Chapter 8 - Extensions of the basic HMM

02433 - Hidden Markov Models

Martin Wæver Pedersen, Henrik Madsen

28 March, 2011

MWP, compiled March 31, 2011

Lunds universitet
Lunds Tekniska Högskola

Interreg IVA
ÖRESUND – KATTEGAT – SKAGERRAK

DTU

Extensions

- Second-order underlying Markov chain.
- Multinomial-like time series.
- Categorical time series.
- Other types of multivariate series.
- Series that depend on covariates.
- Models with additional dependencies.

General univariate time series

Unbounded counts
Possible models: Poisson, negative binomial. Examples: series of breakdowns of technical equipment, earthquakes, sales, insurance claims, accidents reported, defective items, stock trades, etc.

Binary data
Model: Bernoulli. Examples: series of rain/no rain, trading of a share (traded/not traded), airport departures (delayed/not delayed) etc.

Bounded counts
Model: Binomial. Difficult to forecast since we need to know the number of future trials. Examples: Purchasing preference, of a total of n_t sales on a day x_t contains the number of sales of a product, A, say.

Continuous valued data
Use probability densities instead of probability functions. Models: exponential, gamma, normal. Examples: series of share returns (normal), animal movements i.e. length of observed steps over time (exponential or gamma), waiting times (exponential or gamma).

HMM with second-order Markov chain

The dependency of a second-order Markov chain is depicted in Figure 8.1 on page 119 in Zucchini09. Technically a second-order Markov chain $\{C_t\}$ has the property

$$\Pr(C_t = k | C_{t-1}, C_{t-2}, C_{t-3}, \ldots) = \Pr(C_t = k | C_{t-1}, C_{t-2}),$$

with the transition probability

$$\Pr(C_t = k | C_{t-1} = j, C_{t-2} = i) = \gamma(i, j, k)$$

and stationary distribution

$$\Pr(C_t = k, C_{t-1} = j) = u(j, k).$$

A potential problem with higher-order Markov chains is that the size of the probability transition matrix increases rapidly with the number of states m. If a particular structure can be assumed for the transition matrix this problem may be mitigated. Still the size of the matrix for third-order Markov chains and above may become a limiting factor with respect to memory or computation time.

Multivariate time series - properties

With a multivariate time series the observations are represented by a vector $\mathbf{x}_t = \{x_{tj}\}$, where $t \in \{1, \dots, T\}$ and $j \in \{1, \dots, q\}$, i.e. q is the dimension of $\mathbf{x}_t$. The q components of $\mathbf{x}_t$ do not necessarily have the same type of distribution.

With multivariate time series we differentiate the concept of conditional independence into two sub-categories:

Longitudinal conditional independence

$$\Pr(\mathbf{X}_t = \mathbf{x}_t | \mathbf{C}^{(T)}) = \Pr(\mathbf{X}_t = \mathbf{x}_t | C_t)$$

Contemporaneous conditional independence

$$\Pr(\mathbf{X}_t = \mathbf{x}_t | C_t) = \Pr(X_{t1} = x_{t1} | C_t)\Pr(X_{t2} = x_{t2} | C_t) \cdots \Pr(X_{tq} = x_{tq} | C_t)$$

Longitudinal conditional independence comes from the fact that the underlying Markov chain has the (first-order) Markov property. Contemporaneous conditional independence is a convenient assumption when computing the state dependent distribution. It is however not always appropriate in which case the dependence between the q components must be accounted for.

Multivariate time series - examples

Multinomial distribution

Purchasing preference. Of a total of n_t sales on a day $\mathbf{x}_t$ is the vector containing the number of sales of each q products.

Categorical

A DNA base sequence has exactly one symbol at each position in the sequence. So at a position t $\mathbf{x}_t$ is a vector a vector containing $q-1$ zeros and one 1, e.g.

$$\mathbf{x}_t = (0, 0, 1, 0).$$

This is a special case of the multinomial distribution with $n_t = 1$.

Continuous

A weather gauge may at time t record $\mathbf{x}_t = (S_t, P_t, W_t)$, where S_t is temperature, P_t is pressure, and W_t is wind speed. These are all continuous measures, but may follow different distributions. It this case it is not appropriate to assume contemporaneous conditional independence of the observations, since e.g. pressure is not independent from wind speed at a given time.

Series that depend on covariates

Time series that depend on covariates can be modelled by HMMs in which the parameters are functions of the relevant covariates. That is the transition probabilities or the parameters related to the state dependent distributions may be functions of quantities (explanatory variables) that contain information about the data. An obvious example of a covariate is time which is useful for modelling seasonal variation in data.

Simple example: time as covariate for the intensity of a Poisson-HMM

$$\log\ {}_t\lambda_i = \boldsymbol{\beta}_i \mathbf{y}_t' = \beta_{i1} + \beta_{i2} t.$$

Seasonal changes could be incorporated by including sine terms depending on t in the above expression.

In estimating the model we would then not estimate ${}_t\lambda_i$, but instead the parameters in the functional expression β_{i1} and β_{i2}. So, including covariate will lead to more model parameters and probably longer computation times.

Models with additional dependencies

The simplest HMM has the dependencies

$$\Pr(X_t = x_t | C_t), \quad \text{and} \quad \Pr(C_t = i | C_{t-1}).$$

As illustrated in Figure 8.5 in Zucchini09 we may consider HMMs with more complicated dependencies. For example, models with additional dependencies may have the following properties

$$\Pr(X_t = x_t | C_t, X_{t-1}), \quad \text{and} \quad \Pr(C_t = i | C_{t-1}),$$

or

$$\Pr(X_t = x_t | C_t, X_{t-1}, X_{t-2}), \quad \text{and} \quad \Pr(C_t = i | C_{t-1}),$$

or

$$\Pr(X_t = x_t | C_t, C_{t-1}), \quad \text{and} \quad \Pr(C_t = i | C_{t-1}).$$

Such generalizations only lead to minimal modifications in the calculation of the state dependent distributions.

Exercises

3,5,9