ANALYSIS OF ENSEMBLE FILTERS FOR DATA ASSIMILATION AND INVERSE PROBLEMS

Andrew Stuart

Mathematics Institute
University of Warwick

Supported by EPSRC, ERC and ONR

March 24th 2013, KAUST

Outline

References

1 "Ensemble Kalman methods for inverse problems"
MA Iglesias KJH Law and AM Stuart
Inverse Problems **29**(2013), 045001
`arxiv.1209.2736`

2 "Well-posedness of ensemble Kalman filters"
DB Kelly and AM Stuart
In preparation.

Outline

The Filtering Problem

Partially Observed Dynamics

Discrete-time dynamical system

$$z_{n+1} = \Xi(z_n). \tag{1}$$

with linear noisy observations

$$y_{n+1} = Hz_{n+1} + \eta_{n+1} \qquad \text{where } \eta_n \sim N(0, \Gamma).$$

State Estimation

Try to estimate z_n given $\{y_j\}_{j=1}^n$.

Approximate Gaussian Filters

Prediction Step

$$\widehat{z}_{n+1} = \Xi(z_n).$$

Analysis Step

$$z_{n+1} = \mathrm{argmin}_z \left(||C_{n+1}^{-\frac{1}{2}}(z - \widehat{z}_{n+1})||^2 + ||\Gamma^{-1/2}(y_{n+1} - Hz)||^2 \right)$$

Design Parameter

The operators C_{n+1} characterize model uncertainty and are design parameters.

The Ensemble Kalman Filter

Prediction Step

$$\widehat{z}^j_{n+1} = \Xi(z^j_n), \quad j \in \{1, \cdots, J\}.$$

Estimate model uncertainty:

$$\overline{z}_{n+1} = \frac{1}{J}\sum_{j=1}^J \widehat{z}^j_{n+1}$$
$$C_{n+1} = \frac{1}{J}\sum_{j=1}^J \widehat{z}^j_{n+1}(\widehat{z}^j_{n+1})^T - \overline{z}_{n+1}\overline{z}^T_{n+1}$$

Analysis Step

$$S_{n+1} = HC_{n+1}H^T + \Gamma, \qquad K_{n+1} = C_{n+1}H^T S^{-1}_{n+1}$$
$$z^j_{n+1} = (I - K_{n+1}H)\widehat{z}^j_{n+1} + K_{n+1}y^j_{n+1}, \quad j \in \{1, \cdots, J\}.$$

Perturbed Observations Data

$$y^j_n = y_n + \eta^j_n, \qquad \eta^j_n \sim N(0, \Gamma)$$

Outline

The Setting

We work in a setting that includes Lorenz '63, Lorenz '96 and Navier-Stokes on a 2D torus.

The Model

$$\frac{dz}{dt} + Az + B(z, z) = f$$

Assumptions 1

For all $w \in V$

$$\langle Aw, w \rangle \geq \lambda \|w\|^2, \quad \langle B(w, w), w \rangle = 0.$$

Assumptions 2

For all $w_i \in V$

$$\langle B(w_1, w_2), w_2 \rangle \leq K \|w_1\| \|w_2\| |w_2|, \quad \langle B(w_1, w_2), w_3 \rangle \leq K \|w_1\| \|w_2\| \|w_3\|.$$

Discrete Time Filter: Well-Posedness

Assumptions

$$\Gamma = \gamma^2 I, H = I.$$

Theorem (Kelly, AMS)

There is constant β independent of n such that

$$\mathbb{E}|z_n^j - z_n|^2 \leq \exp(2\beta n)\mathbb{E}|z_0^j - z_0|^2 + K(J)\Big(\frac{\exp(2\beta n) - 1}{\exp(2\beta) - 1}\Big)\gamma^2.$$

Continuous Time Limit

Scalings

$$\Gamma = \frac{1}{h}\Gamma_0, \quad z_n = z(nh), \quad z_n^j = z^j(nh), \quad h \ll 1.$$

Limiting SPDEs

$$\frac{dz^j}{dt} + Az^j + B(z^j, z^j) = f + CH^*\Gamma_0^{-1}\Big(\Gamma_0^{\frac{1}{2}}\frac{dW^j}{dt} - Hz\Big).$$

