Journal of Statistical Software

*August 2017, Volume 80, Book Review 1.* *doi: 10.18637/jss.v080.b01*

Reviewer: Abdolvahab Khademi
University of Massachusetts

**Hidden Markov Models for Time Series: An Introduction Using R (2nd Edition)**

Walter Zucchini, Iain L. MacDonald, Roland Langrock
Chapman &Hall/CRC, Boca Raton, 2016.
ISBN 9781482253832. xxviii+370 pp. USD 89.95 (H).
https://www.crcpress.com/p/book/9781482253832

The distribution of estimable parameters of observed data in some stochastic processes may be influenced by another yet unobservable process called the parameter process. Such a parameter process can be serially dependent. A tractable mathematical model to account for serially dependent processes is the Markov chain process, which can be applied to both observed and unobserved processes. A Markov chain process used to model an unobserved serially dependent process is commonly called hidden Markov model (HMM), which is used in fields as diverse as signal processing, machine learning (e.g., handwriting recognition), environmental sciences, economics, and longitudinal data in social or medical sciences. Given the importance and widespread use of HMM, several books and references have been published for students, practitioners, and researchers. *Hidden Markov Models for Time Series: An Introduction Using R (2nd Edition)* is a book on HMM dedicated to time series data for researchers who cannot use a standard time series model for their data, including researchers in animal behavior, epidemiology, finance, hydrology, and sociology, among others.

The book is structured in three parts. Part I of the book (eight chapters), *Model Structure, Properties and Methods*, presents the theoretical and mathematical context for HMM. In Part II (five chapters), *Extensions*, the authors introduce additional complexities to the standard HMM, including continuous variables, covariates, semi-Markov models, and longitudinal data. Applications of hidden Markov models are presented in eleven chapters in Part III of the book, *Applications*. Each chapter in the book ends with exercises on the application of HMM, proofs, data analysis, and practice writing R functions. The authors use a recurring example (earthquake counts) to aid understanding HMM. The book has two appendices that present R code and some mathematical proofs and results. A dedicated website to the book provides the R code used in the book and errata for both the first and second editions of the book. Companion data sets are provided on the book's first edition link in the dedicated website.

In Chapter 1, *Preliminaries: Mixtures and Markov Chains*, the authors introduce independent mixture distributions and Markov chains. An example on modeling earthquakes using Poisson distributions is used to introduce the context where a mixture model is appropriate to capture all observed and underlying variations in data. This is an intuitive introduction to Markov chains and the concept of a latent heterogeneous process. Next, Markov chains are introduced in a very clear and detailed manner along with a simple example to show the mathematics of estimating transition probabilities and higher order Markov chains. The chapter ends with extensive exercises on proof, data analysis, and practice in writing or reworking R functions (often with some tips by the authors).

After justifying the need for unobserved processes, the authors introduce HMM in Chapter 2, *Hidden Markov Models: Definition and Properties*, in two sections. In the first section, HMM is defined, exemplified, and mathematically explained in terms of marginal distributions, moments, and covariance. In the second section, the likelihood for an $m$-state HMM is introduced for estimation purposes and shown for a two-state Bernoulli HMM process. Likelihood estimation and properties for missing data and interval-censored data are also presented.

Estimation of HMM parameters is further extended in Chapter 2, *Estimation by Direct Maximization of the Likelihood.* Numerical underflow, constraints on parameters, and multiple local maxima are stated as potential problems in maximum likelihood estimation, and solutions are proposed to overcome them, such as the scaling algorithm, reparametrization, and embedding a continuous-time Markov model. Computation of standard errors and confidence intervals through Hessian matrix and bootstrap are also elaborated. *Estimation by the EM Algorithm* (or Baum-Welch algorithm) is discussed in Chapter 4. In the context of HMM, the EM algorithm is used when the Markov chain is homogenous but nonstationary. In this chapter, the authors elaborate on the backward and forward probability vectors and their properties. The EM algorithm is presented in both general form and in the HMM context, with examples for discrete and continuous distributions.

The text transitions from theoretical to more practical in Chapter 5, *Forecasting, Decoding and State Prediction.* In this chapter, the authors discuss how to derive conditional distributions for an observation at time $t$ given all other observations. The authors present forecast distributions for a future observation conditioned on past observations for an HMM. Decoding, which is the extraction of information from hidden states, is presented next. Both local and global decoding are explained with an example. Finally, the authors present the application of HMM in classification problems.

