Learning Imprecise Hidden Markov Models

Arthur Van Camp, Gert de Cooman, Jasper De Bock, Erik Quaeghebeur & Filip Hermans
SYSTeMS Research Group, Ghent University
{Arthur.VanCamp,Gert.deCooman,Jasper.DeBock,Erik.Quaeghebeur,Filip.Hermans}@UGent.be

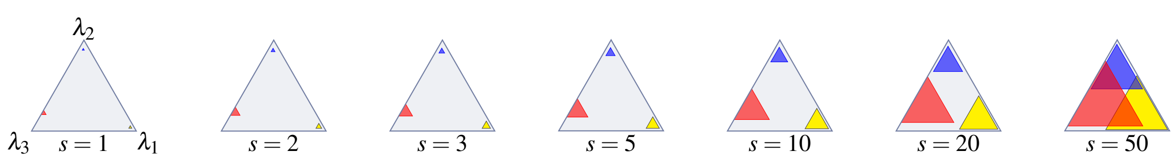
Abstract

Consider a stationary *precise* hidden Markov model (HMM) with n hidden states X_k, taking values x_k in a set $\{1,\ldots,m\}$ and n observations O_k, taking values o_k. Both the marginal model $p_{X_1}(x_1)$, the emission models $p_{O_k|X_k}(o_k|x_k)$ and the transition models $p_{X_k|X_{k-1}}(x_k|x_{k-1})$ are unknown. We can then use the Baum–Welch algorithm [see, e.g., 4] to get a maximum-likelihood estimate of these models. The Baum–Welch algorithm constructs the expected number of transitions $n_{ij} := \sum_{k=2}^{n} p_{X_{k-1},X_k|O_1,\ldots,O_n}(i,j|o_1,\ldots,o_n)$ from state i to state j in the whole Markov chain.

If we do not have enough data to justify a precise probability model, such as the one we obtain using the classical Baum–Welch algorithm, then the approach we present is useful. Our contribution exists of a method for learning *imprecise* transition probabilities in an HMM. We are not aware of another such method in the literature. These transitions from a state $X_{k-1} = i$ to a state $X_k = j$ are multinomial processes. The imprecise Dirichlet model (IDM) is a convenient model for describing uncertainty about such processes [3]. In order to learn using an IDM, we need the number of transitions and a choice for the pseudocounts s. Since the hidden states are unavailable, our method consists in taking the expected[1] number of transitions (positive real numbers instead of natural numbers), derived from the Baum–Welch algorithm, rather than real counts. So, the lower probability for state j conditional on state i is estimated by $\underline{Q}(\{j\}|i) := n_{ij}/(s+\sum_{j=1}^{m} n_{ij})$.

Learning an imprecise HMM, like our method does, is necessary before being able to make imprecise inferences. There are algorithms, like MePICTIr [2] and EstiHMM [1], to make exact inferences in an imprecise HMM.

We applied our method to the following application. Given an observation sequence of counted earthquakes in a number of consecutive years, we are interested in predicting the number of earthquakes in the next years. The observed number of earthquakes in a year k is assumed to be generated by a Poisson process with rate λ_{X_k}. Both the emission probabilities and the marginal probability are kept precise, while the transition probabilities are made imprecise, using our method. The following figure shows the credal sets for the transition model in the case $m = 3$, for different pseudocounts s.



Keywords. hidden Markov model, learning, expected counts, imprecise Dirichlet model

References

[1] J. De Bock and G. de Cooman. State sequence prediction in imprecise hidden markov models. In *Proc. of ISIPTA '11*. To appear.
[2] G. de Cooman, F. Hermans, A. Antonucci, and M. Zaffalon. Epistemic irrelevance in credal nets: The case of imprecise markov trees. *Int. J. Approx. Reason.*, 51:1029–1052, 2010. doi:10.1016/j.ijar.2010.08.011. URL http://hdl.handle.net/1854/LU-1086418.
[3] P. Walley. Inferences from multinomial data: Learning about a bag of marbles. *J. R. Stat. Soc. Ser. B*, 58:3–57, 1996. URL http://www.jstor.org/stable/2346164.
[4] W. Zucchini and I. MacDonald. *Hidden Markov Models for Time Series. An Introduction Using R*. Chapman and Hall, 2009.

[1] The idea of using expected counts in an IDM has also been suggested by and discussed with Marco Zaffalon.