Variable Length Markov Chains and Dynamic Combination of Models

A dissertation submitted to the
Swiss Federal Institute of Technology
Zurich

for the degree of
Doctor of Mathematics

presented by
Fiorenzo Ferrari
Dipl. Math. ETH
born July 29, 1972
citizen of Lodrino TI

accepted on the recommendation of
Prof. Dr. P. Bühlmann, examiner
Prof. Dr. H. R. Künsch, co-examiner

2002
Abstract

We develop new results about a sieve methodology for estimation of minimal state spaces and probability laws in the class of general stationary processes. Using a sieve approximation with variable length Markov chains of increasing order, we first show that for finite categorical spaces, an adapted version of the Context Algorithm yields asymptotically correct estimates for the minimal state space and for the probability distribution.

Variable length Markov chains are essentially Markov chains of higher order, but with the property of having a reduced memory, where some irrelevant states are lumped together. Thus in many situations, the number of parameters to be estimated turns out to be drastically lower compared with full Markov chains. The proposed method of sieves thereby yields a nice graphical tree representation for the potentially infinite dimensional minimal state space of the data generating process, which is very useful for exploration of the memory.

This procedure is also consistent for increasing size countable categorical spaces. Using a quantization procedure based on the inverse distribution function of real-valued variables, we construct an estimator of the distribution of real-valued general stationary processes. Also for this estimator we prove some consistency.

This result, with the associated quantization method, is the basis for the introduction of a new class of models for time series analysis: the Dynamic Combination of Models. The new methodology has a wide range of applicability: essentially for the conditional moments or distribution of a real-valued random variable given its past information (for instance the conditional expectation).

These moments or distributions are represented as a dynamically
weighted sum of some local models (for instance autoregressive models). The weights are the transition probabilities from an approximating variable length Markov chain for the quantization intervals of the underlying process.

Local model estimation is pursued with the maximum likelihood method, or a two step backfitting algorithm, depending whether the local model is finite parametric or infinite dimensional and nonparametric. We show that the maximum likelihood estimator of the dynamic combination of autoregressive models is consistent for the theoretically best dynamic combination of autoregressive models. Model selection is performed with the Akaike information criterion. Either with a "full search" (exhaustive) but simple quantization structure or with a "greedy search" (forward) involving more complex quantizers.

A wide empirical study supports this new methodology. In some situations the performance is evidently better than with more classical approaches.