i(0)$ (π_i in the ergodic case). The emission-time of this first signal, T_0, has the density

$$f_{T_0}(t) = q_i \exp(-q_i t), \quad t > 0 .$$

At time $t = T_0$ a j-type signal is broadcasted ($j \neq i$) with probability q_{ij}/q_i. The emission-time of j is T_1, with the density

$$f_{T_1}(t) = q_j \exp(-q_j t), \quad t > 0 .$$

At time $t = T_0 + T_1$ the process jumps to the signal k ($k \neq j$), with probability q_{jk}/q_j, and so on.

Let $\{Z_0, Z_1, \dots\}$ be the successive states the system passes through ($Z_k \neq Z_{k+1}$, $k = 0, 1, \dots$).

The bivariate discrete-time process $\{(Z_k, T_k), k = 0, 1, \dots\}$ is a Markov process on the cartesian product $A = S \times (0, \infty)$. We call it the imbedded process. Its transition probabilities are

$$P(Z_{k+1} = j, \; T_{k+1} > t \mid Z_k = i, \; T_k = u) = \begin{cases} \dfrac{q_{ij}}{q_i} \exp(-q_j t), & i \neq j \\ 0 & , \; i = j \end{cases}$$

and its initial probabilities are

$$P(Z_0 = i, \; T_0 > t) = p_i(0) \exp(-q_i t), \quad t > 0 .$$

The information contained in the sample $\{X_t, 0 \leqq t \leqq T\}$ is essentially the same as in the sample $\{(Z_0, T_0), \dots, (Z_{N_T - 1}, T_{N_T - 1})\}$, where N_T is the random number of jumps which occur till the time T. The complete sample contains slightly more information than the sample from the imbedded process, the additional information being the length of the time interval from the last jump to the end of observation and the state occupied by the process during this period.

More precisely, the sample $\{X_t(w),\ 0 \leqq t \leqq T\}$ is equivalent, with probability one, to the sample $\{(Z_0(w), T_0(w)), \ldots, (Z_{N_T(w)-1}(w), T_{N_T(w)-1}(w)), Z_{N_T(w)}(w)\}$.

Let λ be the Lebesgue measure on R and c the counting measure on S. We consider $A_n = \prod_{j=1}^{n} [S \times (0, \infty)] \times S$ and let σ_n be the σ-finite measure on A_n defined by

$$\sigma_n = \prod_{j=1}^{n} [c \otimes \lambda] \otimes c .$$

Let $B = \bigcup_{n=0}^{\infty} A_n$ and for each set $M \subset B$ for which $M \cap A_n$ is σ_n-measurable define $\sigma^*(M) = \sum_{n=0}^{\infty} \sigma_n(M \cap A_n)$.

σ^* is defined on the σ-field $\mathscr{F}^*$ which is the smallest σ-field containing all sets $M \subset B$ whose projection on $\mathbb{R}_n$ is a Borel set for each n.

Let μ^* be the probability distribution of the imbedded process, which is absolutely continuous with respect to σ^*. Then the density corresponding to the sample $\{X_t(w),\ 0 \leqq t \leqq T\}$ with $N_T(w) = n$ is

$$f_T(v) = \begin{cases} P(X_0 = z_0) \exp(-q_{z_0} T), & \text{if } v = (z_0) \\ P(X_0 = z_0) \prod_{j=0}^{n-1} q_{z_j z_{j+1}} \exp[-(q_{z_j} - q_{z_n}) t_j - q_{z_n} T], & \\ \quad \text{if } v = ((z_0, t_0), \ldots, (z_{n-1}, t_{n-1}), z_n) \text{ with } t_j \geqq 0, \ \sum_{j=0}^{n-1} t_j < T & \\ 0, \text{ otherwise}. \end{cases}$$

Let $n_T(i, j)$ be the total number of transitions from state i to state j during $[0, T]$ and $r_T(i)$ the total length of the time interval during which state i is occupied (the signal i is broadcasted). Then

$$f_T(v) = k \prod_{\substack{i,j=1 \\ i \neq j}}^{s} [q_{ij}]^{n_T(i,j)} \prod_{i=1}^{s} \exp(-q_i\, r_T(i)) .$$

