The Annals of Applied Statistics
2014, Vol. 8, No. 3, 1281–1313
DOI: 10.1214/14-AOAS742

JOINT MODELING OF MULTIPLE TIME SERIES VIA THE BETA PROCESS WITH APPLICATION TO MOTION CAPTURE SEGMENTATION

BY EMILY B. FOX[1,*], MICHAEL C. HUGHES[1,2,†], ERIK B. SUDDERTH[1,†] AND MICHAEL I. JORDAN[1,‡]

University of Washington[*]*, Brown University*[†] *and University of California, Berkeley*[‡]

We propose a Bayesian nonparametric approach to the problem of jointly modeling multiple related time series. Our model discovers a latent set of dynamical behaviors shared among the sequences, and segments each time series into regions defined by a subset of these behaviors. Using a beta process prior, the size of the behavior set and the sharing pattern are both inferred from data. We develop Markov chain Monte Carlo (MCMC) methods based on the Indian buffet process representation of the predictive distribution of the beta process. Our MCMC inference algorithm efficiently adds and removes behaviors via novel split-merge moves as well as data-driven birth and death proposals, avoiding the need to consider a truncated model. We demonstrate promising results on unsupervised segmentation of human motion capture data.

1. Introduction. Classical time series analysis has generally focused on the study of a single (potentially multivariate) time series. Instead, we consider analyzing *collections* of related time series, motivated by the increasing abundance of such data in many domains. In this work we explore this problem by considering time series produced by motion capture sensors on the joints of people performing exercise routines. An individual recording provides a multivariate time series that can be segmented into types of exercises (e.g., jumping jacks, arm-circles, and twists). Each exercise type describes locally coherent and simple dynamics that persist over a segment of time. We have such motion capture recordings from *multiple* individuals, each of whom performs some subset of a global set of exercises, as shown in Figure 1. Our goal is to discover the set of global exercise types ("behaviors") and their occurrences in each individual's data stream. We would like to take advantage of the overlap between individuals: if a jumping-jack behavior is discovered in one sequence, then it can be used to model data for other individuals.

Received May 2013; revised January 2014.

[1]Supported in part by AFOSR Grant FA9550-12-1-0453 and ONR Contracts/Grants N00014-11-1-0688 and N00014-10-1-0746.

[2]Supported in part by an NSF Graduate Research Fellowship under Grant DGE0228243.

Key words and phrases. Bayesian nonparametrics, beta process, hidden Markov models, motion capture, multiple time series.