Probabilistic Engineering Mechanics 38 (2014) 77–87

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

Iterated stochastic filters with additive updates for dynamic system identification: Annealing-type iterations and the filter bank

Tara Raveendran [a], Debasish Roy [b,*], Ram Mohan Vasu [a]

[a] Department of Instrumentation and Applied Physics, Indian Institute of Science, Bangalore, India
[b] Computational Mechanics Lab, Department of Civil Engineering, Indian Institute of Science, Bangalore, India

ARTICLE INFO

Article history:
Received 13 November 2013
Received in revised form
23 August 2014
Accepted 17 September 2014
Available online 20 September 2014

Keywords:
Stochastic filtering
Iterated additive update
Ensemble square root filter
Gaussian sum approximation
Kushner–Stratonovich equation
Dynamic system identification

ABSTRACT

A nonlinear stochastic filtering scheme based on a Gaussian sum representation of the filtering density and an annealing-type iterative update, which is additive and uses an artificial diffusion parameter, is proposed. The additive nature of the update relieves the problem of weight collapse often encountered with filters employing weighted particle based empirical approximation to the filtering density. The proposed Monte Carlo filter bank conforms in structure to the parent nonlinear filtering (Kushner–Stratonovich) equation and possesses excellent mixing properties enabling adequate exploration of the phase space of the state vector. The performance of the filter bank, presently assessed against a few carefully chosen numerical examples, provide ample evidence of its remarkable performance in terms of filter convergence and estimation accuracy vis-à-vis most other competing filters especially in higher dimensional dynamic system identification problems including cases that may demand estimating relatively minor variations in the parameter values from their reference states.

1. Introduction

Dynamic system identification aims at estimating the hidden state processes that solve the system or process model, often in the form of stochastic ordinary differential equations (SDEs), given a set of noisy partial observations, which are typically characterized by the observation SDEs whose drift fields are known functions of the system (process) states. The 'estimate of a state' often stands for its mean (first moment) with respect to the filtering probability density function (PDF) of the instantaneous state conditioned on the observation history till the current time. Variants of Bayesian filtering, which provide a computationally feasible route in obtaining the filtering PDF, typically involve a two-stage recursive procedure consisting of the prediction and update stages. While the prediction stage recursively propagates the process or system model in time, the predicted solution is modified in the update stage in order to assimilate the currently available observation consistent with a recursive form of the generalized Bayes' formula [1] and thus characterize (marginals of) the filtering PDF (also called the posterior PDF). The Kalman filter (KF) has been a major breakthrough [2], providing for an analytical scheme to arrive at the exact posterior PDF for a linear Gaussian dynamic state space model. Nonlinear dynamical systems with non-Gaussian additive/multiplicative noises may also be dealt with, albeit sub-optimally, with the extended Kalman filter (EKF) that employs linearized approximations to the signal-observation dynamics. But the EKF and its variants [3] may perform quite poorly where the dynamics are significantly nonlinear due to the imprecise Gaussian approximation of the transition law of the signal-observation process. Moreover, unless an extensive tuning operation for the process noise covariance is performed, the evolution of the analytical error covariance in the KF/EKF may become divergent.

With the rapid emergence of cheaply available computing resources, sequential Monte Carlo (SMC) methods such as particle filters (PFs), which provide asymptotically optimal estimates for nonlinear and non-Gaussian filtering problems, are being increasingly used. PFs rely on a first order Markov model for the time-discretized signal-observation processes and implement a recursive Bayesian update by Monte Carlo (MC) simulations [4]. Over a given time-step, they use particles, which are independently sampled and weighted realizations of the random variables (representing the instantaneous filtered states) to approximate the continuous filtering PDF by random (empirical) measures. Here the weights define the likelihood of the current observation given the predicted particles available through time-integration of the process dynamics. Being free from the approximations involving linearizations, PFs are endowed with the universality that have seen their applications in the context of a wide-ranging array of noisy nonlinear dynamical systems encountered in target

* Corresponding author.
E-mail address: royd@civil.iisc.ernet.in (D. Roy).

http://dx.doi.org/10.1016/j.probengmech.2014.09.002
0266-8920/

tracking, digital communications, chemical engineering etc. [5–7]. Efforts to use a form of analyticity characteristic of the KF within the framework of PFs have also led to the development of semi-analytical PFs [8]. Such PFs transform the nonlinear system/observations to an ensemble of piecewise linearized equations so that the KF can be used for each linearized system to yield a family of conditionally Gaussian posterior PDFs whose weighted sum yield the filtering PDF. The accruing advantage of reduced sampling variance however comes at the cost of a substantively increased computational overhead as the current observation must be repetitively assimilated for each linearized system.

