Inference Algorithms for Hidden (Semi) Markov Models

Patrick Aschermayr

Zurich, Switzerland

Abstract

The goal of this thesis is to explore and test new methods to learn, describe and predict economic cycles. In particular, a comparison of inference algorithms for hidden Markov models (HMM) and hidden semi-Markov models (HSMM) is conducted. Both proposed approachesthe popular Expectation-maximization (EM) algorithm and a Markov chain Monte Carlo (MCMC) samplerhave advantages depending on the size and noise of the underlying data as well as whether interval estimation or the addition of data-specific knowledge is desired. Furthermore, HMMs and HSMMs are used to describe financial markets. While the hidden Markov model performs well as a trading tool, it is less suitable to model economic cycles, since its implicit geometric state duration distribution flips states unrealistically often. In contrast, the hidden semi-Markov regime switching model appears to be very promising, demonstrating high potential as a strategic asset allocation (SAA) overlay from a finance perspective and as a model for economic cycle predictions from an economics perspective.

Keywords: Hidden Markov Models, Hidden Semi-Markov Models, Statistics, Machine Learning, Algorithms, EM-Algorithm, MCMC, Economic Cycles, Financial Markets, ETH Zurich, University of Zurich

Executive Summary

How do economies move? This question is of great relevance from a research perspective as economies and large institutions are increasingly connected to each other. This in turn stresses the need of understanding the world as a network of interconnected systems. For instance, the latest financial crisis has undoubtedly shown that knowledge about dependencies of global players and their impact on society is quite limited. Hence, it is essential to develop tools and methods to appropriately learn, describe and predict economic cycles.

A general class of methods that can model such dependencies and still extract valuable inference are probabilistic directed acyclic graphical models, such as Bayesian networks. Perhaps the most well-known example of a Bayesian network is the so called hidden Markov model (HMM). This thesis is primarily concerned with addressing of how to compute appropriate parameters for a 'good' model. The popular EM-Algorithm (Expectation Maximization Algorithm) for hidden Markov models is compared against Markov chain Monte Carlo (MCMC) methods. The large majority of literature still uses the EM-algorithm, but MCMC methods do have advantages, even when obtaining point estimates is the objective. Furthermore, the model is expanded to a hidden semi-Markov model (HSMM) by separately modelling the state duration of the transition matrix, and thus greatly improving modelling capabilities for real world applications.

In a first step, different scenarios are analyzed in order to compare the EM-algorithm and MCMC methods and their suitability for HMMs. When working with larger amounts of data, the EM-algorithm performs excellently and operates on similar performance levels as the MCMC sampler by running the algorithm several times, even though less amount of time is needed to fit the model. Thus, if only a point estimate is desired for the HMM, the EM-algorithm is the simplest and quickest solution. However, because it is typically run several times, the computing time advantage diminishes slightly. The EM-Algorithm also fails to converge on occasion. The MCMC method outperforms the EM-algorithm for smaller amount of data, but struggles to obtain accurate parameters for data with a lot of noise. Moreover, the computing time takes far longer for the latter approach. As financial markets are typically very noisy & only point estimates are required for forecasting, the EM-algorithm will ultimately be used as the preferred method when working with empirical data.

Finally, the models' ability to capture real data dynamics is tested in a financial markets setting. While the hidden Markov model performs well as a trading tool, it is less suitable to model economic cycles, since its implicit geometric state duration distribution flips states unrealistically often. In contrast, the hidden semi-Markov model appears to be very promising, demonstrating high potential as a strategic asset allocation (SAA) overlay from a finance perspective and as a model for economic cycle predictions from an economics perspective.