Modelling three-dimensional movements using multivariate hidden Markov models

Project supervisor: Timo Adam[1]

[1]tiad@math.ku.dk

April 17, 2023

Picture by: Andy Morffew

1 Introduction

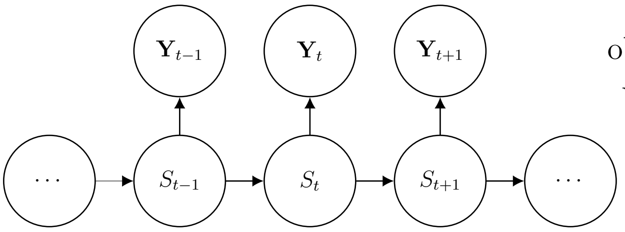
In recent years, HMMs have emerged as a versatile class of statistical models for time series where the observed variables are driven by hidden states. In ecological applications, HMMs are often used to infer animals' behaviours (e.g., resting, foraging, and travelling) and their various drivers from tracking data. In this project, you will use multivariate HMMs to model three-dimensional movements (i.e., horizontal and vertical movements) of eagles.

In their basic form, HMMs comprise two stochastic processes, one of which is hidden and the other is observed:

- The (hidden) state process, which is denoted by $\{S_t\}_{t=1,\dots,T}$, is modelled by a discrete-time, N-state Markov chain with initial probabilities $\delta_i = \Pr(S_1 = i), i = 1, \dots, N$, and transition probabilities $\gamma_{ij} = \Pr(S_{t+1} = j | S_t = i), i, j = 1, \dots, N, t = 1, \dots, T-1$.

- The (observed) state-dependent process, which is denoted by $\{\mathbf{Y}_t\}_{t=1,\dots,T}$, is modelled by N state-dependent distributions with P-dimensional probability density function (p.d.f.) $f_i(\mathbf{y}_t), i = 1, \dots, N$, where P denotes the number of variables considered.

The state that is active at time t, S_t, determines which state-dependent distribution generates the corresponding observations:



In practice, it is often assumed that the different variables (e.g., horizontal and vertical distances travelled) are conditionally independent of each other, given the states (this has the advantage that the state-dependent p.d.f. can then be written as a product of P univariate densities, e.g., the p.d.f. of a gamma distribution for horizontal distances travelled, which can only be positive, and the p.d.f. of a normal distribution for vertical distances travelled, which can be positive and negative). In this project, you will investigate possible consequences of making this assumption and extend the basic HMM by allowing for variables that are still correlated under the states (this requires a multivariate state-dependent p.d.f., e.g., the p.d.f. of a bivariate normal-gamma distribution).

A comprehensive introduction to HMMs and their application in ecology can be found in McClintock *et al.* (2020). A useful R package that implements HMMs for animal movement data, momentuHMM, is explained in McClintock and Michelot (2018). Further details on HMMs can be found in Zucchini *et al.* (2016) (accompanying R code is provided at http://hmms-for-time-series.de/second/appendix-a/ZMLcode.txt).

2 Objectives and preliminary plan of the project

In this project, you will:

- Study the relevant literature to learn about HMMs and their application in ecology, particularly animal movement modelling.

- Fit basic HMMs to eagle data and interpret the results, particularly with regard to behaviours exhibited by the animals, and investigate the model fit.

- Perform a simulation study to investigate possible consequences of assuming that the different variables are conditionally independent of each other, given the states. Therefore, you will simulate data that are similar to the original data but where the different variables exhibit different degrees of correlation (e.g., no correlation, some correlation, and high correlation).

- Extend the basic HMM by allowing for variables that are still correlated under the states, fit it to eagle data, interpret the results, and investigate the model fit.

(Optional) Depending on your interests, you can extend the above plan in various ways. For instance, you can include other variables (e.g., turning angles), incorporate covariates into the state process, or use non-parametric methods (e.g., splines) for the state-dependent distributions.

3 Data

In this project, you will work with three-dimensional movement data of eagles (see Pirotta *et al.* (2018) for an example analysis of such data) that will be provided at the beginning of the project.

4 Prerequisites

You should be familiar with Markov chains, common distributions, programming with R (or in a different environment), maximum likelihood estimation, and numerical optimisation (using optimisers such as nlm or optim in R).

References

McClintock, B.T., Langrock, R., Gimenez, O., Cam, E., Borchers, D.L., Glennie, R., and Patterson, T.A. (2020). Uncovering ecological state dynamics with hidden Markov models. *Ecology Letters*, 23(12), 1878–1903.

McClintock, B.T., and Michelot, T. (2018). momentuHMM: R package for generalized hidden Markov models of animal movement. *Methods in Ecology and Evolution*, 9(6), 1518–1530.

Pirotta, E., Katzner, T., Miller, T.A., Duerr, A.E., Braham, M.A., and New, L. (2018). State-space modelling of the flight behaviour of a soaring bird provides new insights to migratory strategies. *Functional Ecology*, 32(9), 2205–2215.

Zucchini, W., MacDonald, I.L., and Langrock, R. (2016): *Hidden Markov models for time series: an introduction using R, 2nd Edition.* Chapman and Hall/CRC, Boca Raton.