Guest editor's introduction to the special issue on "Hidden Markov Models: Theory and Applications"

Jan Bulla¹ · Roland Langrock² · Antonello Maruotti³

Received: 2 July 2019 / Accepted: 12 July 2019 / Published online: 17 July 2019
© Sapienza Università di Roma 2019

1 Introduction

In the 1970s, hidden Markov models (HMMs) gained prominence as useful tools for speech recognition, i.e. for the automatic translation of speech into text. In such a setting, an HMM would consider segmented speech signals, for example obtained by spectral analysis, to be noisy versions of the actual phonemes spoken, which are to be inferred by the computer. The class of HMMs became very successful in this area due to their appealing combination of immense versatility on the one hand, which allows the models to be tailored to various types of sequential data, and the relative mathematical ease coupled with the availability of efficient training and decoding algorithms on the other hand. While originally applied predominantly by engineers and computer scientists, the wide applicability and tractability has resulted in HMMs becoming increasingly popular also in applied statistical modelling. In particular, since the 1990s, HMMs have been successfully applied to diverse statistical problems in disciplines such as finance [17], psychology [19], medicine [12], volcanology [5], ecology [8], bioinformatics [9], and marketing [14].

The mathematical theory of HMMs, and in particular the various methods to fit an HMM to a given data set—most notably maximum likelihood estimation, either by direct numerical optimisation or using the expectation-maximisation algorithm [21], and within a Bayesian framework via Markov chain Monte Carlo [18]—are now well established in the statistical literature. Indeed, over the last two decades, we have witnessed HMMs establishing themselves within the toolboxes of many applied researchers. To some extent this was also driven by the provision of specialised HMM software (e.g. the R packages `depmixS4` by Visser and Speekenbrink [20], `msm` by Jackson [11], `HiddenMarkov` by Harte [10], `LMest` by Bartolucci et al. [4], `moveHMM` by Michelot et al. [16], and `momentuHMM` by McClintock and Michelot [15]). Nevertheless, some disciplines are only beginning to discover the utility of HMMs. Furthermore, there is still a need to adapt the HMM framework to increasingly

✉ Roland Langrock
roland.langrock@uni-bielefeld.de

¹ University of Bergen, Bergen, Norway

² Bielefeld University, Bielefeld, Germany

³ Libera Università Maria Ss Assunta, Rome, Italy

large and complex data sets, which in recent years have once more fuelled methodological research into HMMs.

The aim of this special issue is to highlight several current directions of statistical research related to HMMs, both regarding novel model formulations and estimation approaches, but also innovative types of applications. In what follows, we briefly summarise the contribution of each article in this special issue.

2 Overview of the articles in this special issue

For many years now, the collection and thus the analysis of electronic health records has seen major progress. Such data hold great promise for medical research as well as personalised treatments. Patient-specific data are typically sequential, and with the underlying health state only indirectly being observed, HMMs are natural models for (medical) classification as well as forecasts of disease progression. In this special issue, Amoros et al. [3] present an application of HMMs for cancer surveillance, thereby showcasing the great potential of these models for complex health record data. In particular, the paper addresses the temporal irregularity of the sampling by using a continuous-time HMM formulation, and addresses patient heterogeneity using a hierarchical model with random effects, which is fitted in a Bayesian framework. While tailored to the application at hand, the concepts presented in this paper are widely applicable in health research.

Cole [7] describes three different approaches for investigating identifiability and detecting parameter redundancy in the context of HMMs. Given the fairly complex structure of HMMs in many applied settings, with not only two stochastic processes but also different dependence structures that can be assumed, the issue of estimability in general is clearly of much interest when working with HMMs. Two of the procedures discussed in Cole [7], the Hessian method and the log-likelihood profile method, are based on numerical techniques, while the third uses symbolic algebra. The paper provides guidance on the practical feasibility of these approaches.

Lember et al. [13] discuss global state decoding for HMMs fitted within a Bayesian framework. The Viterbi algorithm provides an efficient recursive scheme for finding the most likely sequence of states to have given rise to the observations, conditional on a *point* estimate of the model's parameter vector. Lember et al. [13] give an overview of corresponding global state decoding approaches in a Bayesian framework, where the outcome is a (posterior) *distribution* (rather than a point estimate) of the model's parameter vector. In particular, they suggest a new decoding approach in the Bayesian framework, the segmentation expectation-maximisation algorithm, which is designed to directly yield a Viterbi path, i.e. without the need to first of all fit the model. The new algorithm regards the model parameters as nuisance parameters, thus avoiding the usual two-stage approach to decoding the states (first fitting the model, then applying Viterbi to the point estimate).

While non-parametric inference in HMMs has been studied intensively for the case of continuous-valued time series, corresponding work on discrete-valued time series has so far been lacking. Adam et al. [2] fill this gap by proposing a maximum penalised likelihood approach for fitting HMMs to discrete-valued time series without making any distributional assumptions for the observed state-dependent process. In simulations and case studies, it is shown that the new approach provides a good balance between completely unrestricted estimation of the probability mass functions, which often results in overfitting, and parametric model formulations using say the Poisson or the negative binomial, which are often too

inflexible. The new approach is applicable effectively to any univariate discrete time series with at least ordinal data, and is available to users within the R package `countHMM` [1].

