Applied Mathematics and Computation 233 (2014) 507–521

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Sequential Monte Carlo sampling in hidden Markov models of nonlinear dynamical systems ☆,☆☆

X. Zeng [a,b], M. Anitescu [b,*]

[a] Mathematics Department, Shanghai University, Shanghai, PR China
[b] Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Avenue, Building 240, Argonne, IL 60439-4844, USA

ARTICLE INFO

Keywords:
Hidden Markov model
Ordinary differential equation
Sequential Monte Carlo methods
Chemical process

ABSTRACT

We investigate the issue of which state functionals can have their uncertainty estimated efficiently in dynamical systems with uncertainty. Because of the high dimensionality and complexity of the problem, sequential Monte Carlo (SMC) methods are used. We investigate SMC methods where the proposal distribution is computed by maximum likelihood or by a linearization approach. We prove that the variance of the SMC method is bounded linearly in the number of time steps when the proposal distribution is truncated normal distribution. We also show that for a moderate large number of steps the error produced by approximation of dynamical systems linearly accumulates on the condition that the logarithm of the density function of noise is Lipschitz continuous. This finding is significant because the uncertainty in many dynamical systems, in particular, in chemical engineering systems, can be assumed to have this nature. We demonstrate our findings for a simple test case from chemical engineering. The theoretical findings provide a foundation for the parallel software SISTOS.

1. Introduction

Accountancy of expensive or consequential materials in the production process for chemical plants is a critical endeavor. For example, nuclear fuel reprocessing [1] facilities are expensive to build, are costly to operate, and have multiple stages and subsystems [2–4], losses of even minute amounts of the nuclear fuel can be therefore consequential. Currently, the Unite States has more than 15,000 chemical plant sites, which are required to file a risk management plan with the U.S. Environmental Protection Agency. Such plans consider both worst-case scenarios and alternative-case scenarios. Alternatives include the abnormal release of controlled materials or illegal plant interference. Whereas worst-case scenarios can be determined relatively efficiently with mass-balance approaches, it is difficult to detect long-term, slow releases, which are likely to be confused with measurement noise or other uncertain information. Hence, having superior accountancy will enable early detection of situations that may affect the safety of such plants and that may otherwise be hard to detect.

☆ Preprint ANL/MCS-P1852-0311, Argonne National Laboratory, Mathematics and Computer Science Division.

* Corresponding author.
E-mail address: anitescu@mcs.anl.gov (M. Anitescu).

http://dx.doi.org/10.1016/j.amc.2014.02.012
0096-3003/

Aiming for incorporating more information in the accountancy process, we investigate issues connected with modeling chemical reactions, in addition to achieving a superior estimation of the state. *The problem is the one of estimating functionals of the state of a nonlinear dynamical system, given measurements with noise.* To that end, we investigate the use of a hidden Markov model to estimate important functionals of the state of the system [5]. However, because of the high dimensionality and the nonlinear nature of the chemical reaction, to produce samples from the target distribution is difficult. Not the least of the difficulties is the fact that the large dimension of high-fidelity models may not allow us to store in memory all estimates at one time. Therefore, we investigate sequential sampling plans. Specifically, sequential Monte Carlo (SMC) methods (also called particle filters) produce scenarios of past and current time so that function of interest can be estimated. The most famous filter is the Kalman filter [6], which is used to make estimates in a Bayesian framework for linear Gaussian problems. For nonlinear systems, ensemble Kalman filters [7,8] or unscented Kalman filters [9] have emerged. One difficulty of such filters is their need to maintain at each step an approximation of the full covariance matrix and to update each of its entries, which results in exceedingly large memory and computational requirements [10].

In this paper, we investigate SMC methods that account for full nonlinearity. The proposal density is obtained by approximating the Hessian information of the logarithm of the targeted density at modal scores and can be loosely thought of as using a Kalman concept as a "preconditioner" for the sampling. By doing so, we get more accurate estimates while approaching the performance of Kalman methods. Our main goal is to determine whether SMC works for dynamical systems of the type encountered in chemical plants. We will thus justify the various assumptions we make about the dynamical systems by typical features encountered in real systems.

The key hypotheses investigated are that (1) the variance of the estimates produced by the method guarantees apriori the approach is tractable and (2) the proposal distributions we derive are sufficiently practical to be implemented. We demonstrate our findings on a synthetic example around the methane steam reforming reaction. We implement SISTOS, the Sequential Importance Sampling Toolkit of ODE Systems for statistical estimates of dynamical systems.

2. The statistical/uncertainty model

2.1. Hidden Markov model

The hidden Markov model (HMM) is a widely used approach in temporal pattern recognition, such as finance [11]; handwriting [12]; speech [13]; bioinformatics [14]; and, perhaps the area closest to that used here, weather forecast [10].

HMM is based on a state space model that includes two components. The first describes the state evolution of the dynamical system; while the second models noisy observations that are dependent on the state. Both parts include uncertainties. The mathematical formulation is described as follows

$$\boldsymbol{x}_i = M(\boldsymbol{x}_{i-1}) + \boldsymbol{\mu}_i, \tag{1}$$

$$\boldsymbol{z}_i = H(\boldsymbol{x}_i) + \boldsymbol{v}_i, \tag{2}$$

where i is the discrete time index. Here $\boldsymbol{x}_i \in \mathcal{X}$ is the state variable (for example, the moles of the reactants in the reaction at the beginning of step i) and $\boldsymbol{\mu}_i$ is a random variable that quantifies the uncertainty in the model (for example, in the outside feed or in the model itself). The variable $\boldsymbol{z}_i \in \mathcal{Z}$ are the observed quantities (for example, the outputs after some time period t_f) and $\boldsymbol{v}_i$ is a random variable that quantifies measurement error i. *The variables $\boldsymbol{\mu}_i$ and $\boldsymbol{v}_i$ are the core of the uncertainty model.* Their definition controls the lack of information about both the computational model and the measurement process, essentially the entire uncertainty space.

Another component of the model is $M : \mathcal{M} \subseteq \mathcal{X} \to \mathcal{X}$, a nonlinear function describing the evolution of the dynamical system over the period between two observations (for simplicity we assume it models an autonomous system, though that assumption can easily be relaxed). The mapping $H : \mathcal{X} \to \mathcal{Z}$ denotes the response of the measuring process. It is natural to assume that the $\mathcal{X} \subset \mathcal{R}^{n_x}$ and $\mathcal{Z} \subset \mathcal{R}^{n_z}$ are compact because many dynamical systems such as chemical plants have limited capacity and most state variables should be nonnegative.

It immediately follows that the system described in (1) has the Markov property

$$p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}, \boldsymbol{x}_{k-2}, \ldots, \boldsymbol{x}_0) = p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}).$$

To simplify the notation, we let $\boldsymbol{x}_{0:k} := (\boldsymbol{x}_0, \ldots, \boldsymbol{x}_k)^T$ and $\boldsymbol{z}_{0:k} := (\boldsymbol{z}_0, \ldots, \boldsymbol{z}_k)^T$. From the Markov property we obtain that the prior density is

$$p(\boldsymbol{x}_{0:k}) = \prod_{i=0}^{k} p(\boldsymbol{x}_i|\boldsymbol{x}_{i-1}). \tag{3}$$

From (2) we compute the density of the collection of observations conditioned on $\boldsymbol{x}_{0:k}$ as

$$p(\boldsymbol{z}_{0:k}|\boldsymbol{x}_{0:k}) = \prod_{i=0}^{k} p(\boldsymbol{z}_i|\boldsymbol{x}_i). \tag{4}$$