INFERENCE ON THE STATE DISTRIBUTION IN PERIODIC HIDDEN MARKOV MODELS

Carlina C. Feldmann[1], Sina Mews[1], Rouven Michels[1] and Roland Langrock[1]

[1] Department of Business Adminstration and Economics, Bielefeld University, (e-mail: carlina.feldmann@uni-bielefeld.de, sina.mews@uni-bielefeld.de, r.michels@uni-bielefeld.de, roland.langrock@uni-bielefeld.de)

ABSTRACT: We present an exact solution for the time-varying state distribution in hidden Markov models (HMMs) with periodic state-switching dynamics. In a case study using African elephant data, the approach is shown to be superior to commonly applied alternatives.

KEYWORDS: Markov chain, movement ecology, periodic stationarity.

1 Introduction

When inferring latent states and their dynamics from an observed time series, periodic effects such as diel variation or seasonality are often of primary interest. In applications such as movement ecology or climatology, hidden Markov models (HMMs) with cyclic components are commonly used to address periodic variation in the latent state process (see, e.g., Nagel *et al.*, 2021). Inference then often focuses on the periodically varying probabilities of occupying the different states. These can, in principle, be taken as the empirical distribution of states per time point, as obtained using decoding algorithms such as Viterbi (see, e.g., Schwarz *et al.*, 2021).

To avoid the noise associated with this approach, especially for shorter time series, it may however be desirable to instead evaluate the time-varying state distribution as implied under the fitted model. Here we show how to exploit the periodic stationarity of corresponding HMMs to arrive at an analytic solution for the time-varying state distribution. In a case study on elephant movement, we demonstrate the superiority of our approach over commonly applied alternatives.

2 Methods

We consider an HMM comprising a state-dependent process $\{X_t\}_{t=1,\dots,T}$ (where X_t can be a vector) and a latent state process $\{S_t\}_{t=1,\dots,T}$, with S_t selecting which of N possible component distributions generates X_t. The state process $\{S_t\}$ is assumed to be an N–state Markov chain, characterised by its initial state distribution and the time-varying transition probability matrix (t.p.m.)

$$\Gamma^{(t)} = (\gamma_{ij}^{(t)}), \text{ with } \gamma_{ij}^{(t)} = \Pr(S_t = j | S_{t-i} = i),$$

$t = 1, \dots, T$. We consider a setting with periodically varying state-switching dynamics, such that

$$\Gamma^{(t)} = \Gamma^{(t+L)} \tag{1}$$

for all $t = 1, \dots, T$, with L denoting the length of a cycle. For hourly data and $N = 2$, we could for example model time-of-day variation ($L = 24$) as

$$\operatorname{logit}\big(\gamma_{ij}^{(t)}\big) = \beta_1^{(ij)} \sin\left(\frac{2\pi t}{24}\right) + \beta_2^{(ij)} \cos\left(\frac{2\pi t}{24}\right), \text{ for } i \neq j. \tag{2}$$

The interpretation of such transition probabilities as functions of time can be tedious, especially when $N > 2$. Therefore, it has become common practice to instead consider a simpler summary statistic, namely the (periodically varying) distribution of the states at time t,

$$\delta^{(t)} = \big(\Pr(S_t = 1), \dots, \Pr(S_t = N)\big),$$

as a function of time $t = 1, \dots, L$. The latter is usually approximated by the hypothetical stationary distribution $\rho^{(t)}$ of the Markov chain that would result if the t.p.m. was homogeneous with $\Gamma = \Gamma^{(t)}$, which is the solution to $\rho^{(t)} = \rho^{(t)}\Gamma$ subject to $\sum_{i=1}^N \rho_i^{(t)} = 1$ (Patterson *et al.*, 2009). This approximation of $\delta^{(t)}$ will in general be biased because it ignores the preceding process dynamics as implied by $\Gamma^{(t-1)}, \Gamma^{(t-2)}, \dots$ and instead pretends that the process has been following the dynamics as implied by a constant $\Gamma^{(t)}$ for a considerable time.

However, for periodically inhomogeneous Markov chains as defined in (1), there is in fact no need for such an approximation. To see this, consider for fixed t the thinned Markov chain $S_t, S_{t+L}, S_{t+2L}, \dots$, which is homogeneous with constant t.p.m.

$$\tilde{\Gamma}_t = \Gamma^{(t+1)} \cdot \ldots \cdot \Gamma^{(t+L)}.$$

Provided that this thinned Markov chain is irreducible, it has a unique stationary distribution $\delta^{(t)}$, which is the solution to

$$\delta^{(t)} = \delta^{(t)} \tilde{\Gamma}_t$$

(see also Ge *et al.*, 2006 and Kargapolova & Ogorodnikov, 2012). Provided that the Markov chain starts in its stationary distribution, $\delta^{(t)}$ is the state distribution at time t we are interested in (and otherwise it will be at least approximately correct as the thinned Markov chain will converge to its stationary distribution).

