Analysis of the Ensemble Kalman Filter for Inverse Problems

Matthieu Bulté

17.01.2019

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 [Iglesias et al.(2013)Iglesias, Law, and Stuart], 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 [Schillings and Stuart(2017)]. 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*. For simplicity, we will only consider the case where $\mathcal{Y}$ is a finite dimensional space, but the result presented can mostly be extended to support general Hilbert spaces. 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 [Stuart(2010)] 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 $\eta \sim \mathrm{N}(0, \Gamma)$, 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

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

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

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

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

Replacing the dynamic (5) in the standard update formulate of the EnKF gives the following update rule for the ensemble of particles

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

with

$$\begin{aligned} C^{\mathrm{up}} &= \hat{\mathrm{cov}}(u, \mathcal{G}(u)), \\ C^{\mathrm{pp}} &= \hat{\mathrm{cov}}(\mathcal{G}(u), \mathcal{G}(u)), \end{aligned} \tag{8}$$

where $\hat{\mathrm{cov}}$ is the empirical covariance computed from 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 [Goldstein and Wooff(2007), Law et al.(2016)Law, Tembine, and Tempone]. 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 [Ernst et al.(2015)Ernst, Sprungk, and Starkloff].

3 Study of the continous time limit

3.1 EnKF as a gradient flow

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 that the equation corresponds to an Euler-Maruyama discretization of the following coupled system of SDEs

$$\frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} = C^{\mathrm{up}}(u)\Gamma^{-1}\left(y - \mathcal{G}(u^{(j)}) + \sqrt{\Sigma}\frac{\mathrm{d}W^{(j)}}{\mathrm{d}t}\right), \tag{9}$$

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

Furthermore, assuming $\mathcal{G}(u) = Au$ is a linear map and $\Sigma = 0$, we obtain a deterministic system of coupled ODEs which, given the definition of the empirical covariance is

$$\begin{aligned} \frac{\mathrm{d}u^{(j)}}{\mathrm{d}t} &= \hat{\mathrm{cov}}(u, Au)\Gamma^{-1}(y - Au^{(j)}) \\ &= \frac{1}{J}\sum_{k=1}^{J}\langle A(u^{(k)} - \overline{u}), y - Au^{(j)}\rangle_\Gamma (u^{(k)} - \overline{u}) \\ &= A^T\Gamma^{-1}(y - Au^{(j)})\hat{\mathrm{cov}}(u, u) \\ &= -\hat{\mathrm{cov}}(u, u)\mathrm{D}_u\Phi(u^{(j)}; y). \end{aligned} \tag{10}$$

This shows that in its continuous time limit, the EnKF algorithm performs a set of gradient descents for $\Phi(\cdot; y)$, coupled through a preconditioning of the ensemble by its empirical covariance. This result draws a link between the optimization and Bayesian approaches to inverse problems.

3.2 Properties of the gradient flow

We are now interested in studying properties of the coupled gradient flows presented in the previous section.

4 Numerical results

5 Conclusion

References

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

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

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

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

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

[Ernst et al.(2015)Ernst, Sprungk, and Starkloff] 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.