Journal of Computational Physics 258 (2014) 833–855

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Adaptive approximation of higher order posterior statistics

Wonjung Lee

Oxford Centre for Collaborative Applied Mathematics (OCCAM), Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford, OX1 3LB, UK

ARTICLE INFO

Article history:
Received 21 May 2013
Received in revised form 10 October 2013
Accepted 13 November 2013
Available online 19 November 2013

Keywords:
Nonlinear filtering
Data assimilation
Wiener chaos expansion
Uncertainty quantification

ABSTRACT

Filtering is an approach for incorporating observed data into time-evolving systems. Instead of a family of Dirac delta masses that is widely used in Monte Carlo methods, we here use the Wiener chaos expansion for the parametrization of the conditioned probability distribution to solve the nonlinear filtering problem. The Wiener chaos expansion is not the best method for uncertainty propagation without observations. Nevertheless, the projection of the system variables in a fixed polynomial basis spanning the probability space might be a competitive representation in the presence of relatively frequent observations because the Wiener chaos approach not only leads to an accurate and efficient prediction for short time uncertainty quantification, but it also allows to apply several data assimilation methods that can be used to yield a better approximate filtering solution. The aim of the present paper is to investigate this hypothesis. We answer in the affirmative for the (stochastic) Lorenz-63 system based on numerical simulations in which the uncertainty quantification method and the data assimilation method are adaptively selected by whether the dynamics is driven by Brownian motion and the near-Gaussianity of the measure to be updated, respectively.

1. Introduction

The assimilation of partial noisy observations into a hidden dynamical system for an improved estimate of a single trajectory has many significant and practical applications in science and engineering. Examples include navigational and guidance systems, radar tracking, sonar ranging, satellite and airplane orbit determination, the spread of hazardous plumes or pollutants, prediction of weather and climate in atmosphere–ocean dynamics [23,24,26,21,15,1,14]. When the underlying system is a continuous Markov process and the associated observations arrive sequentially in time, it is desirable to incorporate the data in real time via a recursive algorithm for practical reasons. One approach for this online estimation is the alternate application of the prediction to quantify the uncertainty delivered by the evolving system (or to approximate the prior measure) and the updating to incorporate the latest observation into the previous state estimation of the system (or to approximate the posterior measure) in a sequential fashion.

In the simplest case of linear dynamics together with linear observations corrupted by an independent Gaussian noise, the filtering problem can explicitly be solved by the Kalman filter which describes how the mean and covariance of the posterior measure evolve given the observations [23,24]. For nonlinear systems where the conditioned measure cannot be characterized by a Gaussian distribution, there occasionally exist analytic solutions but the applicability is quite limited [3]. As a result, work has been done to approximate the filtering solution in case of nonlinear systems. Traditional filters include the extended Kalman filter [15], the ensemble Kalman filter [14], the bootstrap filter [17,13], and the Gaussian mixture filter

E-mail address: leew@maths.ox.ac.uk.

0021-9991/$ – see front matter
http://dx.doi.org/10.1016/j.jcp.2013.11.015

[7,37,19,18]. Though they can in principle be applied to general physical models, each algorithm has its own theoretical and computational problems in practical implementation [36,38,13].

The extended Kalman filter provides an analytic mapping of the first two moments like the Kalman filter does. However, even when it is stable, significant error might occur while applying the extended Kalman filter to systems with strong nonlinearities because the algorithm uses a linearization or a closure model of the state equations for the prediction step. It is important to note that the updating method of the extended Kalman filter is the one from the Kalman filter and is developed for the Gaussian prior measure, and hence in general not suitable for nonlinear systems. It is also important to note that knowledge of the lower order first two moments, i.e., the mean and the covariance, is not sufficient to track the trajectory in the non-Gaussian scenario because being unable to accurately describe the tail behavior of the posterior measure implies that one is likely to lose the tracked object all together at some point over a long span of time.