3 Case study: elephant movement

We consider a complete movement track of an African elephant with hourly GPS data between October 2008 and June 2009. Based on consecutive locations, we calculate the Euclidean step lengths as well as the turning angles and model them in a 3-state HMM with gamma and von Mises distributions, respectively. To investigate diel variation in the state-switching dynamics we model the transition probabilities as trigonometric functions of the time of day (see Equation 2). The fitted model features an "encamped" state with short step lengths and frequent reversals in direction (state 1), an "exploratory" state with higher persistence in direction and medium step lengths (state 2), and a "travelling" state with highly directed and fast movement (state 3).

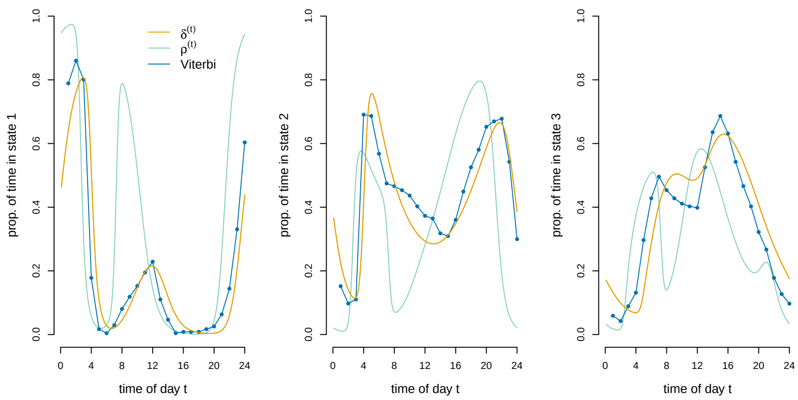

Figure 1. *Proportion of time spent in each state according to the model-implied periodic stationary distribution, the approximated stationary distribution, and the Viterbi state decoding.*

Based on the fitted HMM, we derive the proportions of time spent in each state using the model-implied periodic stationary distribution $\delta^{(t)}$, the approximated stationary distribution $\rho^{(t)}$, and the Viterbi-decoded states. The corresponding results are compared in Figure 1. The hypothetical stationary distribution $\rho^{(t)}$ differs greatly from the exact solution $\delta^{(t)}$ and is therefore a poor approximation in this example. Concerning the proportion of time spent in each state obtained using the Viterbi algorithm, the results are similar to the analytically derived periodic stationary distribution $\delta^{(t)}$. The advantage of the latter, however, is that it is less affected by noise in the data and instead offers a smooth function of time, even for short time series.

References

GE, HAO, JIANG, DA-QUAN, & QIAN, MIN. 2006. A Simple Discrete Model of Brownian Motors: Time-periodic Markov Chains. *Journal of Statistical Physics*, **123**(4), 831–859.

KARGAPOLOVA, N. A., & OGORODNIKOV, V. A. 2012. Inhomogeneous Markov chains with periodic matrices of transition probabilities and their application to simulation of meteorological processes. *Russian Journal of Numerical Analysis and Mathematical Modelling*, **27**(3), 213–228.

NAGEL, REBECCA, MEWS, SINA, ADAM, TIMO, STAINFIELD, CLAIRE, FOX-CLARKE, CAMERON, TOSCANI, CAMILLE, LANGROCK, ROLAND, FORCADA, JAUME, & HOFFMAN, JOSEPH I. 2021. Movement patterns and activity levels are shaped by the neonatal environment in Antarctic fur seal pups. *Scientific Reports*, **11**(1), 14323.

PATTERSON, TOBY A., BASSON, MARINELLE, BRAVINGTON, MARK V., & GUNN, JOHN S. 2009. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *Journal of Animal Ecology*, **78**(6), 1113–1123.

SCHWARZ, JONAS F. L., MEWS, SINA, DERANGO, EUGENE J., LANGROCK, ROLAND, PIEDRAHITA, PAOLO, PÁEZ-ROSAS, DIEGO, & KRÜGER, OLIVER. 2021. Individuality counts: A new comprehensive approach to foraging strategies of a tropical marine predator. *Oecologia*, **195**(2), 313–325.

ZUCCHINI, W., MACDONALD, I.L., & LANGROCK, R. 2016. *Hidden Markov Models for Time Series: An Introduction Using R*. Boca Raton, FL: Chapman and Hall/CRC Press.