OPTIMAL ESTIMATION AND CRAMÉR-RAO BOUNDS FOR PARTIAL NON-GAUSSIAN STATE SPACE MODELS *

NICLAS BERGMAN1**, ARNAUD DOUCET2 AND NEIL GORDON3***

1Division of Automatic Control, Linköping University, S-581 83 Linköping, Sweden
2Signal Processing Group, Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, U.K., e-mail: ad2@eng.cam.ac.uk
3Defence Evaluation and Research Agency, St. Andrews Road, Malvern, Worcestershire WR14 3PS, U.K., e-mail: N.Gordon@signal.dera.gov.uk

(Received May 1, 2000; revised August 4, 2000)

Abstract. Partial non-Gaussian state-space models include many models of interest while keeping a convenient analytical structure. In this paper, two problems related to partial non-Gaussian models are addressed. First, we present an efficient sequential Monte Carlo method to perform Bayesian inference. Second, we derive simple recursions to compute posterior Cramér-Rao bounds (PCRB). An application to jump Markov linear systems (JMLS) is given.

Key words and phrases: Optimal estimation, Bayesian inference, sequential Monte Carlo methods, posterior Cramér-Rao bounds.

1. Introduction

1.1 Background

A partial non-Gaussian state-space model is a linear model whose parameters evolve with time according to an unobserved stochastic process $s_t$. Let $t$ denote the discrete-time index, then one has

(1.1) \[ x_t = A_t(s_t)x_{t-1} + B_t(s_t)u_t + F_t(s_t)w_t \]
(1.2) \[ y_t = C_t(s_t)x_t + D_t(s_t)w_t + G_t(s_t)u_t, \]

where $x_t \in \mathbb{R}^{n_x}$, $y_t \in \mathbb{R}^{n_y}$, $u_t \in \mathbb{R}^{n_u}$, $v_t \in \mathbb{R}^{n_v}$ and $w_t \in \mathbb{R}^{n_w}$. Given $s_t$, $A_t(s_t)$, $B_t(s_t)$, $C_t(s_t)$, $D(s_t)$, $F_t(s_t)$ and $G_t(s_t)$ are known matrices of appropriate dimension and $D_t(s_t)D_t^T(s_t) > 0$ for any $s_t$. $x_t$ is an unobserved state, $y_t$ is the observation process and $u_t$ is an exogenous control term. The noise sequences $v_t \overset{\text{i.i.d.}}{\sim} N(0, I_{n_v})$, $w_t \overset{\text{i.i.d.}}{\sim} N(0, I_{n_w})$ are independent Gaussian sequences, mutually independent and independent of the initial state $x_0 \sim N(m_0, P_0)$.

Conditional upon $s_t$, (1.1)–(1.2) is thus a standard linear Gaussian state-space model. However, the process $s_t$ is itself an unobserved random process. For the sake of simplicity, it is assumed to be a first-order Markov process of initial distribution $p(s_0)$ and Markov transition kernel $p(s_t \mid s_{t-1})$.

**Now at SaabTech Systems, Nettovägen 6, 175 88 Järfalla, Sweden, e-mail: ncbe@celsiustech.se
***The research of Neil Gordon was sponsored by UK MOD Corporate Research programme TG9.
This class of models has numerous applications as illustrated in the two following examples, see Kitagawa and Gersch (1996), Shephard (1994) and West and Harrison (1997) for many other examples.

Example 1. Jump Markov linear system. Assume that the process $s_t$ is a finite state-space Markov chain, then the resulting model is a so-called JMLS; that is a linear Gaussian system whose parameters evolve according to an unobserved finite state-space Markov chain. Such systems are widely used in digital communications, econometrics and target tracking (see Bar-Shalom and Li (1995)).

Example 2. Time-varying autoregressive (TWAR) process. The TWAR coefficients $a_t$ are reparametrised into the partial correlation coefficients $s_t \in \mathbb{R}^{n_x}$ and $s_t$ is assumed to follow a simple Gaussian random walk: $s_t = s_{t-1} + \varepsilon_t$ where $s_0 \sim N(0, I_{n_x})$ and $\varepsilon_t \sim N(0, I_{n_x})$. From $s_t$, one can compute $a_t$ through the standard Levinson recursion. Let $u_t = \sum_{i=1}^{n_x} a_{t,i} u_{t-i} + \sigma_w v_t$ and $y_t = u_t + \sigma_w w_t$ where $v_t \sim N(0,1)$ and $w_t \sim N(0,1)$. Then, by setting $x_t \equiv (u_t, \ldots, u_{t-n_x})$, one can put $(s_t, x_t, y_t)$ in the state-space form (1.1)–(1.2).

In this paper, we propose an efficient sequential Monte Carlo (SMC) method to perform Bayesian inference for partial non-Gaussian state-spaces and we derive simple recursions allowing easy computation of some PCRB.

1.2 Sequential Bayesian estimation

We denote for any sequence $z_t$, $z_{i:j} \equiv (z_i, z_{i+1}, \ldots, z_j)$. We are interested in estimating sequentially in time $t$ the posterior distribution of the state of the system given by $p(x_{0:t}, s_{0:t} | y_{1:t})$, or some of its characteristics such as the filtering distribution $p(x_t | y_{1:t})$. There is no closed-form expression for this class of models and one needs to use computational methods to perform Bayesian inference. In a batch framework, several authors have exploited the structure of partial non-Gaussian state-space models so as to develop efficient Markov chain Monte Carlo (MCMC) algorithms (see Carter and Kohn (1994, 1996), Frühwirth-Schnatter (1994) and Shephard (1994)). However MCMC methods are not suited to sequential estimation. Recently there has been a surge of interest in SMC methods for nonlinear/non-Gaussian time series analysis (Doucet et al. (2001)). These methods, initiated in Gordon et al. (1993) and Kitagawa (1996), utilise a random sample (or particle) based representation of the posterior probability distributions: the particles are propagated over time using a combination of sequential importance sampling and resampling steps. Related early work by West (1993a, 1993b) develops weighted mixtures of kernel densities as the proposal distribution for sequential importance sampling. However, in their standard forms, these algorithms do not use all the salient structure of partial non-Gaussian state-space models.

We show here how it is possible to use this structure to develop an efficient SMC algorithm to perform sequential Bayesian estimation. This algorithm combines sequential importance sampling, a selection scheme and MCMC methods. In particular, variance reduction is achieved by Rao-Blackwellisation using the Kalman filter as discussed in Doucet (1997) and Doucet et al. (2000). However, we further improve the algorithm by using other variance reduction methods and sampling schemes. A generalization of the backward-forward algorithm of Carter and Kohn (1996) is also given: it allows an exact initialization of the backward recursion, and requires neither the state covariance matrix to be strictly positive nor $A(s_t)$ to be regular.