Summary 12: An MCMC Sampling Approach to Estimation of Non-stationary Hidden Markov Models

Susan Cherry

This paper extends the MCMC-based methods to estimation of Non-Stationary MCMC (NSMCMC). Hidden Markov Models (HMM) are important in many areas of statistical analysis, such as signal processing and speech processing. They are also used in the Human GenomeProject to extract information from large amounts of data and for the automatic classification of protein sequences. Despite their usefulness, HMM have a weakness: they are inflexible to model state durations. Their state durations are fixed geometric distributions. To overcome this weakness, variable duration HMMs have been introduced. While VDHMMs are more complex than traditional HMMs, they are also more flexible. Later, a non-stationary HMM NSHMM) was introduced, in which the state transition probabilities are modeled as functions of time. While VDHMMs and NSHMMs are equivalent, NSHMMs are usually more tractable for use.

The paper begins by giving a review of traditional HMMs and the main problems associated with them. In particular, let d be the duration of a particular state. the probability of d is $S_k(d) = a_{kk}^{d-1}(1 - a_{kk})$, which is geometric. Next, it moves on to non-stationary HMMs. State duration probability mass functions are introduced. The method for generating state sequences of VDHMMs is specified. NSHMMs are created by allowing all of the transition probabilities $a_{i,j}$ to be functions of d, so $a_{i,j}(d)$ is the probability that the system will switch from S_i to S_j. Proposition 1 states the relationship between the probability mass functions $P_i(d)$ and the self-transition probabilities $a_{ii}(d)$. Proposition 2 states that the NSHMM with constant state transition weights is equivalent to VDHMM. Though they are equivalent, NSHHMs are more convenient to use because they are more tractable for analysis.

Next, the paper discusses the estimation of NSHMMs by MCMC Sampling. MCMC sampling allows us to work with high dimensional and complex models. The paper specified the priors and goes through the details of the Gibbs Sampling Procedure. It also discusses the problem of label switching. One of the main issues with MCMC is the assessment of convergence. Here, the authors adopt a method based on multiple chance and use the between-sequence and within sequence variances.

Finally the paper discusses simulation results. The authors perform two experiments where the number of states was $N = 3$ and the number of emission variables was L=7. The specific the emission parameters in both models and use a MAP estimator. In the first experiment, there were only two mismatches, as can be seen in Figure 5. In the second experiment there were 23 mismatches. The results obtained using Expectation Maximization are show in Figure 6. Overall, the authors found: 1) The results from EM+Viterbi algorithm are not as accurate as the MCMC method and 2) The EM+Viterbi algorithm is sensitive to initializations. The posterior probabilities are show in Figure 7. The second experiment needed more than 2000 iterations to converge. The convergence depends on the parameters of the NSHMM. In summary, this paper shows that the NCHMM is easy to implement because it is straightforward to draw samples from the conditional distributions. The experiments show quick converge and good accuracy of the model parameters.