The ensemble Kalman filter and the bootstrap filter employ a number of weighted Dirac masses, called particles, to characterize the evolving conditional probability distribution. In both filters, the prior measure is approximated in a straightforward manner via a repetitive application of a pre-developed integrator for each realization of the prescribed random inputs. Unfortunately the particle method can be impractical in case of high dimension because one may need a significantly large number of particles to achieve a certain degree of accuracy but the forward integration of the model is often very time-consuming and hence puts an upper limit on the number of particles [36,38,13]. Furthermore, it is not guaranteed in the ensemble Kalman filter that the ensemble members come from the true posterior distribution [14]. While the bootstrap filter resolves the convergence issue, the algorithm prunes particles with small weight in order to prevent the degeneracy in the weights [13]. This additional process subsequently causes other theoretical and computational problems in the successive filtering. The Gaussian mixture filter, in which Gaussian kernels are used instead of particles, also suffers from the degeneracy in the weights.

Our goal is to build a filtering approach that approximates not only the mean and covariance but the higher order moments of the posterior with high accuracy and efficiency. To this end we first focus on the data assimilation step which is indeed crucial for an accurate approximation of the filtering solution. From a careful examination of the data assimilation methods employed by the traditional filters, we are guided not to insist on a particular assimilation but to adaptively apply one of multiple updating algorithms depending on the characteristics of the prior measure for an improved assimilation performance. In order to achieve the adaptivity of the assimilation algorithm, it is inadequate to use moments, discrete measures and Gaussian kernels because conversion from one to another is not immediate. Hence the necessity for the parametrization of the conditioned measure in terms of something different arises and we here use the Wiener chaos expansion (WCE) which is an expansion of the system variables using a Galerkin projection of the model variables onto the space spanned by a fixed orthonormal polynomial basis in the random variables [5,34,12]. From the truncated WCE, one can immediately derive a variety of measure representations and the adaptation follows from applying one of the updating methods in the traditional filters.

The adaptivity enabled by the WCE is not limited to the updating step. The uncertainty quantification in the Wiener chaos approach is equivalent to finding the Wiener chaos coefficients due to its separation nature of deterministic effects from randomness in a rigorous and effective manner. In some cases the numerical methods based on the WCE are more efficient and accurate for the short time solutions than those based on the Monte Carlo simulations (see, e.g. [31,20] and the references therein). This higher prediction performance of the WCE is the main reason for the successful filtering of nonlinear deterministic systems in [27,43] where the data assimilation method is modeled upon the one used in the ensemble Kalman filter. Importantly, the WCE allows for multiple ways to numerically approximate the coefficients forward in time [27]. Therefore we can extend the adaptivity of the WCE further to the prediction step and solve the filtering problem by applying an optimal choice from several combinations of the two subsequent operations provided the conditions under which each algorithm outperforms are carefully studied. Our suggestion of the adaptive filtering is tested on the (stochastic) Lorenz-63 model that can possibly be driven by three independent Brownian motions and shown to significantly reduce the computational cost while retaining the accuracy that one can achieve from the bootstrap filter with a large number of particles.

The current paper is arranged as follows. Some traditional filters are reviewed in Section 2. The Wiener chaos expansion and its application to uncertainty quantification are introduced in Section 3. Section 4 is devoted to designing numerous data assimilation methods for the WCE. Numerical simulations are performed in Section 5 and conclusions are drawn in Section 6.

2. Filtering of dynamical systems

Consider the differential equation for a d-dimensional dynamical variable $X(t)$, $t \in \mathbb{R}^+ \cup \{0\}$,

$$\textbf{Forward model} \quad dX(t) = b\big(t, X(t)\big)\,dt + s\big(t, X(t)\big)\,dW(t) \tag{1}$$

where $b \in \mathbb{R}^d$ is the drift, $s \in \mathbb{R}^{d \times l}$ is the volatility, and Eq. (1) can possibly be driven by finitely many independent Brownian motions $W = (W_1, W_2, \ldots, W_l)^t$. Here the superscript t denotes the transpose. Suppose that the partial noisy observations y_n, $n \in \mathbb{N}$, associated with $X_n = X(n\Delta)$ for an inter-observation time $\Delta > 0$ satisfying

$$\textbf{Observational data} \quad y_n = PX_n + \eta_n, \quad \eta_n \sim \mathcal{N}(\mathbf{0}, R_n) \tag{2}$$