Distributions of pattern statistics in sparse Markov models

Donald E. K. Martin[1]

Abstract

Markov models provide a good approximation to probabilities associated with many categorical time series, and thus they are applied extensively. However, a major drawback associated with them is that the number of model parameters grows exponentially in the order of the model, and thus only very low-order models are considered in applications. Another drawback is lack of flexibility, in that Markov models give relatively few choices for the number of model parameters. Sparse Markov models are Markov models with conditioning histories that are grouped into classes such that the conditional probability distribution for members of each class is constant. The model gives a better handling of the trade-off between bias associated with having too few model parameters and variance from having too many. In this paper, methodology for efficient computation of pattern distributions through Markov chains with minimal state spaces is extended to the sparse Markov framework.

Keywords Auxiliary Markov chain · Pattern distribution · Sparse Markov model · Variable length Markov chain

✉ Donald E. K. Martin
    demarti4@ncsu.edu

[1] Department of Statistics, North Carolina State University, 4272 SAS Hall, 2311 Stinson Drive, Raleigh, NC 27695-8203, USA