Journal of Multivariate Analysis 140 (2015) 19–30

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Multiple hidden Markov models for categorical time series

R. Colombi [a], S. Giordano [b,*]

[a] Department of Engineering, University of Bergamo, Italy

[b] Department of Economics, Statistics and Finance, University of Calabria, Italy

ARTICLE INFO

*Article history:*
Received 31 March 2014
Available online 12 April 2015

*AMS subject classifications:*
68R10
60J10
62H99

*Keywords:*
Conditional independence
Granger noncausality
Graphical models
Marginal models
Markov properties
Multivariate Markov chains

ABSTRACT

We introduce multiple hidden Markov models (MHMMs) where a multivariate categorical time series depends on a latent multivariate Markov chain. MHMMs provide an elegant framework for specifying various independence relationships between multiple discrete time processes. These independencies are interpreted as Markov properties of a mixed graph and a chain graph associated respectively to the latent and observation components of the MHMM. These Markov properties are also translated into zero restrictions on the parameters of marginal models for the transition probabilities and the distributions of observable variables given the latent states.



## 1. Introduction

In several applications involving time series, it is interesting to describe how the evolution of variables over time depends on latent characteristics or the focus may be on the dynamics of unobservable characteristics measured by variables observed on consecutive time occasions. These issues are addressed by hidden Markov models and a widespread application of them has occurred in several fields such as speech recognition, signal processing, digital communication, biology, reliability, etc., standard references are [3,13,15,16], among others.

Basically, a hidden Markov model assumes that an observable time series depends on a latent Markov chain in such a way that the joint process is also Markovian. Note that, throughout the paper, the term observable never refers to the observability property of state-space models (see [18] and [21], among others), but it is used with its literal meaning to distinguish the variables that can be directly observed from the latent ones.

In this work, we focus on hidden Markov models in discrete time with a multivariate categorical process depending on a multivariate latent chain. In these models, several variables are observed and their distribution is supposed to be affected by one or more latent variables. For the multidimensionality of latent and observation components, we will refer to these models as multiple hidden Markov models (MHMMs).

The MHMM can be seen as a variant of the usual hidden Markov model (HMM) that allows modeling opportunities not available in the standard approach with the same clarity, interpretability and parsimony.

In particular, MHMMs enable us to formulate meaningful independence structures for the latent component and for the observable variables given the latent states. Such independence hypotheses are restrictions on the transition

* Corresponding author.
*E-mail addresses:* roberto.colombi@unibg.it (R. Colombi), sabrina.giordano@unical.it (S. Giordano).

http://dx.doi.org/10.1016/j.jmva.2015.04.002

probabilities of the latent chain and on the state-dependent distributions of the observation process. The same hypotheses can be investigated using the classical hidden Markov models, but formulating these hypotheses as simple conditional independencies and turning them into tractable constraints on parameters are much more demanding than in the MHMM context.

MHMMs can be well suited to applications where all the observable time series are affected by one common unobservable factor (general effect) and each observable variable is also governed by its specific latent variable. MHMMs may also handle time series data where an unobservable aspect influences the marginal dynamics of each observable variable while another latent factor influences the association among them. In the framework of MHMMs, different sets of observable categorical time series are allowed to depend on different sets of unobservable processes and the observable variables are not required to be independent given the latent states (local independence assumption), but the association between them is also modeled. Moreover, in MHMMs, the multivariate latent process can satisfy the hypotheses of Granger noncausality and contemporaneous independence described by Colombi and Giordano [7] for multivariate Markov chains.

So great is the variety of roles (specific, generic, association-affecting, etc.) assigned to the unobserved component in MHMMs, that it is hard to imagine this variety being easily handled by any classical HMM where only one latent process is allowed.

This approach, for example, responds to the shortcomings highlighted in Zucchini and Guttorp [23] who proposed a model for describing the sequence of wet and dry days at 5 sites without taking into account the spatial dependence among sites situated in closed locations and without allowing for a multivariate state process with sites in different regions responding to different components of the latent process. Other examples illustrated in the literature can be enriched by more flexible and realistic hypotheses using MHMMs.

To appreciate the possibility offered by MHMMs of modeling various hypotheses on the temporal evolution of the latent components and on the influence of the latent states on the observable variables, it is crucial to formulate such hypotheses by compact and simple expressions. The methodology of graphical models (Lauritzen [14]) serves this need. In fact, the graphs can visually represent the complex independence structures related to the latent and observation components of the MHMM and all the hypotheses can be described and handled by careful combination of simple elements in graphical models. This is the reason why we take advantage of the graphical models by presenting the independence conditions behind the MHMMs as Markov properties of the associated graphs and testing them as simple linear constraints on parameters.

