Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 124 (2014) 4149–4170

stochastic processes and their applications

www.elsevier.com/locate/spa

Exponential bounds for convergence of entropy rate approximations in hidden Markov models satisfying a path-mergeability condition

Nicholas F. Travers*

Department of Mathematics, University of California, Davis, United States

Received 5 February 2014; received in revised form 6 June 2014; accepted 9 July 2014
Available online 17 July 2014

Abstract

A hidden Markov model (HMM) is said to have *path-mergeable states* if for any two states i, j there exist a word w and state k such that it is possible to transition from both i and j to k while emitting w. We show that for a finite HMM with path-mergeable states the block estimates of the entropy rate converge exponentially fast. We also show that the path-mergeability property is asymptotically typical in the space of HMM topologies and easily testable.

Keywords: Hidden Markov model; Entropy rate; Exponential convergence; Path-mergeable

1. Introduction

Hidden Markov models (HMMs) are generalizations of Markov chains in which the underlying Markov state sequence (S_t) is observed through a noisy or lossy channel, leading to a (typically) non-Markovian output process (X_t). They were first introduced in the 50s as abstract mathematical models [15,7,6], but have since proved useful in a number of concrete applications, such as speech recognition [20,25,3,19] and bioinformatics [26,13,21,14,4].

* Correspondence to: Department of Mathematics, Technion—Israel Institute of Technology, Israel. Tel.: +972 585055757.
E-mail address: travers@tx.technion.ac.il.

http://dx.doi.org/10.1016/j.spa.2014.07.011
0304-4149/

One of the earliest major questions in the study of HMMs [6] was to determine the entropy rate of the output process

$$h = \lim_{t\to\infty} H(X_t|X_1, \ldots, X_{t-1}).$$

Unlike for Markov chains, this actually turns out to be quite difficult for HMMs. Even in the finite case no general closed form expression is known, and it is widely believed that no such formula exists. A nice integral expression was provided in [6], but it is with respect to an invariant density that is not directly computable.

In practice, the entropy rate h is instead often estimated directly by the finite length block estimates

$$h(t) = H(X_t|X_1, \ldots, X_{t-1}).$$

Thus, it is important to know about the rate of convergence of these estimates to ensure the quality of the approximation.

Moreover, even in cases where the entropy rate can be calculated exactly, such as for unifilar HMMs, the rate of convergence of the block estimates is still of independent interest. It is important for numerical estimation of various complexity measures, such as the excess entropy and transient information [11], and is critical for an observer wishing to make predictions of future output from a finite observation sequence $X_1, \ldots, X_t$. It is also closely related to the rate of memory loss in the initial condition for HMMs, a problem that has been studied extensively in the field of filtering theory [27,1,2,22,8,12,9]. Though, primarily in the case of continuous real-valued outputs with Gaussian noise or similar, rather than the discrete case we study here.

No general bounds are known for the rate of convergence of the estimates $h(t)$, but exponential bounds have been established for finite HMMs (with both a finite internal state set $\mathcal{S}$ and finite output alphabet $\mathcal{X}$) under various positivity assumptions. The earliest known, and mostly commonly cited, bound is given in [5], for finite functional HMMs with strictly positive transition probabilities in the underlying Markov chain. A somewhat improved bound under the same hypothesis is also given in [17]. Similarly, exponential convergence for finite HMMs with strictly positive symbol emission probabilities and an aperiodic underlying Markov chain is established in [24] (both for state-emitting and edge-emitting HMMs) using results from [22].

Without any positivity assumptions, though, things become substantially more difficult. In the particular case of unifilar HMMs, we have demonstrated exponential convergence in our recent work on synchronization with Crutchfield [29,28]. Also, in Ref. [16] exponential convergence is established under some fairly technical hypotheses in studying entropy rate analyticity. But, no general bounds on the convergence rate of the block estimates $h(t)$ have been demonstrated for all finite HMMs.

Here we prove exponential convergence for finite HMMs (both state-emitting and edge-emitting) satisfying the following simple path-mergeability condition: *For each pair of distinct states i, j there exist a word w and state k such that it is possible to transition from both i and j to k while emitting w*. We also show that this condition is easily testable (in computational time polynomial in the number of states and symbols) and asymptotically typical in the space of HMM topologies, in that a randomly selected topology will satisfy this condition with probability approaching 1, as the number of states goes to infinity. By contrast, the positivity conditions assumed in [5,17,24], as well as the unifilarity hypothesis we assume in [29,28], are satisfied