In Chapter 6, *Model Selection and Checking*, the authors present AIC, BIC, and psuedo-residuals as tools for checking model fit. In Chapter 7, *Bayesian Inference for Poisson-Hidden Markov Models*, the audience learn how to perform an HMM in the Bayesian framework using the Gibbs sampler. Parallel sampling is presented as a model selection method in the Bayesian approach to HMM. Chapter 8, *R Packages* introduces the R packages commonly used in HMM.

Part II of the book, comprising five chapters, is the major addition and change to the first edition of the book. In this part, the basic HMM model gains additional complexities through continuous-valued state processes, hidden semi-Markov models, and Markov models for longitudinal data.

In Chapter 9, *HMMs with General State-Dependent Distributions*, general univariate, multinomial, and multivariate state-dependent distributions are discussed. HMMs for unbounded counts, binary data, bounded counts, continuous-valued series, proportions, and HMMs for circular-valued series are presented in the univariate section of the chapter. Multinomial, categorical, and compositional HMM models are discussed in the second section of the chapter.

In the multivariate section, the authors discuss longitudinal conditional independence and contemporaneous conditional independence.

HMMs with covariates are explained in Chapter 10, *Covariates and Other Extra Dependencies*, with a focus on covariate time series. Covariates are examined in both state-dependent distributions and transition probabilities. In the state-dependent case, HMMs are modified to accommodate the effect of covariates. In the transition probability case, covariates are accommodated in the HMM model by dropping the homogeneity assumption of the Markov chain and instead assuming that transition probabilities are functions of covariates. In another extension to the standard HMM model, the hidden Markov chain is treated as second order.

The primary variable type in the text is discrete data, but in Chapter 11, *Continuous-Valued State Processes*, the authors extend the HMM to allow for continuous processes where the number of states becomes infinitely large. The mathematics of likelihood estimation through numerical integration and discretization is presented. An example is used to show the discretization of continuous-valued process.

When the latent Markov process is allowed to have a lag larger than one, the number of parameters to estimate increases. Also, in case of first-order latent Markov models, the dwell time is constrained to equal 1, which may be restrictive. Hidden semi-Markov models are one alternative method to solve these two problems. *Hidden Semi-Markov Models and Their Representation as HMMs* in Chapter 12 introduces the audience to this solution. Modeling panel or longitudinal data is presented in Chapter 13, *HMMs for Longitudinal Data*. Longitudinal data are introduced along with random and constant parameters across component series for both continuous-valued and discrete variables. This chapter ends Part II of the book.

In Part III of the book, *Applications*, the authors show how the HMM framework can be applied to various fields where a latent parameter process underlies the stochastic process. Chapter 14, *Introduction to Applications*, provides a reference table for the applications of HMM and related published works in those areas. Chapters 15–24 present some applications of HMM in fields as diverse as neuroscience, environment, animal behavior, finance, and public health.

*Hidden Markov Models for Time Series: An Introduction Using* R *(2nd Edition)* is a well-balanced text on the theory, application, and computation of the HMMs. Although the title of the book reads *Introduction*, the book is essentially written for advanced students, researchers, and practitioners who are well-grounded in probability theory, mathematics, and advanced statistics. The background knowledge plays an important role in not only understanding the text but also doing the chapter exercises. In addition, a working knowledge of R programming language is required to understand, rework, and apply the functions in the text and exercises.

The book is organized in twenty-four dense chapters. Although the authors have tried their best to simplify the concepts, the subject matter is inherently challenging and the readers are encouraged to read them in depth and in times to obtain a clear understanding of HMMs. However, Part III of the book (*Applications*) can be read in tandem with Part I chapters to get a faster understanding of the concepts. Another feature of the book that remarkably helps clarify the concepts and the algorithms is the set of R-coded functions that represent some of the algorithms and mathematical functions in the text. Working through the R code can significantly help the readers to understand the book and be able to apply the models to their own data.

As for the computation aspect of their work, the authors have dedicated an entire chapter for the introduction of R packages used in the book for modeling HMMs. This is a very convenient feature for the researchers who may not be familiar with the common tools in HMM computation.

Overall, this a dense but rich, and very well-balanced book that presents the theory and mathematics of HMMs to the advanced student in fields where observations are collected over time with a substantive underlying parameter process. In addition, the book provides chapter-length examples of how and where HMMs can be implemented. Researchers new to HMM may find the applications in the book inspiring and will look at their data with a new perspective.

**Reviewer:**

Abdolvahab Khademi
University of Massachusetts
Department of Mathematics and Statistics
Amherst MA 01002, United States of America
E-mail: khademi@math.umass.edu

*Journal of Statistical Software* http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
August 2017, Volume 80, Book Review 1 *Published:* 2017-08-29
doi:10.18637/jss.v080.b01