Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-18 (2016) 199–204

Reduced Order Gaussian Smoothing for Nonlinear Data Assimilation

Sarah A. King * **Wei Kang** ** **Kazufumi Ito** ***

* *Naval Research Laboratory, Monterey, CA 93943 USA (e-mail: sarah.king@nrlmry.navy.mil).*
** *Naval Postgraduate School, Monterey, CA 93943 USA (e-mail: wkang@nps.edu)*
*** *North Carolina State University, Raleigh, NC 24695 (e-mail: kito@ncsu.edu)*

Abstract: We investigate Gaussian filtering for data assimilation in numerical weather prediction (NWP). Data assimilation is the process of combining prior forecasts and observations to produce a system estimate. The prevailing data assimilation method in operational NWP centers is variational data assimilation. This method involves solving a cost function over a time window forming a maximum likelihood estimate. This method, however, requires the use of linearized models which in practice are difficult to produce and maintain. As an alternative we propose Gaussian smoothing for derivative-free, nonlinear data assimilation. Gaussian filters and their corresponding smoothers use numerical integration to evaluate the recursive Bayesian formulas for optimal filtering under Gaussian assumptions. This numerical integration typically requires many model evaluations making conventional Gaussian filtering/smoothing impractical for use in NWP. We will present a reduced order method for forming a Rauch-Tung-Striebel (RTS) type smoother. To do so we review the Bayesian filtering and smoothing equations and discuss an efficient numerical method for evaluating them. We will then discuss a numerical example using the Korteweg-de Vries equation to compare our technique to the standard variational approach.

Keywords: State Estimation, Large-Scale Systems, Order Reduction, Smoothing, Filtering

1. INTRODUCTION

Atmospheric data assimilation is used in numerical weather prediction (NWP) to create initial conditions for models to forecast future weather conditions. Data assimilation combines background information from climatology, or a previous forecast, and observations optimally to estimate the system state. In NWP applications this estimation problem is ill-posed: in a six hour time window 10^7 observations may be used to estimate 10^8 state variables used in models. These observations come from a variety of sources for example satellites, radiosondes, aircraft, and land surface instrumentation. The operational NWP models are often in the form of software packages, nonlinear, and expensive to run. The current Navy global environmental model (NAVGEM) (Hogan et al. (2014)) uses a Gaussian grid with $1,080 \times 540$ points and 50 vertical levels which corresponds to about a resolution of 37 km. NAVGEM includes physical parameterizations of nonlinear processes such as subgrid-scale moist processes and radiation. The physical parameterizations in NWP models can be stochastic and discontinuous and are difficult to linearize.

Currently the Navy's operational data assimilation system is a four-dimensional variational (4D-Var) system (Xu et al. (2005)). This is a maximum likelihood estimate of the state and is of the form of the minimization of a cost function. It assumes the evolving dynamical system is quasi-linear and requires the computation of the model Jacobian or tangent linear model. Creating and maintaining an accurate analytical tangent linear model for NWP use is difficult and expensive. For models with strong nonlinearity tangent linear models are only valid for a short time window. An added complication of this method is it not amenable to parallel computer architectures. This has the implication that while NWP models are running at progressively higher resolutions data assimilation is performed at much lower resolutions.

Gaussian filters provide a nonlinear state estimate without the use of linearized models and are highly parallelizable. They solve the nonlinear filtering problem by Gaussian density approximations, that is, we assume the filtering densities are well approximated by Gaussian distributions via moment matching. Using Bayes' formula with respect to the first two moments one can obtain the Gaussian filter. A variety of quadratures, e.g. Gauss Hermite (Ito and Xiong (2000)) and cubature (Wu et al. (2006)), are then used to evaluate the Bayesian formulas. The extended Kalman filter as well as the unscented Kalman filter may be viewed as special cases of Gaussian filters (Särkkä (2013)).

We develop a reduced order Gaussian smoother for data assimilation in order to condition our filtering solution on all measurements during our assimilation window. The solution based from the smoother is then comparable to

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.163

the solution from 4D-Var. For linear models with Gaussian error 4D-Var is equivalent to the Kalman smoother (Li and Navon (2001)). As with the Kalman smoother after the forward Gaussian filter the smoother runs a backward recursive correction over the forward filtering result. To form the smoother we discuss the Bayesian smoothing equations and then apply efficient quadrature to evaluate the equations to form the filter/smoother combination.

We begin Section 2 by reviewing the optimal filtering based on Bayes' equation and discussing central differences based quadrature. In Section 3 we discuss variational data assimilation. We present a numerical example in Section 4 using Rossby waves. We conclude with a discussion of the opportunities for Gaussian filtering and smoothing in data assimilation.

2. GAUSSIAN SMOOTHING

We begin by reviewing the nonlinear estimation problem using the Bayesian smoothing equations. Consider the discrete system modeled by

$$x(k) = f(x(k-1)) + w(k),\ x(0) = x_0 \tag{1}$$

where x_0 has covariance P_0 with the observation process given by

$$y(k) = h(x(k)) + v(k)$$

where $w(k)$ is the model error with covariance Q and $v(k)$ is the observation error with variance R. The nonlinear filtering problem is to find the conditional expectation $E[x(k)|Y_k]$ given the observations up to the current time k, i.e., $Y_k = \{y(j),\ 1 \leq j \leq k\}$. The nonlinear smoothing problem is to find $E[x(k)|Y_T]$ given all the observation data across a window including future observations, i.e., $Y_T = \{y(j),\ 1 \leq j \leq T\}$. After the Gaussian filter performs a forward filtering pass, the smoother recursively computes corrections in a backward pass. The smoothing of the state estimate will then be conditioned on all measurements.

