STUDY CENTER
GERZENSEE

Hidden Markov models in time series, with applications in economics

Sylvia Kaufmann

Working Paper 16.06

Hidden Markov models in time series, with applications in economics

Sylvia Kaufmann[*][†]

September 2016

Abstract

Markov models introduce persistence in the mixture distribution. In time series analysis, the mixture components relate to different persistent states characterizing the state-specific time series process. Model specification is discussed in a general form. Emphasis is put on the functional form and the parametrization of time-invariant and time-varying specifications of the state transition distribution. The concept of mean-square stability is introduced to discuss the condition under which Markov switching processes have finite first and second moments in the indefinite future. Not surprisingly, a time series process may be mean-square stable even if it switches between bounded and unbounded state-specific processes. Surprisingly, switching between stable state-specific processes is neither necessary nor sufficient to obtain a mean-square stable time series process. Model estimation proceeds by data augmentation. We derive the basic forward-filtering backward-smoothing/sampling algorithm to infer on the latent state indicator in maximum likelihood and Bayesian estimation procedures. Emphasis is again laid on the state transition distribution. We discuss the specification of state-invariant prior parameter distributions and posterior parameter inference under either a logit or probit functional form of the state transition distribution. With simulated data, we show that the estimation of parameters under a probit functional form is more efficient. However, a probit functional form renders estimation extremely slow if more than two states drive the time series process. Finally, various applications illustrate how to obtain informative switching in Markov switching models with time-invariant and time-varying transition distributions.

JEL classification: C11,C22,C24,C32,C34,E24,E32,E52
Key words: Bayesian inference, EM algorithm, Markov switching, prior information.

[*]Study Center Gerzensee, Foundation of the Swiss National Bank, Dorfstrasse 2, P.O. Box 21, CH-3115 Gerzensee, `sylvia.kaufmann@szgerzensee.ch`

[†]I thank Markus Pape for valuable comments and suggestions. All remaining errors and omissions are mine.

1 Introduction

Hidden Markov models are mixture models with sequential dependence or persistence in the mixture distribution. For a finite, discrete number of G components, persistence in distribution is induced by specifying a latent component indicator which follows a Markov process. The transition probabilities for the Markov process may either be time-invariant or time-varying. In the latter case, Markov models extend mixture of experts model (see chapter II.5 of this volume) by introducing persistence in the mixtures.

Hidden Markov models in time series econometrics became very popular after the publications of Hamilton (1989, 1990). He transferred earlier regression based approaches like Goldfeld and Quandt (1973) into time series analysis by recognizing their usefulness in capturing asymmetric conditional moments or asymmetric dynamic properties of time series. In section 2 we start by setting out the framework and terminology. In time series analysis, components are usually called states or regimes, and the transition between states is termed regime switch or regime change. This wording will be used in this chapter to be consistent with the econometrics literature. We discuss in separate sections the basic modelling choice of specifying the transition distribution of states. Hamilton (1989, 1990) introduced the model with time-invariant or constant transition distribution, and most of the following literature stayed with this specification.

This is not as restrictive as it may seem at first sight, given that more sophisticated models can be built by imposing either restrictions on the state transition probabilities or by combining multiple latent state indicators in a dynamical or hierarchical way. Change-point models (Chib 1996, Pesaran et al. 2007, Bauwens et al. 2015) are nested in Markov switching models by imposing appropriate zero restrictions on the transition distribution. Linking multiple latent state indicators dynamically, we can capture many leading/lagging features in multivariate analysis (Phillips 1991, Paap et al. 2009, Kaufmann 2010). Linking state indicators hierarchically, we obtain hierarchical Markov mixture models e.g. to disentangle long-term from short-term changing dynamics (Geweke and Amisano 2011, Bai and Wang 2011). Nevertheless, constant or exogenous transition distributions do not incorporate an explicit explanation or interpretation of the driving forces underlying the transition distribution.

Including covariates effects into the transition distribution renders it time-varying and yields at least an indication, if not a driving cause, of the regime switches. One of the first proposals is Diebold et al. (1994). Applications followed in business cycle analysis in Filardo (1994) and Filardo and Gordon (1998). Both probit and logit functional forms were used for the transition distribution. Under the assumption of independence between state alternatives, both parameterizations yield essentially the same estimation results. Later on, Koop and Potter (2007) introduced duration dependent time-varying probabilities into a change-point model. An interesting alternative is presented in Billio and Casarin (2011), who use a beta autoregressive process to model time-varying transition probabilities.

Against this background, we outline various extensions that are available within the general framework we present. Given that covariates may have state-dependent effects on the transition distribution, we elaborate on various considerations that may flow into

the specific parametrization of time-varying transition probabilities. Section 2 closes with the discussion of an attractive feature of Markov switching models that has so far, to our knowledge, not been exploited in time series analysis. So far, these models have been applied under the assumption that the conditional, i.e. state-dependent, distributions are stationary or, in other words, have finite moments in every period t. This need not be the case, however. Many real phenomena are consistent with a process that alternates between a stationary and a non-stationary state-specific distribution. Think of the recent financial crisis, during which dynamics across economic variables may have engaged transitorily on an unsustainable path. Francq and Zakoïan (2001), and more recently Farmer et al. (2009a), derive conditions under which the unconditional distribution of multivariate time series processes in the indefinite future has finite moments, even if some state- and period-specific conditional distributions may not have finite moments. Moreover, they show that state-specific stationary distributions are not sufficient for a multivariate process to approach finite moments in the indefinite future.

In section 3 we outline the estimation of Markov switching models, where the emphasis is on Bayesian estimation. Maximum likelihood estimation and variants of it are based on the EM algorithm, in which the 'E' step takes explicitly into account the state-dependence in the mixture to infer about the state indicator (Hamilton 1990). Extensions to multivariate models followed in Krolzig et al. (2002) and Clements and Krolzig (2003). The forward-filtering backward-sampling algorithm provides the basis for data augmentation in Bayesian estimation (McCulloch and Tsay 1994; Chib 1996). Markov chain Monte Carlo methods prove very useful to estimate models with multiple latent variables, like factor models with Markov switching factor mean or factor volatility (Kim and Nelson 1998).

Hidden Markov models endorse all issues concerning mixture modelling, as comprehensively exposed in Frühwirth-Schnatter (2006). In the present chapter, we therefore discuss in detail the design of state-invariant prior distributions for time-invariant and time-varying transition probabilities (Kaufmann 2015; Burgette and Hahn 2010). We then set out the posterior random permutation sampler to obtain draws from the unconstrained, multimodal posterior (Frühwirth-Schnatter 2001). To sample the parameters of the logit functional form, we borrow from data augmentation algorithms outlined in Frühwirth-Schnatter and Frühwirth (2010) which render the non-linear, non-Gaussian model in latent utilities linear Gaussian. Parameters are sampled from full conditional distributions rather than by Metropolis-Hastings (Scott 2011; Holmes and Held 2006). The approach of Burgette and Hahn (2010) proves very useful to sample parameters of the probit functional form. Instead of normalizing the error covariance of latent utilities with respect to a specific element (McCulloch et al. 2000; Imai and van Dyk 2005), they propose to restrict the trace of the normalized error covariance of the latent utilities, whereby normalization occurs in each iteration of the sampler with respect to a randomly chosen latent state utility. To conclude section 3, we compare estimation time and sampler efficiency between using the logit and the probit functional form to estimate the data generating process of a univariate series driven by 2 and 3 hidden Markov mixtures. We briefly illustrate that posterior state-identification is obtained by post-processing the