Coupling

Coupled through the empirical covariance C

Continuous Time Limit: Well-Posedness

Assumptions

$$\Gamma_0 = \gamma^2 I, H = I.$$

Theorem (Kelly, AMS)

There is constant β independent of t such that

$$\mathbb{E}|z^j(t) - z(t)|^2 \leq \exp(2\beta t)\mathbb{E}|z^j(0) - z(0)|^2 + K(J)\Big(\frac{\exp(2\beta t) - 1}{\exp(2\beta) - 1}\Big)\gamma^2.$$

Outline

The Inverse Problem

X, Y Banach spaces and $G : X \to Y$.

$$y = G(u) + \eta,$$
$$\eta \sim N(0, \Gamma)$$

Prior knowledge of subsurface properties, for example:

$$u \sim \mu_0.$$

Model/Data mismatch:

$$\Phi(u; y) = \frac{1}{2}||\Gamma^{-1/2}(y - G(u))||^2$$

Example 1. Porous Media Flow: Forward Problem

State Variable

h: pore pressure (head)

Parameter

$e^u = K$: permeability (hydraulic conductivity)

Single-phase Darcy Flow

$$\begin{aligned} -\nabla \cdot e^u \nabla h &= f, \quad x \in D \\ -e^u \nabla h \cdot \mathbf{n} &= 0, \quad x \in \partial D \end{aligned}$$

Example 1. Porous Media Flow: Inverse Problem

Pressure at well locations x^ℓ:

$$G^\ell(u) = h(x^\ell), \quad \ell \in \{1, \dots, L\},$$

Measurement operator:

$$G(u) = \Big(G^1(u), \dots, G^L(u)\Big)$$

Unknown

$e^u = K$: permeability. $u = \log(K) \in X := L^\infty(D)$.

Data

$$y = G(u) + \eta \in Y := \mathbb{R}^L.$$

Example 1. Porous Media Flow: Applications

Estimation of subsurface properties

- Hydrology
- Fossil fuel extraction: oil, shale gas
- Carbon sequestration
- Compressed air storage
- Nuclear waste burial

Example 2. Navier Stokes Equation: Forward Problem

State Variable

v: fluid velocity

Parameter

$u =:$ initial fluid velocity

Navier-Stokes Equation as an ODE in $\dot{L}^2_{\mathrm{div}}(\mathbb{T}^2)$

$$\frac{dv}{dt} + \nu A v + B(v, v) = f, \quad v(0) = u$$

Example 2. Navier Stokes Equation: Inverse Problem

Fluid velocity at a finite set of points in space-time:

$$G^{j,k}(u) = v(x_j, t_k), \quad (j,k) \in \{1, \cdots, J\} \times \{1, \cdots, K\}.$$

Measurement operator:

$$G(u) = \big(G^{1,1}(u), \cdots, G^{J,K}(u)\big),$$

Unknown

$$u \in X := \dot{L}^2_{\mathrm{div}}(\mathbb{T}^2).$$

Data

$$y = G(u) + \eta \in Y := \mathbb{R}^{JK}.$$

Example 2. Navier Stokes Equation: Applications

Determination of Initial Fluid Velocity Field

- Weather Forecasting
- Oceanography
- Atmospheric Chemistry

Filtering and Inverse Problems

Artificial Dynamics

Define

$$z = \begin{pmatrix} u \\ p \end{pmatrix}, \qquad \Xi(z) = \begin{pmatrix} u \\ G(u) \end{pmatrix}, \qquad z_{n+1} = \Xi(z_n).$$

and then data is, for $H = (0, I)$,

$$y_n = Hz_n + \eta_n \qquad \text{where } \eta_n \sim N(0, \Gamma).$$

We Are In General Setting Above

We can estimate z_n given $\{y_j\}_{j=1}^n$ and, in this particular inverse problems setting, u_n given $\{y_j\}_{j=1}^n$.