3. INFORMATIONAL PROPERTIES

Boltzmann's entropy for the observation interval $[0, T]$ is

$$H_T = -\int f_T(v) \ln f_T(v)\, d\sigma^*(v) .$$

$$H_T = -\ln k - \sum_{\substack{i,j=1\\i\neq j}}^{s} E(n_T(i,j)) \ln q_{ij} + \sum_{i=1}^{s} E(r_T(i))\, q_i\,,$$

$$H_T = -\ln k - \sum_{\substack{i,j=1\\i\neq j}}^{s} \left(\int_0^T P(X_t = i)\,\mathrm{d}t\right) q_{ij} \ln q_{ij} + \sum_{i=1}^{s} \left(\int_0^T P(X_t = i)\,\mathrm{d}t\right) q_i\,.$$

We define *the entropy* of a Markov pure-jump process as

$$H = \lim_{T\to\infty} \frac{1}{T} H_T\,.$$

Property 1. For an ergodic Markov pure-jump process the entropy H is finite,

$$H = \frac{1}{\varrho} \sum_{i,j=1}^{s} Q^{ii} q_{ij}(1 - \ln q_{ij})\,,$$

where Q^{ii} is the (i, i) cofactor of Q and ϱ is the product of the non-zero eigenvalues of Q.

Proof. For an ergodic process we have

$$\int_0^T P(X_t = i)\,\mathrm{d}t = \frac{1}{\varrho} Q^{ii} T + o(T) \quad \text{as } T \to \infty\,.$$

Then

$$H = \lim_{T\to\infty} \frac{1}{T}\left(-\ln k - \sum_{\substack{i,j=1\\i\neq j}}^{s} \frac{1}{\varrho} Q^{ii} T q_{ij} \ln q_{ij} + \sum_{i=1}^{s} \frac{1}{\varrho} Q^{ii} T q_i\right),$$

$$H = \frac{1}{\varrho} \sum_{i,j=1}^{s} Q^{ii} q_{ij}(1 - \ln q_{ij})\,.$$

The cross-entropy (Kullback-Leibler information) for the observation interval $[0, T]$ is

$$I_T(Q; R) = E_Q\left(\ln \frac{f_{T,Q}(v)}{f_{T,R}(v)}\right),$$

where $f_{T,Q}(v)$ $(f_{T,R}(v))$ is the density under the infinitesimal generator $Q(R)$ and E_Q denotes the expectation under Q.

The cross-entropy $I(Q; R) = \lim_{T\to\infty} I_T(Q; R)/T$ measures "the distance" between the probability distribution of the process under Q and the distribution under R.

Property 2. For ergodic Markov pure-jump processes, the cross-entropy $I(Q; R)$ is finite,

$$I(Q, R) = \frac{1}{\varrho_Q}\left[\sum_{\substack{i,j=1\\i\neq j}}^{s} Q^{ii} q_{ij} \ln \frac{q_{ij}}{r_{ij}} - \sum_{i=1}^{s} Q^{ii}(q_i - r_i)\right].$$

The property follows by direct calculation.

Now, let us introduce the Markov pure-jump information source, with the infinitesimal generator Q.

$A = S \times (0, \infty)$ will be the alphabet of the source. Let $\mathscr{A} = \mathscr{P}(S) \otimes \mathscr{B}$ be the σ-field on A and $c \otimes \lambda$ the probability on $\mathscr{A}$.

We consider the product measurable space $(Y, \mathscr{Y}) = \bigotimes_{n=-\infty}^{\infty} (A_n, \mathscr{A}_n)$, with $A_n = A$, $\mathscr{A}_n = \mathscr{A}$ for every n. Let $\{\mathscr{T}^n : Y \to Y,\ n = 0, \pm 1, \ldots\}$ be the group of shift transformations on Y.

Let us consider $\mathscr{Y}_{s,t} = \bigotimes_{j=s}^{s+t-1} \mathscr{A}$, $\sigma_{s,t} = \bigotimes_{j=s}^{s+t-1} (c \otimes \lambda)$ and let σ be the product probability such that if restricted to $\mathscr{Y}_{s,t}$ it coincides with $\sigma_{s,t}$. We notice that σ is of the product (independent) type and is time stationary.

