Statistics and Probability Letters 141 (2018) 62–67

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Velocity formulae between entropy and hitting time for Markov chains

Michael C.H. Choi

Institute for Data and Decision Analytics, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, PR China

ARTICLE INFO

Article history:
Received 6 May 2018
Accepted 25 May 2018
Available online 7 June 2018

MSC:
60J10

Keywords:
Entropy
Hitting time
Commute time
Eigentime
Random walk

ABSTRACT

In the absence of acceleration, the velocity formula gives "distance travelled equals speed multiplied by time". For a broad class of Markov chains such as circulant Markov chains or random walk on complete graphs, we prove a probabilistic analogue of the velocity formula between entropy and hitting time, where distance is the entropy of the Markov trajectories from state i to state j in the sense of Ekroot and Cover (1993), speed is the classical entropy rate of the chain, and the time variable is the expected hitting time between i and j. This motivates us to define new entropic counterparts of various hitting time parameters such as average hitting time or commute time, and prove analogous velocity formulae and estimates between these quantities.

1. Introduction and main results

Suppose a particle moves from a point i and to another point j. In elementary physics, the classical velocity formula yields the distance between i and j is equal to the speed of the particle multiplied by the time taken, provided that the particle has no acceleration. For the class of Markov chains with constant row entropy, the main aim of this note is to prove analogues of the velocity formula where "distance" is replaced by various entropic quantities, "speed" is the entropy rate associated with the chain and "time" is substituted by various hitting time related parameters such as average hitting time and commute time.

Before we discuss our main results in Theorems 1.1 and 1.2, we first fix our notations and provide a quick review on the relevant background. Our notations follow closely those of Kafsi et al. (2013), Cover and Thomas (2006) and Ekroot and Cover (1993). Throughout this paper, we consider a discrete-time homogeneous irreducible finite Markov chain $X = (X_n)_{n\in\mathbb{N}}$ on state space $\mathcal{X}$ with transition matrix $P = (P_{i,j})_{i,j\in\mathcal{X}}$ and stationary distribution $\pi = (\pi_i)_{i\in\mathcal{X}}$. The entropy rate $H(X)$ of the Markov chain X is defined to be

$$H(X) := -\sum_{i,j\in\mathcal{X}} \pi_i P_{i,j} \log P_{i,j} = \sum_{i\in\mathcal{X}} \pi_i H(P_{i,\cdot}),$$

where $H(P_{i,\cdot}) := -\sum_{j\in\mathcal{X}} P_{i,j} \log P_{i,j}$ is the one-step local entropy at state i, and the usual convention of $0 \log 0 = 0$ applies. $H(X)$ can be broadly interpreted as the average entropy produced by a single step of X, and this interpretation is particularly useful in understanding our main results. Another entropic quantity that we are interested in is the so-called entropy of the Markov trajectories $H_{i,j}$ from state i to state j, as studied by Ekroot and Cover (1993) and Kafsi et al. (2013). Define a trajectory $T_{i,j}$ from i to j as a path with initial state i, final state j with no intervening state equal to j. We denote such

E-mail address: michaelchoi@cuhk.edu.cn.

https://doi.org/10.1016/j.spl.2018.05.026

trajectory by $T_{i,j} = ix_1x_2 \dots x_{k-1}j$. The probability of $T_{i,j}$ is $p(T_{i,j}) := P_{i,x_1}P_{x_1,x_2} \dots P_{x_{k-1},j}$. Writing $\mathcal{T}_{i,j}$ as the set of all possible trajectories from i to j, $H_{i,j}$ is then defined to be

$$H_{i,j} = H_{i,j}(X) := -\sum_{T_{i,j} \in \mathcal{T}_{i,j}} p(T_{i,j}) \log p(T_{i,j}).$$

We now move on to discuss a few hitting time related parameters of X. Define $\tau_j := \inf\{n \geqslant 0;\ X_n = j\}$ to be the first hitting time of the state j, and $\tau_j^+ := \inf\{n \geqslant 1;\ X_n = j\}$ to be the first return time of the state j. The usual convention applies in these definitions with $\inf \emptyset = \infty$.

