DiscreteTS : two hidden-Markov models for time series of count data

J. Alerini\textsuperscript{a}, M. Oltean\textsuperscript{b} and J. Ridgway\textsuperscript{b}

\textsuperscript{a} PIREH (Pôle Informatique de Recherche et d’Enseignement en Histoire)
Université Paris 1
1 Rue Victor Cousin, 75005 Paris, France
julien.alerini@univ-paris1.fr

\textsuperscript{b} SAMM (Statistique, Analyse et Modélisation Multidisciplinaire), EA 4543
Université Paris 1
90 Rue de Tolbiac, 75013 Paris, France
madalina.oltean@univ-paris1.fr, James.Ridgway@ensae-paristech.fr


Time series of count data are encountered often in Humanities and Social Sciences. Modeling this kind of data is a challenging topic for the statistician : autoregressive structure, over-dispersion in zero, existence of several unobserved regimes controlling the process.

One common approach used for modeling integer-valued time series are the hidden Markov models. However, the available R packages such as HiddenMarkov [1] or HMM [2] are implemented for usual distributions only. Moreover, none of this packages performs estimation for autoregressive Markov-switching models.

Two new models were recently introduced in [3] and [4] :

1. ZIP-HMM (Hidden Markov models with zero-inflated Poisson distributions) were proposed in order to take into account the over-dispersion in zero. This model is a usual hidden Markov model, except that the Poisson distribution of the observed process conditionally to the hidden state was replaced by a mixture of a Poisson and a Dirac distribution.

2. INAR\((p)\)-HMM (Integer-valued autoregressive models with Markov-switching regimes) were introduced as a parallel to the autoregressive hidden-Markov models existing already in the continuous case [5]. The observed process is supposed to behave as an integer-valued autoregressive INAR\((p)\) [6], whose parameters are controlled by the states of a hidden Markov chain.

For both models, the estimation procedure is achieved through the EM algorithm. These models were implemented in a R-package called DiscreteTS. The package provides the possibility of either simulating these models, or of estimating them starting from a given time-series. A toy example on medieval historical data is also provided.

References