Despite its universality and algorithmic simplicity, a PF is beset with the generic problem of particle impoverishment in applications involving higher dimensional process models as the weights tend to collapse to a point mass [9]. Indeed, the necessary sample size needed to counter such weight degeneracy could be practically unattainable even with the best of computing resources. A way out of this degeneracy, which is also the primary focus of this article, could be provided through additive updates that may be contrasted with the multiplicative, weight-based updates used with the PFs. One such prominent example, the ensemble Kalman filter (EnKF) that may be loosely viewed as an MC version of the KF implementing additive gain-type updates, has indeed found applications in higher dimensional filtering problems in oceanographic and atmospheric modeling [10]. The EnKF uses an ensemble of system states predicted through the process dynamics, thus avoiding the EKF-type Gaussian closure through linearization in the prediction stage. However the additive update term, derived based on an MC-version of the Kalman gain formula, brings back a Gaussian closure approximation to the empirical filtering density.

As a sequel to our recent work on an iterated gain-based stochastic filter (IGSF) [14] incorporating an iterative form of additive updates on the predicted particles, our present aim is to propose a substantively modified version of the algorithm in order to introduce an explicit non-Gaussian representation of the filtering density and an improved exploration of the process state space during the iterated updating stage. As with the IGSF, the iterations over a given time-step here are also aimed at driving the innovation term to a zero-mean random variable. This is consistent with the original aim of a stochastic filter as described by the Kushner–Stratonovich (KS) equation [11], which is generally achieved by designing the temporal recursion such that the innovation process is reduced to a zero-mean martingale. The first part of the current proposal is to develop the iterative and additive update through an annealing-type parameterization using an artificial diffusion parameter (ADP). In addition, non-Gaussian representations of the prediction and filtering densities are now provided through Gaussian sums. Specifically, the iterations in the update stage require ADP-parameterized repetitive computations of gain-like coefficient matrices $\boldsymbol{K}_i^l$ (i being the temporal recursion step and l the iteration index for a fixed i), consistent with the nonlinear KS equation, with the initial guess $\boldsymbol{K}_i^0$ evaluated on similar lines as in an ensemble square root filter (EnSRF) [15]. In addition to capturing non-Gaussianity in the posterior density, the Gaussian sum filter bank [16] also helps exploring the phase space of the state variables better and the added diversity in the particles enables easier adaptation of the process dynamics with the measured variables. The ADP, which may be lowered to zero over successive iterations at a much faster rate (allowing even for a discontinuous scheduling) than is feasible with the conventional simulated annealing, also helps enhance the so called 'mixing property' [17] of the iterative update kernels. An attempt is made to provide adequate numerical evidence of the enhanced filter performance with the introduction of some of these novel elements.

2. Statement of the problem

Let (Ω, F, P) be a complete probability space with F_t, $t \geq 0$, being the σ-algebra generated by all the noise processes involved in the presentation to follow at a given time t. The collection of sets $\mathbf{N}_t := \{F_s: 0 \leq s \leq t\}$ defines the so called increasing 'filtration' as t increases. Also the time interval of interest $[0, \tau]$ is discretized as $0 = t_0 < t_1 ..... < t_i .... t_L = \tau$ with $\Delta t_i = (t_i, t_{i+1}]$. The process model describing the evolution of the so-called 'hidden' states of a continuous-time dynamical system containing an additive Brownian noise term (which may, among others, account for modelling errors) may be represented by the Ito stochastic differential equation (SDE) [18]

