Analysis of the Ensemble Kalman Filter for Inverse Problems

Matthieu Bulté

Technische Universität München

Abstract

The EnKF (Ensemble Kalman Filter) is a widely adopted methodology for state estimation in partially and noisily observed dynamical systems. While its adaptation for estimating solutions of Bayesian Inverse Problems (BIPs) has shown many positive empirical results, the theoretical understanding of the method is still poor. In this paper, we present the rigorous analysis of the EnKF proposed by C. Shillings and A. Stuart [1] and demonstrate their theoretical results through numerical experiments.

Keywords: EnKF, Inverse Problems, Bayesian statistics, Optimisation

1. Introduction

The ensemble Kalman filter (EnKF) has had a large impact in the natural sciences over the past years. Using an ensemble of particles, the method was orginally used to approximate the solution of data assimilation problems [2], but can also be extended to approximate the solution of Bayesian inverse problems. While the method is well understood in the large ensemble limit, the method is often used by practitioners because of its robustness - even with a small number of particles. It is thus relevant to study different properties of the algorithm without considering the large ensemble limit.

This article is a review of the work of C. Shillings and A. Stuart [1]. Section 2 presents the EnKF applied to inverse problems and will draw links between the method and the optimization approach to inverse problems. A selection of theoretical results will then be presented and proved in Section 3, and empirically tested with numerical experiments in Section 4.

2. A fresh look at the EnKF

2.1. Inverse Problems

In this section, we will consider the EnKF for solving inverse problems. Given a continuous map $\mathcal{G} : \mathcal{X} \to \mathcal{Y}$ between two Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$, we would like to identify $u \in \mathcal{X}$ solving

$$y = \mathcal{G}(u) + \eta, \tag{1}$$

where η is the *observational noise*. A central element in solving inverse problems is the *least squares functional* $\Phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, measuring the model-data misfit given by

$$\Phi(u; y) = \frac{1}{2} \left\| \Gamma^{-1/2} \left(y - \mathcal{G}(u) \right) \right\|_{\mathcal{Y}}^2. \tag{2}$$

In this expression, Γ is a positive semi-definite operator describing the covariance structure of the observational noise η. In the general case, this problem is *ill-posed* and minimization of the data-model misfit through minimization of the least squares operator is not possible. However, solvability of this system can be improved by use of *regularization* methods.

As thouroughly described by A. Stuart [3] one successful approach to regularization of inverse problems is to consider the problem from a Bayesian point of view. There, the problem in (1) is replaced by considering the pair (u, y) as a joint random variable over $\mathcal{X} \times \mathcal{Y}$. Assuming that we can encode current knowledge about u in a *prior* distribution μ_0 with $u \sim \mu_0$ independent of the observational noise $\eta \sim \mathrm{N}(0, \Sigma)$, the solution of the *Bayesian inverse problem* is given by $u|y \sim \mu$ with

$$\mu(\mathrm{d}u) \propto \exp\left(-\Phi(u; y)\right) \mu_0(\mathrm{d}u). \tag{3}$$

We now investigate how the EnKF method can be used to estimate the solution given in (3) of the inverse problem.

2.2. The EnKF for inverse problems

Since the EnKF algorithm is designed to solve data assimilation problem, an artificial dynamic must be defined in order to apply the method to inverse problems. To that end, we consider a discrete dynamic over the space $\mathcal{X} \times \mathcal{Y}$ defined by the operator

$$\Xi(u, p) = (u, \mathcal{G}(u)). \tag{4}$$

We then complete this artificial process by defining the observational process, such that for each time step $n \in \mathbb{N}$ with associated state (u_n, p_n), the observation y_n is given by

$$y_n = p_n + \xi_n, \tag{5}$$

2

where $\{\xi_j\}_{j\in\mathbb{N}}$ is an i.i.d. sequence distributed according to $N(0,\Sigma)$, in which Σ is a positive semidefinite matrix.

TODO: it's really kind of weird here, because the discretization approaches from [2] and [1] do not really match.

The EnKF can then be used to estimate each step $n \in \mathbb{N}$ of the dynamical system through an unweighted set of J particles $\{u_n^{(j)}\}_{j=1}^J$ constructing the empirical measure

$$\mu_n \approx \hat{\mu}_n := \frac{1}{J}\sum_{j=1}^J \delta_{u_n^{(j)}}. \tag{6}$$

Furthermore, replacing the dynamic (5) in the standard update formulate of the EnKF given the following update rule for the ensemble of particles

$$u_{n+1}^{(j)} = u_n^{(j)} + C^{\mathrm{up}}(u_n)[C^{\mathrm{pp}}(u_n) + h^{-1}\Gamma]^{-1}(y_{n+1}^{(j)} - G(u_n^{(j)})) \tag{7}$$

with $C^{\mathrm{up}} = \mathrm{c\hat{o}v}(u,\mathcal{G}(u))$ and $C^{\mathrm{pp}} = \mathrm{c\hat{o}v}(\mathcal{G}(u),\mathcal{G}(u))$, in which $\mathrm{c\hat{o}v}$ is the empirical covariance based on the approximate measure $\hat{\mu}_n$. We can then consider cases where $\Sigma = \Gamma$, or where $\Sigma = 0$, i.e. where no artificial noise is added to the data. In this article, we will only consider the second case.

This application of the EnKF for solving inverse problems can be shown to properly approximate the distribution μ of $u|y$ in the large ensemble limit when the model $\mathcal{G}$ is linear and the noise η is Gaussian, see [4, 5]. However, since the EnKF relies on a linear approximation, the error term resulting from the model estimation is independent of the number of particles, see [6].

3. Study of the continous time limit

While we have mentionned some properties of the large ensemble limit, it is also interesting to consider what happens for $h \to 0$, that is, we want to consider the continuous time limit of the algorithm. With that in mind, we can take a carefull look at the update rule (7) and recognize an Euler-Maruyama discretization of the following coupled system of SDEs

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = C^{\mathrm{up}}(u)\Gamma^{-1}\left(y - \mathcal{G}(u^{(j)})\right) + C^{\mathrm{up}}(u)\Gamma^{-1}\sqrt{\Sigma}\frac{\mathrm{d}W^{(j)}}{\mathrm{d}t}, \tag{8}$$

where $\{W^{(j)}\}_{j=1}^J$ are independent Browninan motions.

4. Numerical results

5. Conclusion

References

[1] C. Schillings, A. M. Stuart, Analysis of the ensemble kalman filter for inverse problems, SIAM Journal on Numerical Analysis 55 (2017) 1264–1290.

[2] M. A. Iglesias, K. J. Law, A. M. Stuart, Ensemble kalman methods for inverse problems, Inverse Problems 29 (2013) 045001.

[3] A. M. Stuart, Inverse problems: a bayesian perspective, Acta Numerica 19 (2010) 451–559.

[4] M. Goldstein, D. Wooff, Bayes linear statistics: Theory and methods, volume 716, John Wiley & Sons, 2007.

[5] K. J. Law, H. Tembine, R. Tempone, Deterministic mean-field ensemble kalman filtering, SIAM Journal on Scientific Computing 38 (2016) A1251–A1279.

[6] O. G. Ernst, B. Sprungk, H.-J. Starkloff, Analysis of the ensemble and polynomial chaos kalman filters in bayesian inverse problems, SIAM/ASA Journal on Uncertainty Quantification 3 (2015) 823–851.