For Ensemble Methods

$$y_n^j = y + \eta_n^j, \qquad \eta_n^j \sim N(0, \Gamma)$$

Key Theoretical Result

Linear Span of Initial Ensemble

$$\mathcal{A} = \operatorname{span}\{u_0^j\}_{j=1}^J.$$

Theorem (Iglesias, Law, AMS)

$$u_n^j \in \mathcal{A} \text{ for all } (n, j) \in \mathbb{N} \times \{1, \cdots, J\}.$$

Implications

- Compare EnKF with Best Approximation (BA) in $\mathcal{A}$.
- Compare EnKF with Least Squares (LSQ) in $\mathcal{A}$.
- Study Effect of choice of $\mathcal{A}$.

Specific Case of Invariant Subspace Property

G. Li and A. Reynolds. *An iterative ensemble Kalman filter for data assimilation*. SPE Annual Technical Conference, 2007.

Sketch Proof

$$z_0^j = \begin{pmatrix} u_0^j \\ p_0 \end{pmatrix}.$$

Prediction Step

$$\begin{pmatrix} \widehat{u}_{n+1}^j \\ \widehat{p}_{n+1}^j \end{pmatrix} = \begin{pmatrix} u_n^j \\ G(u_n^j) \end{pmatrix}$$

Compute empirical covariance to estimate model uncertainty:

$$C_{n+1} = \begin{pmatrix} C_{n+1}^{uu} & C_{n+1}^{up} \\ (C_{n+1}^{up})^T & C_{n+1}^{pp} \end{pmatrix}$$

Analysis Step

$$u_{n+1}^j = u_n^j + C_{n+1}^{up}(C_{n+1}^{pp} + \Gamma)^{-1}\Big(y_{n+1}^j - G(u_n^j)\Big)$$

$$p_{n+1}^j = G(u_n^j) + C_{n+1}^{pp}(C_{n+1}^{pp} + \Gamma)^{-1}\Big(y_{n+1}^j - G(u_n^j)\Big)$$

Sketch Proof (Continued)

Define

$$\widetilde{p}_n^j = \widehat{p}_n^j - \frac{1}{J}\sum_{k=1}^{J} \widehat{p}_n^j$$
$$d_{n+1}^j \equiv (C_{n+1}^{pp} + \Gamma)^{-1}\Big(y_{n+1}^j - G(u_n^j)\Big)$$

EnKF updates can be written as

$$u_{n+1}^j = u_n^j + \frac{1}{J}\sum_{k=1}^{J} \langle \widetilde{p}_{n+1}^k, d_{n+1}^j \rangle u_n^k$$

EnKF mean (parameter) at the final time

$$\overline{u} = \sum_{k=1}^{J} \alpha_j u_0^j \in \mathcal{A} \equiv \mathrm{span}\{u_0^1, \ldots, u_0^J\}.$$

Outline

Best Approximation (BA)

Best Approximation in a Compact Set

$$
\begin{aligned}
u^\dagger &= \text{truth} \\
\mathcal{A} &= \text{compact subset of } X \\
u_{\mathrm{BA}} &= \operatorname{argmin}_{u \in \mathcal{A}} \left(\| u - u^\dagger \| \right)
\end{aligned}
$$

Regularized Least Squares (LSQ)

Minimization Over a Compact Set

$$\begin{aligned}\mathcal{A} &= \text{compact subset of } X \\ u_{\mathrm{LSQ}} &= \mathrm{argmin}_{u\in\mathcal{A}}\Big(\Phi(u;y)\Big)\end{aligned}$$

Truncated Iteration

$$\begin{aligned}\Phi(u_{k+1};y) &\le \Phi(u_k;y) \\ u_{\mathrm{LSQ}} &= u_K\end{aligned}$$

Recap of Algorithms

EnKF

$$\overline{u} = \sum_{j=1}^{J} \alpha_j u_0^j \in \mathcal{A} \equiv \mathrm{span}\{u_0^1, \dots, u_0^J\}.$$