The paper is organized as follows. MHMMs are presented in Section 2 and Section 3 illustrates that the transition probabilities of the latent model and the distributions of observable variables given the latent states of MHMMs can be required to obey the Markov properties of a mixed and a chain graph, respectively. In Section 4, the conditional independencies, defining the MHMMs and interpreted as Markov properties of graphs in Section 3, are shown to be equivalent to linear constraints on parameters of Gloneck and McCullagh models [12] for the transition probabilities and the state-dependent distributions. This last result is extremely important for fitting and testing MHMMs. Finally, in the last section several MHMMs are used to analyze two data sets.

## 2. Multiple hidden Markov models

Let $\mathbf{E}_{\mathcal{U}}$ be a $r$-variate process of categorical variables, $\mathbf{E}_{\mathcal{U}} = \{E_{\mathcal{U}}(t) : t \in \mathbb{N}\} = \{E_i(t) : t \in \mathbb{N}, i \in \mathcal{U}\}$, $\mathcal{U} = \{1, \ldots, r\}$, $\mathbb{N} = \{0, 1, 2, \ldots, \}$ and let $\mathbf{F}_{\mathcal{V}}$ be a $s$-dimensional process of categorical variables $\mathbf{F}_{\mathcal{V}} = \{F_{\mathcal{V}}(t) : t \in \mathbb{N}\} = \{F_j(t) : t \in \mathbb{N}, j \in \mathcal{V}\}$, $\mathcal{V} = \{1, \ldots, s\}$. The random variables $E_i(t)$, $F_j(t)$ take values in finite sets $\mathcal{E}_i$ and $\mathcal{F}_j$, $i \in \mathcal{U}, j \in \mathcal{V}$.

For every subset $\mathcal{T} \subset \mathcal{U}$ and $\mathcal{R} \subset \mathcal{V}$, marginal processes are represented by $\mathbf{E}_{\mathcal{T}} = \{E_i(t) : i \in \mathcal{T}, t \in \mathbb{N}\}$ and $\mathbf{F}_{\mathcal{R}} = \{F_j(t) : j \in \mathcal{R}, t \in \mathbb{N}\}$. Univariate marginal processes will be denoted by $\mathbf{E}_i$, $\mathbf{F}_j$, $i \in \mathcal{U}, j \in \mathcal{V}$.

The following definition states when $(\mathbf{E}_{\mathcal{U}}, \mathbf{F}_{\mathcal{V}})$ is an MHMM.

**Definition 1.** The joint process $(\mathbf{E}_{\mathcal{U}}, \mathbf{F}_{\mathcal{V}})$ is an MHMM if

(a) $\mathbf{E}_{\mathcal{U}}$ is unobservable
(b) $(\mathbf{E}_{\mathcal{U}}, \mathbf{F}_{\mathcal{V}})$ is a first order multivariate Markov chain
(c) $E_{\mathcal{U}}(t) \perp\!\!\!\perp F_{\mathcal{V}}(t-1) | E_{\mathcal{U}}(t-1)$
(d) $F_{\mathcal{V}}(t) \perp\!\!\!\perp E_{\mathcal{U}}(t-1), F_{\mathcal{V}}(t-1) | E_{\mathcal{U}}(t)$.

In particular, condition (c) implies that $\mathbf{E}_{\mathcal{U}}$ is a first order Markov chain (see [6]).

A marginal process $(\mathbf{E}_{\mathcal{T}}, \mathbf{F}_{\mathcal{R}})$, $\mathcal{T} \subset \mathcal{U}$ and $\mathcal{R} \subset \mathcal{V}$, in general, is not a hidden Markov model. The following theorem clarifies when the properties of an MHMM are preserved after marginalizing the latent and observation processes. This helps to specify under which restrictions all the attractive features of the hidden models (forecast distributions, smoothing and filtering algorithms, etc.) are still valid for an MHMM with marginalized components.

**Theorem 1.** *Let $\mathbf{E}_{\mathcal{T}}$ and $\mathbf{F}_{\mathcal{R}}$, $\mathcal{T} \subset \mathcal{U}$ and $\mathcal{R} \subset \mathcal{V}$, be marginal processes of the latent and observation components of an MHMM $(\mathbf{E}_{\mathcal{U}}, \mathbf{F}_{\mathcal{V}})$. The marginal process $(\mathbf{E}_{\mathcal{T}}, \mathbf{F}_{\mathcal{R}})$ is still an MHMM if the following conditions are fulfilled for all $t \in \mathbb{N} \setminus \{0\}$*

$$E_{\mathcal{T}}(t) \perp\!\!\!\perp E_{\mathcal{U} \setminus \mathcal{T}}(t-1) | E_{\mathcal{T}}(t-1) \tag{1}$$

$$F_{\mathcal{R}}(t) \perp\!\!\!\perp E_{\mathcal{U} \setminus \mathcal{T}}(t) | E_{\mathcal{T}}(t). \tag{2}$$