Smoothing Algorithms for Variable Rate Models

Pete Bunch, Simon Godsill, *Member, IEEE,*

***Abstract*—The abstract goes here.**

***Index Terms*—**

I. Introduction

THE objective of sequential Bayesian inference is to estimate an imperfectly observed quantity as it varies over time. This is accomplished through the use of probabilistic models for the state evolution and measurement processes. Often, the latent state is a continuously varying quantity, whereas the observations are made at a discrete set of times. In these circumstances, it is simplest to discretise the state onto the same set of times as the observations. When the system is also Markovian, this leads to the standard "fixed rate" hidden Markov model (HMM). The standard HMM is poorly suited to systems where the state evolution contains discontinuities; for example, the price of a financial asset which may display large jumps at random times between periods of diffusion-like behaviour, or the kinematic state of a manoeuvring vehicle which may have sudden changes in the acceleration when turns begin or end. Such problems can be handled more naturally using a "variable rate" model, in which the state dynamics are conditioned upon a set of random changepoints which characterise transitions in behaviour.

In a variable rate model, the set of changepoints and associated parameters are modelled as a marked point process (MPP), the mathematical properties of which are thoroughly set out in [**?**]. Conditional upon this MPP the state evolves according to some benign dynamics. In [**?**], [**?**], the conditional state evolution is treated as deterministic, while in [**?**], [**?**] a conditionally linear-Gaussian state model is considered.

The posterior distribution for the changepoint MPP is inherently nonlinear, and cannot be calculated analytically. Instead, inference must be conducted using numerical approximations. The particle filter (introduced by [**?**]) and smoother (see [**?**], [**?**]) are schemes which approximate a posterior distribution using a set of samples, or "particles", drawn sequentially from it. A thorough introduction to particle filtering and smoothing methods can be found in [**?**], [**?**]. In [**?**], [**?**], [**?**], the particle filter was adapted for use with variable rate models, resulting in the variable rate particle filter (VRPF).

The VRPF allows the changepoint sequence – and hence the current state – to be estimated sequentially as observations are received. However, estimates can often be improved later once further observations have been made. In this paper, we address the problem of smoothing in variable models, i.e. the estimation of the changepoint sequence and latent state given all the observations. This is achieved with an efficient backward sweep through the observations, in a similar manner to the method for standard HMMs described in [**?**]. Two new schemes are introduced: one for conditionally linear-Gaussian models which exploits the method of Rao-Blackwellisation; and a second for use with conditionally-deterministic models which uses an augmented target distribution in the style of an SMC sampler [**?**].

Add a bit about structure of paper.

P. Bunch and S. Godsill are with the Department of Engineering, Cambridge University, UK. email: {pb404,sjg30}@cam.ac.uk

Manuscript received January 01, 1901; revised January 02, 1901.

II. Variable Rate Models

We consider a general model from time 0 to T, between which observations, $\{y_1 \dots y_N\}$, are made at times $\{t_1 \dots t_N = T\}$. During this period, an unknown number of changepoints, K, occur at times $\{\tau_0 = 0, \tau_1 \dots \tau_K\}$, each with associated changepoint parameters, $\{u_0, u_1 \dots u_K\}$. The pairs $\{\tau_k, u_k\}$ are the elements of an marked point process (MPP). We will refer to the times of the MPP as the "change-time sequence" and marks as "parameter sequence", and both together as the "changepoint sequence". The latent state is a continuous-time process denoted $x(t)$. Discrete sets containing multiple values over time will be written as, e.g. $y_{1:n} = \{y_1 \dots y_n\}$.

The objective for inference will be to estimate the changepoint sequence. This will be denoted as $\theta = \{\tau_{0:K}, u_{0:K}\}$. At a particular time t_n, the sequence will be divided into past $\theta_n = \{\tau_j, u_j \forall j : 0 \leq \tau_j < t_n\}$, and future $\theta_n^+ = \{\tau_j, u_j \forall j : t_n \leq \tau_j < T\}$. It will also be useful to define a variable for the changepoints which occur in the interval $[t_{n-1}, t_n)$, $\theta_{n \setminus n-1} = \{\tau_j, u_j \forall j : t_{n-1} \leq \tau_j < t_n\}$.

For notational simplicity, a counting variable is used to keep track of the most recent changepoint to have occurred,

$$\nu_t = \max(k : \tau_k < t). \tag{1}$$

The changepoint sequence is assumed to be a Markov process.

$$\{\tau_k, u_k\} \sim p(u_k|\tau_k, \tau_{k-1}, u_{k-1}) p(\tau_k|\tau_{k-1}, u_{k-1}) \tag{2}$$

The changepoint density will be constructed such that $P(\tau_k < \tau_{k-1}) = 0$.

In the manner of [**?**], a survivor function is defined as the probability that no new changepoint occurs before a given time,

$$\begin{aligned} S(\tau_k, u_k, t) &= P(\tau_{k+1} > t|\tau_k, u_k) \\ &= 1 - \int_{\tau_k}^{t} p(\xi|\tau_k, u_k) d\xi. \end{aligned} \tag{3}$$

It is now possible to write down a prior for the changepoint sequence, where we use the convention that $\tau_0 = 0$. The existence of such a density for a MPP is addressed in [**?**].

$$p(\theta_n) = S(\tau_{\nu_{t_n}}, t_n) p(u_0) \prod_{k=1}^{\nu_{t_n}} p(\tau_k, u_k | \tau_{k-1}, u_{k-1}) \quad (4)$$

A. Conditionally Linear-Gaussian Models

A useful class of variable rate models is those whose state dynamics are linear-Gaussian conditional on the changepoint sequence. Such a model may be discretised onto the set of observation times in exactly the same manner as a standard the standard HMM.

$$x_n = A_n(\theta_n) x_{n-1} + w_n \quad (5)$$

$$y_n = C_n(\theta_n) x_n + v_n. \quad (6)$$

The random variables w_n and v_n have a zero-mean Gaussian distribution with covariance matrices $Q_n(\theta_n)$ and $R_n(\theta_n)$ respectively.

Conditionally linear-Gaussian variable rate models were introduced in [**?**] for a financial inference algorithm. Change-points correspond to jumps in the value or trend of a security, at which points the covariance, $Q_n(\theta_n)$ is inflated.

B. Conditionally Deterministic Models