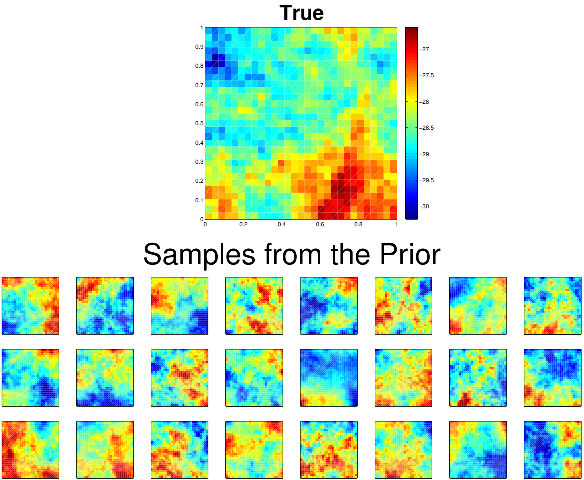
BA

$$u_{\mathrm{BA}} = \operatorname*{argmin}_{\alpha \in \mathbb{R}^J} \left\| u^\dagger - \sum_{j=1}^{J} \alpha_j u_0^j \right\|^2, \qquad u^\dagger = \text{truth}$$

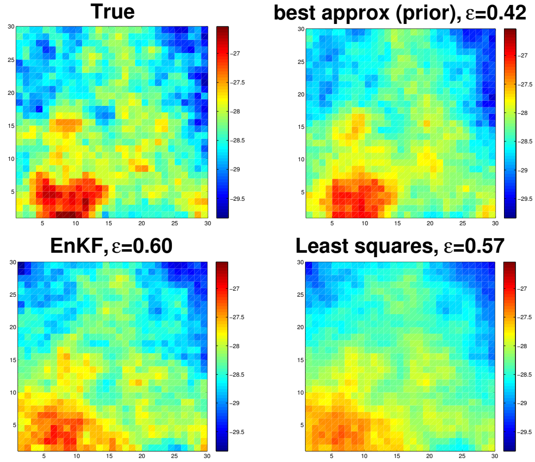
LSQ Variants On

$$u_{\mathrm{LSQ}} = \operatorname*{argmin}_{\alpha \in \mathbb{R}^J} \quad \Phi\left(\sum_{j=1}^{J} \alpha_j u_0^j ; y^\dagger\right), \qquad y = \text{data}$$

EnKF



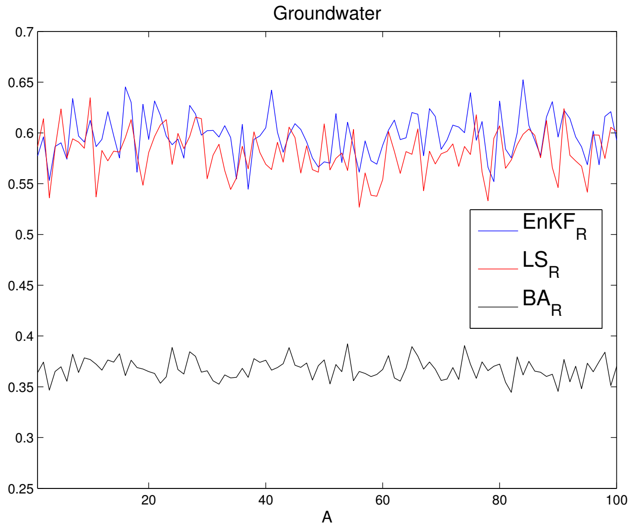
EnKF (Elliptic)



Cost EnKF= 1×10^2 forward models.