In our main results below, we primarily consider Markov chains with constant row entropy. In essence, this means that the Markov chain has zero entropic acceleration as it moves from one state to another since each state gives the same local entropy $H(P_{i,\cdot})$.

Assumption 1.1 (*Constant Row Entropy*). We assume that X has constant row entropy, i.e. for all $i, j \in \mathcal{X}$, $H(P_{i,\cdot}) = H(P_{j,\cdot})$.

Examples of such Markov chains can be found in Section 3, where we apply our results to two-state Markov chains (Example 3.1), random walk on complete graphs (Example 3.2), rank-one Markov chains (Example 3.3) and simple random walks on n-cycle (Example 3.4). Note that random walk on regular graphs and circulant Markov chains (Avrachenkov et al., 2013) also fall into this category.

With the above notations and setting, we are now ready to state our main result. In a broad sense, it can be interpreted as the entropy of the trajectories from i to j equals the entropy per step times the mean hitting time between the two states.

Theorem 1.1 (*Velocity Formula Between Entropy and Hitting Time*). *Assume that X satisfies Assumption 1.1 with constant row entropy. For any $i, j \in \mathcal{X}$, we have*

$$H_{i,j} = \begin{cases} \mathbb{E}_i(\tau_j)H(X), & \text{for } i \neq j, \\ \mathbb{E}_i(\tau_i^+)H(X), & \text{for } i = j. \end{cases}$$

Note that for a deterministic Markov chain X, Theorem 1.1 trivially holds since $H_{i,j} = H(X) = 0$. Motivated by the relation between $H_{i,j}$ and $\mathbb{E}_i(\tau_j)$, we proceed to define a few new entropic parameters which are similar to their hitting time counterparts. We refer interested readers to Levin et al. (2009), Aldous and Fill (2002) and Montenegro and Tetali (2006) for excellent discussion on these parameters as well as their estimates.

Definition 1.1 (*Average Entropy H^{av}, Average Hitting Time t^{av} and Relaxation Time t^{rel}*). The average entropy and average hitting time are defined to be respectively

$$H^{av} = H^{av}(X) := \sum_{i,j \in \mathcal{X}} \pi_i \pi_j H_{i,j}, \quad t^{av} = t^{av}(X) := \sum_{i,j \in \mathcal{X}} \pi_i \pi_j \mathbb{E}_i(\tau_j).$$

For reversible Markov chain X, a closely related parameter is the relaxation time

$$t^{rel} := \frac{1}{1 - \lambda_2},$$

where $1 = \lambda_1 > \lambda_2 \geqslant \dots \geqslant \lambda_n$ are the eigenvalues of reversible P arranged in non-increasing order and $n := |\mathcal{X}|$.

Definition 1.2 (*Commute Entropy $H^c_{i,j}$ and Commute Time $t^c_{i,j}$*). For any $i, j \in \mathcal{X}$, the commute entropy and commute time between i and j are defined to be respectively

$$H^c_{i,j} = H^c_{i,j}(X) := H_{i,j} + H_{j,i}, \quad t^c_{i,j} = t^c_{i,j}(X) := \mathbb{E}_i(\tau_j) + \mathbb{E}_j(\tau_i).$$

We note that average entropy and average hitting time are both global parameters, while commute entropy and commute time are parameters associated with a given pair of states. In our second main result below, we give velocity formula between these parameters and carry a few results of hitting time to their entropic counterparts.

Theorem 1.2.

(1) *(Commute entropy velocity formula) For any $i \neq j \in \mathcal{X}$, we have*

$$H^c_{i,j} = t^c_{i,j} H(X).$$

Note that this holds in general and does not require the constant row entropy assumption.

(2) *(Average entropy velocity formula) Under the constant row entropy Assumption 1.1, we have*

$$H^{av} = (t^{av} + 1)H(X).$$

If in addition X is reversible, then

$$(t^{rel} + 1)H(X) \leqslant H^{av} \leqslant ((|\mathcal{X}| - 1)t^{rel} + 1) \log|\mathcal{X}|.$$