Emily B. Fox, Erik B. Sudderth, Michael I. Jordan, and Alan S. Willsky

Bayesian Nonparametric Learning of Markov Switching Processes

Stochastic process priors for dynamical systems

Markov switching processes, such as the hidden Markov model (HMM) and switching linear dynamical system (SLDS), are often used to describe rich dynamical phenomena. They describe complex behavior via repeated returns to a set of simpler models: imagine a person alternating between *walking, running*, and *jumping* behaviors, or a stock index switching between regimes of high and low volatility. Classical approaches to identification and estimation of these models assume a fixed, prespecified number of dynamical models. We instead examine Bayesian nonparametric approaches that define a prior on Markov switching processes with an unbounded number of potential model parameters (i.e., Markov modes). By leveraging stochastic processes such as the beta and Dirichlet process (DP), these methods allow the data to drive the complexity of the learned model, while still permitting efficient inference algorithms. They also lead to generalizations that discover and model dynamical behaviors shared among multiple related time series.

© PHOTODISC

INTRODUCTION

A core problem in statistical signal processing is the partitioning of temporal data into segments, each of which permits a relatively simple statistical description. This segmentation problem arises in a variety of applications, in areas as diverse as speech recognition, computational biology, natural language processing, finance, and cryptanalysis. While in some cases the problem is merely that of detecting temporal changes (in which case the problem can be viewed as one of changepoint detection), in many other cases the temporal segments have a natural meaning in the domain and the problem is to recognize recurrence of a meaningful entity (e.g., a particular speaker, a gene, a part of speech, or a market condition). This leads naturally to state-space models, where the entities that are to be detected are encoded in the state.

The classical example of a state-space model for segmentation is the HMM [1]. The HMM is based on a discrete state variable and on the probabilistic assertions that the state transitions are Markovian and that the observations are conditionally independent given the state. In this model, temporal segments are equated with states, a natural assumption in some problems but a limiting assumption in many others. Consider, for example, the dance of a honey bee as it switches between a set of turn

Digital Object Identifier 10.1109/MSP.2010.937999

1053-5888/10/$26.00©2010IEEE