ESTIMATION IN MARKOV PROCESSES

Narayana Prasad Santhanam

Department of Electrical Engineering,
University of Hawaii, Honolulu HI 96822
nsanthan@hawaii.edu

ABSTRACT

We observe a length-n sample generated by an unknown, stationary ergodic Markov process (”model”) over a finite alphabet A. Motivated by applications in what are known as backplane channels, given any string w of symbols from A we want estimates of the conditional probability distribution of symbols following w (”model parameters”), as well as the stationary probability of w.

Two distinct problems that complicate estimation in this setting are (i) long memory, and (ii) slow mixing, which could happen even with only one bit of memory. For our problem, any consistent estimator can only converge pointwise over the class of all ergodic Markov models. Namely, given any estimator and any sample size n, the underlying model could be such that the estimator performs poorly on a sample of size n with high probability. But can we look at a length-n sample and identify if an estimate is likely to be accurate?

Since the memory is unknown a-priori, a natural approach is to estimate a potentially coarser model with memory k_n=O(log n). As n grows, estimates get refined and this approach is consistent, with the above scaling of kn also k_nown to be essentially optimal. But while effective asymptotically, the situation is quite different when we want the best answers possible with a length-n sample, rather than just consistency. Combining results in universal compression with Aldous’ coupling arguments, we obtain sufficient conditions on the length-n sample (even for slow mixing models) to identify when naive (i) estimates of the model parameters and (ii) estimates related to the stationary probabilities are accurate; and also bound the deviations of the naive estimates from true values.