Cost LS= 3.6×10^3 forward models

Porous Media Flow

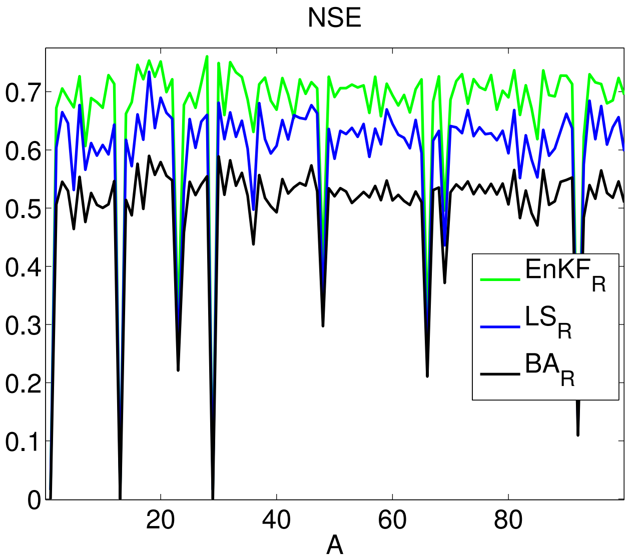


Navier-Stokes

Choice of Initial Ensemble

- For Porous media flow used draws from Gaussian prior.
- For Porous media flow also used Karhunen-Loeve (KL) basis.
- For NSE use draws from the attractor.
- For NSE also use draws from KL basis of empirical Gaussian.

Navier-Stokes



Outline

Scaling Limit

If we scale the noise covariance $\Gamma \to h^{-1}\Gamma_0$ and consider the limit $h \to 0$ then we obtain the following system of SDEs:

$$\frac{du^j}{dt} = \frac{1}{J}\sum_{k=1}^{J} \left\langle G(u^k) - \overline{G}, \Gamma_0^{-1}\left(\frac{dz^j}{dt} - G(u^j)\right)\right\rangle$$

$$\frac{dz^j}{dt} = y + \sqrt{\Gamma_0}\frac{dW^j}{dt}$$

$$\overline{G} = \frac{1}{J}\sum_{k=1}^{J} G(u^k).$$

Outline

Summary

- The EnKF is well-posed, in both discrete and continuous time settings, in fully observed case; filter instability reported in the literature involves interaction with numerical instability.

Summary

- The EnKF is well-posed, in both discrete and continuous time settings, in fully observed case; filter instability reported in the literature involves interaction with numerical instability.
- The EnKF for inverse problems is a derivative-free optimization technique which produces an approximation in the linear span of the initial ensemble $\mathcal{A}$.

Summary

- The EnKF is well-posed, in both discrete and continuous time settings, in fully observed case; filter instability reported in the literature involves interaction with numerical instability.
- The EnKF for inverse problems is a derivative-free optimization technique which produces an approximation in the linear span of the initial ensemble $\mathcal{A}$.
- The error incurred by EnKF is similar to that of derivative-based LSQ optimization techniques in $\mathcal{A}$.

Summary

- The EnKF is well-posed, in both discrete and continuous time settings, in fully observed case; filter instability reported in the literature involves interaction with numerical instability.
- The EnKF for inverse problems is a derivative-free optimization technique which produces an approximation in the linear span of the initial ensemble $\mathcal{A}$.
- The error incurred by EnKF is similar to that of derivative-based LSQ optimization techniques in $\mathcal{A}$.
- Both EnKF and LSQ produce errors of the same magnitude as BA.

Summary

- The EnKF is well-posed, in both discrete and continuous time settings, in fully observed case; filter instability reported in the literature involves interaction with numerical instability.
- The EnKF for inverse problems is a derivative-free optimization technique which produces an approximation in the linear span of the initial ensemble $\mathcal{A}$.
- The error incurred by EnKF is similar to that of derivative-based LSQ optimization techniques in $\mathcal{A}$.
- Both EnKF and LSQ produce errors of the same magnitude as BA.
- The choice of the initial ensemble $\mathcal{A}$ can have considerable impact on accuracy of EnKF.

http : //homepages.warwick.ac.uk/ ∼ masdr

In addition to the two papers highlighted at the start, upon which the talk is based, the following papers are also of background interest:

- **Elliptic Inverse Problem**: M.Dashti and A.M. Stuart, “Uncertainty quantification and weak approximation of an elliptic inverse problem.” SIAM J. Numerical Analysis 49(2011), 2524–2542.
- **Navier-Stokes Inverse Problem**: S.L. Cotter, M. Dashti and A.M.Stuart. "Bayesian inverse problems for functions and applications to fluid mechanics". Inverse Problems 25 (2009) 115008.
- **Invariant Subspace Property**: G. Li and A. Reynolds. *An iterative ensemble Kalman filter for data assimilation*. In SPE Annual Technical Conference and Exhibition, 2007.