We will discuss the nonlinear gaussian smoother in the context of the Rauch-Tung-Striebel (RTS) smoother for linear systems. That is, a forward-backward smoothing over a fixed interval. Applying Bayes' rule to the smoother density $p(x_k, Y_T)$ gives

$$\begin{aligned} p(x_k, Y_T) &= \int_{\mathbb{R}^n} p(x_k, \xi | Y_T) d\xi \\ &= \int_{\mathbb{R}^n} p(\xi|Y_T) p(x_k|\xi, Y_T) d\xi. \end{aligned} \tag{2}$$

The state x_k is independent of future measurements $y_{k+1}, y_{k+2}, y_{k+3}, \ldots$ etc making it Markovian and

$$p(x_k|x_{k+1}, Y_T) = p(x_k|x_{k+1}, Y_k). \tag{3}$$

We may then apply (3) to (2) to arrive at

$$\begin{aligned} p(x_k|Y_T) &= \int_{\mathbb{R}^n} p(\xi|Y_T) p(x_k|\xi, Y_T) d\xi \\ &= p(x_k|Y_k) \int_{\mathbb{R}^n} \frac{p(\xi|Y_T) p(\xi|x_k)}{p(\xi|Y_k)} d\xi. \end{aligned} \tag{4}$$

The forward filtering pass computes $p(x_k|Y_k)$ and the backward smoothing pass recursively computes the smoothing density $p(x_k|Y_T)$. While evaluating (4) exactly is not practical we may approximate it using generalized framework developed in Särkkä (2008) and Särkkä and Hartikainen (2010) based on assumed density filtering. The joint Gaussian approximation is given by

$$p\left(\begin{pmatrix} x_k \\ x_{k+1} \end{pmatrix} \middle| Y_k\right) = N\left(\begin{pmatrix} x_{k|k} \\ x_{k+1|k} \end{pmatrix}, \begin{pmatrix} P_{k|k} & P_{xk} \\ P_{xk}^T & P_{k+1|k} \end{pmatrix}\right).$$

Then

$$x_{k+1|k} = \int_{\mathbb{R}^n} c_1 \cdot f(\xi) e^{-\frac{1}{2}(\xi - x_{k|k})^T P_{k|k}^{-1}(\xi - x_{k|k})} d\xi$$

$$P_{k+1|k} = Q + \int_{\mathbb{R}^n} c_1 \cdot (f(\xi) - x_{k+1|k})(f(\xi) - x_{k+1|k})^T \cdot e^{-\frac{1}{2}(\xi - x_{k|k})^T P_{k|k}^{-1}(\xi - x_{k|k})} d\xi$$

where

$$c_1 = \frac{1}{\sqrt{(2\pi)^n det(P_{k+1|k})}}.$$

The cross-covariance P_{xk} is given by

$$P_{xk} = \int_{\mathbb{R}^n} c_2 \cdot (\xi - x_{k+1|k})(f(\xi) - x_{k+1|k})^T \cdot e^{-\frac{1}{2}(\xi - x_{k+1|k})^T P_{k+1|k}^{-1}(\xi - x_{k+1|k})} d\xi$$

where

$$c_2 = \frac{1}{\sqrt{(2\pi)^n det(P_{k+1|k})}}.$$

Then the Gaussian approximation of $p_{k|k}^s$ with mean x_k^s is given by

$$x_k^s = x_{k|k} + G_k(x_{k+1}^s - x_{k+1|k}) \tag{5}$$

$$P_k^s = P_{k|k} + G_k(P_{k+1}^s - P_{k+1|k})G_k^T \tag{6}$$

where

$$G_k = P_{xk} P_{k+1|k}^{-1}.$$

2.1 Gaussian filter

In order to compute $p(x_k|Y_k)$ and $p(x_{k+1}|Y_T)$ we form the forward Gaussian filter assuming

$$p_{k-1|k-1} = N(x_{k-1|k-1}, P_{k-1|k-1}).$$

Then the Gaussian approximation of $p_{k|k-1}$ has mean $x_{k|k-1}$ with covariance $P_{k|k-1}$. The prediction step of the filter is

$$x_{k|k-1} = \int_{\mathbb{R}^n} c_1 \cdot f(\xi) e^{-\frac{1}{2}(\xi - x_{k-1|k-1})^T P_{k-1|k-1}^{-1}(\xi - x_{k-1|k-1})} d\xi$$

and

$$P_{k|k-1} = Q + \int_{\mathbb{R}^n} c_1 \cdot (f(\xi) - x_{k|k-1})(f(\xi) - x_{k|k-1})^T \cdot e^{-\frac{1}{2}(\xi - x_{k-1|k-1})^T P_{k-1|k-1}^{-1}(\xi - x_{k-1|k-1})} d\xi$$

where

$$c_1 = \frac{1}{\sqrt{(2\pi)^n det(P_{k-1|k-1})}}.$$

We assume $E_{k|k-1}[h(x(k))]$ may be approximated by a Gaussian with mean z and covariance P_{zz} defined by

$$z = \int_{\mathbb{R}^n} c_2 \cdot h(\xi) e^{-\frac{1}{2}(\xi - x_{k|k-1})^T P_{k|k-1}^{-1}(\xi - x_{k|k-1})} d\xi$$

and

$$P_{zz} = \int_{\mathbb{R}^n} c_2 \cdot (h(\xi) - z)(h(\xi) - z)^T \cdot e^{-\frac{1}{2}(\xi - x_{k|k-1})^T P_{k|k-1}^{-1}(\xi - x_{k|k-1})} dt$$