Now, let μ be a probability measure on $(Y, \mathscr{Y})$ and $\mu_{s,t}$ its restriction to $\mathscr{Y}_{s,t}$, such that $\mu_{s,t}$ is absolutely continuous with respect to $\sigma_{s,t}$. We denote by $f_{s,t}(v)$ the corresponding density and set

$$f_{0,n}(v) = f_{0,n}((z_0, t_0), \ldots, (z_{n-1}, t_{n-1})) =$$
$$= \pi_{z_0}\Big[\prod_{j=0}^{n-2} q_{z_j z_{j+1}} \exp(-q_{z_j} t_j)\Big]\, q_{z_{n-1}} \exp(-q_{z_{n-1}} t_{n-1}),$$

where $\{\pi_i,\ i \in S\}$ is the stationary distribution of the ergodic Markov pure-jump process under consideration.

The probability space $(Y, \mathscr{Y}, \mu)$ is called *the ergodic Markov pure-jump information source.*

Property 3. The information source $(Y, \mathscr{Y}, \mu)$ has the generalized equipartition property, that is

$$\lim_{n \to \infty} \left(-\frac{1}{n} \ln f_{0,n}\right) = H$$

in the sense of the convergence in the mean with respect to the measure μ.

Proof. $H_{0,n} = E_{(\sigma_{0,n})}(-f_{0,n} \ln f_{0,n}) < \infty$, where $E_{(\sigma_{0,n})}$ denotes the expectation under $\sigma_{0,n}$.

Since Perez's condition for the validity of the Statement (μ, σ) ([8]) is satisfied, the property follows as a consequence of Perez's result.

This convergence in the mean implies, in particular, the convergence in μ-probability of $-1/n \ln f_{0,n}$ to the entropy H of the considered ergodic process.

References

[1] *A. Albert:* Estimating the infinitesimal generator of a continuous time finite state Markov process. Ann. Math. Statist. *38* (1962), 727—753.

[2] *I. V. Basawa, B. L. S. Prakasa Rao:* Statistical inference for stochastic processes. Academic Press, London, 1980.

[3] *P. Billingsley:* Statistical inference for Markov processes. University of Chicago Press 1961.
[4] *J. L. Doob:* Stochastic processes. John Wiley, New York, 1953.
[5] *S. Guiaşu:* Information Theory with applications. McGraw-Hill Inc., Great Britain, 1977.
[6] *A. Perez:* Sur la théorie de l'information dans le cas d'un alphabet abstrait. Trans. First Prague Conf. Inf. Theory, Statist. Dec. Functions, Random Proc. (1956), Prague 1957, 209—243.
[7] *A. Perez:* Information Theory with an abstract alphabet. Generalized forms of McMillan's limit theorem for the case of discrete and continuous time. Teor. verojatn. i primen. T IV, (1959), 105—109.
[8] *A. Perez:* Extentions of Shannon-McMillan's limit theorem to more general stochastic processes. Trans. Third Prague Conf. Inf. Theory, Statist. Dec. Functions, Random Proc. (1962), Prague 1964, 545—574.

Souhrn

NĚKTERÉ INFORMAČNÍ VLASTNOSTI MARKOVOVSKÝCH ČISTĚ SKOKOVÝCH PROCESŮ

MONICA BAD DUMITRESCU

Je konstruována Boltzmannova entropie pro vnořený proces a je dokázáno, že ergodický markovovský čistě skokový informační zdroj má zobecněnou equipartiční vlastnost.

Резюме

НЕКОТОРЫЕ ИНФОРМАЦИОННЫЕ СВОЙСТВА ЧИСТО СКАЧКООБРАЗНЫХ ПРОЦЕССОВ МАРКОВА

MONICA BAD DUMITRESCU

Строится энтропия Болцмана для погруженного процесса и доказывается, что эргодический марковский чисто скачкообразный информационный источник обладает обобщённым свойством равномерного распределения.

Author's address: Department of Theory of Probabilities, Faculty of Mathematics, University of Bucharest, 14 Academiei St. 70109, Bucharest, Roumania.