Finally, Chiappa and Paquet [6] tackle the problem of identifying and simultaneously tracking multiple moving objects from video data, i.e. pixel arrays, in an unsupervised setting. While recurrent neural networks have been very successfully used in such learning tasks, the corresponding methods also have some shortcomings, including a lack of interpretability as well as limitations when dealing with missing data. Chiappa and Paquet [6] embed the learning task at hand in a state-space model formulation, thereby enabling probabilistic reasoning. In their modelling framework, linear latent processes describe the positions and velocities of the multiple objects to be tracked, with the video being regarded as a sequence of noisy observations of the actual positions.

References

1. Adam, T.: `countHMM`: Penalized estimation of flexible hidden Markov models for time series of counts, R package version 0.1.0. <https://CRAN.R-project.org/package=countHMM> (2019)
2. Adam, T., Langrock, R., Weiß, C.H.: Penalized estimation of flexible hidden Markov models for time series of counts. *METRON* (2019). <https://doi.org/10.1007/s40300-019-00153-6>
3. Amoros, R., King, R., Toyoda, H., Kumada, T., Johnson, P.J., Bird, T.G.: A continuous-time hidden Markov model for cancer surveillance using serum biomarkers with application to hepatocellular carcinoma. *METRON* (2019). <https://doi.org/10.1007/s40300-019-00151-8>
4. Bartolucci, F., Pandolfi, S., Pennoni, F.: LMest: An R package for latent Markov models for longitudinal categorical data. *J. Stat. Softw.* **81**(4), 1–38 (2017)
5. Bebbington, M.: Identifying volcanic regimes using hidden Markov models. *Geophys. J. Int.* **171**(2), 921–942 (2007)
6. Chiappa, S., Paquet, U.: Unsupervised separation of dynamics from pixels. *METRON* (2019). <https://doi.org/10.1007/s40300-019-00155-4>
7. Cole, D.: Parameter redundancy and identifiability in hidden Markov models. *METRON* (2019). <https://doi.org/10.1007/s40300-019-00156-3>
8. DeRuiter, S.L., Langrock, R., Skirbutas, T., Goldbogen, J.A., Calambokidis, J., Friedlaender, A.S., Southall, B.L.: A multivariate mixed hidden Markov model for blue whale behaviour and responses to sound exposure. *Ann. Appl. Stat.* **11**(1), 362–392 (2017)
9. Eddy, S.R.: Profile hidden Markov models. *Bioinformatics* **14**(9), 755–763 (1998)
10. Harte, D.: `HiddenMarkov`: Hidden Markov models, R package version 1.8-11. <https://CRAN.R-project.org/package=HiddenMarkov> (2017)
11. Jackson, C.H.: Multi-state models for panel data: The `msm` Package for R. *J. Stat. Softw.* **38**(8), 1–29 (2011)
12. Langrock, R., Swihart, B.J., Caffo, B.S., Crainiceanu, C.M., Punjabi, N.M.: Combining hidden Markov models for comparing the dynamics of multiple sleep electroencephalograms. *Stat. Med.* **32**(19), 3342–3356 (2013)
13. Lember, J., Gasbarra, D., Koloydenko, A., Kuljus, K.: Estimation of Viterbi path in Bayesian hidden Markov models. *METRON* (2019). <https://doi.org/10.1007/s40300-019-00152-7>
14. Maruotti, A., Bulla, J., Mark, T.: Assessing the influence of marketing activities on customer behaviors: A dynamic clustering approach. *METRON* **77**(1), 19–42 (2019)
15. McClintock, B.T., Michelot, T.: `momentuHMM`: R package for generalized hidden Markov models of animal movement. *Methods Ecol. Evol.* **9**(6), 1518–1530 (2018)
16. Michelot, T., Langrock, R., Patterson, T.A.: `moveHMM`: An R package for the statistical modelling of animal movement data using hidden Markov models. *Methods Ecol. Evol.* **7**(11), 1308–1315 (2016)
17. Rydén, T., Teräsvirta, T., Åsbrink, S.: Stylized facts of daily return series and the hidden Markov model. *J. Appl. Econom.* **13**(3), 217–244 (1998)
18. Scott, S.L.: Bayesian methods for hidden Markov models: Recursive computing in the 21st century. *J. Am. Stat. Assoc.* **97**(457), 337–351 (2002)
19. Visser, I.: Seven things to remember about hidden Markov models: A tutorial on Markovian models for time series. *J. Math. Psychol.* **55**(6), 403–415 (2011)
20. Visser, I., Speekenbrink, M.: `depmixS4`: An R package for hidden Markov models. *J. Stat. Softw.* **36**(7), 1–21 (2010)

21. Zucchini, W., MacDonald, I.L., Langrock, R.: Hidden Markov Models for Time Series—An Introduction Using R, 2nd edn. Chapman and Hall/CRC, Boca Raton (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.