$$dX(t) = \mathcal{F}(X(t), \mu(t), t)dt + G(t)dB(t) \tag{2.1}$$

where the state vector $X(t) \in \mathbb{R}^{n_x}$ is a time-continuous signal, $\mathcal{F}: \mathbb{R}^{n_x} \times \mathbb{R}^{n_\mu} \times \mathbb{R}^+ \to \mathbb{R}^{n_x}$ is the system transition function, $B(t)=\{B^{(r)}(t): r \in [1, q]\}$ is a q-dimensional vector of independently evolving zero-mean F_t-Brownian motion processes with $B^{(r)}(0) = 0$ and $E\{(B^{(r)}(t) - B^{(r)}(s))^2\} = t - s$, where E denotes the expectation with respect to the probability measure P, and $G: \mathbb{R}^+ \to \mathbb{R}^{n_x \times r}$ is the diffusion or volatility co-efficient matrix. System identification typically involves estimating the uncertain or inadequately known parameters $\mu(t) \in \mathbb{R}^{n_\mu}$ in the system model and a solution, within the stochastic filtering framework, requires declaring $\mu(t)$ as additional states. The original state space model (SSM) is thus augmented by allowing $\mu(t)$ to artificially evolve as a vector Brownian motion, as depicted through the following system of zero-drift SDEs:

$$d\mu(t) = G_\mu dB_\mu(t) \tag{2.2}$$

where $G_\mu \in \mathbb{R}^{n_\mu \times n_\mu}$ is the diffusion coefficient matrix and $B_\mu(t) \in \mathbb{R}^{n_\mu}$, a zero-mean Brownian noise vector process. In fact, restricting Eq. (2.2) over different time sub-intervals $\{(t_i, t_{i+1}], i = 0, 1, ... \}$, $\mu(t)$ may be interpreted as a collection of local Brownian motions (i.e. different mean vectors over different sub-intervals), or, more generally, as local martingales (see [1] for a definition of local martingales). The augmented state vector (with parameters as additional states) is now denoted as $\tilde{X} := [X^T, \mu^T]^T = \{\tilde{X}^{(j)} | j \in [1, J]\} \in \mathbb{R}^J; J = n_x + n_\mu$. The response of the dynamic system is partially observed through the noisy and continuous measurement process given by the SDEs (written below in the integral form):

$$Z(t) = \int_0^t A(\tilde{X}, s)ds + G_z^* B_z(t) \tag{2.3a}$$

or more appropriately, since the measurements arrive in a time-discrete manner, by a discrete algebraic counterpart of the above equation:

$$Z(t_{i+1}) := Z_{i+1} = \mathcal{H}(\tilde{X}_{i+1}, t_{i+1}) + G_z \iota \tag{2.3b}$$

Here $\tilde{X}_{i+1} = \tilde{X}(t_{i+1})$, $Z = \{Z^{(m)}: m \in [1, d]\} \in \mathbb{R}^d$ denotes the vector of measurements, ι is a d-dimensional vector of N(0, 1) independent normal random variables with coefficient matrix $G_z \in \mathbb{R}^{d \times d}$. Thus the covariance matrix of the discrete measurement noise vector $G_z \iota$ is given by $G_z G_z^T \in \mathbb{R}^{d \times d}$. The measurement vector function:

$$\mathcal{H} := \{H^{(k)}(\tilde{X}, t): \mathbb{R}^{n_x + n_\mu} \times \mathbb{R}^+ \to \mathbb{R}; k \in [1, d]\}$$

maps the signal process $\tilde{X}(t)$ to $\mathbb{R}^d$. Let $Z_{1:i} := \{Z_1, ..., Z_i\}^T$ denote the set of measurement vectors till $t = t_i$. The process Eqs. (2.1) and (2.2) may now be combined to yield the nonlinear SSM:

$$d\tilde{X}(t) = \tilde{\mathcal{F}}(\tilde{X}, t)dt + \tilde{G}(t)d\tilde{B}(t) \tag{2.4}$$

where $\tilde{\mathcal{F}} := \{\tilde{f}^{(j)}, j \in [1, J]\} \in \mathbb{R}^J$ and $\tilde{G}(t) \in \mathbb{R}^{J \times J}$ are respectively the nonlinear drift vector and the diffusion coefficient matrix. The