Analysis of the EnKF for inverse problems

Plan

- Inverse problems (and ways to solve them)
- Adapting the EnKF method
- The twist ! (*)
- Results (theoretical & numerical)

Goals

me	you
keep you awake	stay awake
get you excited about (*)	keep big picture in mind.
you learn something	

Inverse problems

$$y = G(u) + \eta \qquad \text{(IP)}$$

with $y \in Y$, $u \in X$, $\eta \sim N(0, \Gamma)$.

How well can we reconstruct u from y?

Key element.

$$\Phi(u;y) = \frac{1}{2} \| \Gamma^{-1}(y - G(u)) \|_Y^2 \quad , \text{ the least squares functional}$$

Optimization

$$u^* = \underset{u \in X}{\operatorname{argmin}} \ \Phi(u,y) + \|u - u_0\|_C^2$$

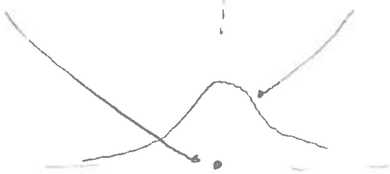
Often solved by some form of gradient descent.

Bayesian

$u \sim \mu_0$, $\eta \sim N(0, \Gamma)$ indep.

$$u^* = u \mid y \sim \mu$$

with $\mu(du) \propto \exp(-\Phi(u,y))\, \mu_0(du)$



NOTE. For $\mu_0 = N(u_0, C)$, we get $\mu(du) \propto \exp\left[-\left(\Phi(u;y) + \|u - u_0\|_C^2\right)\right] du$

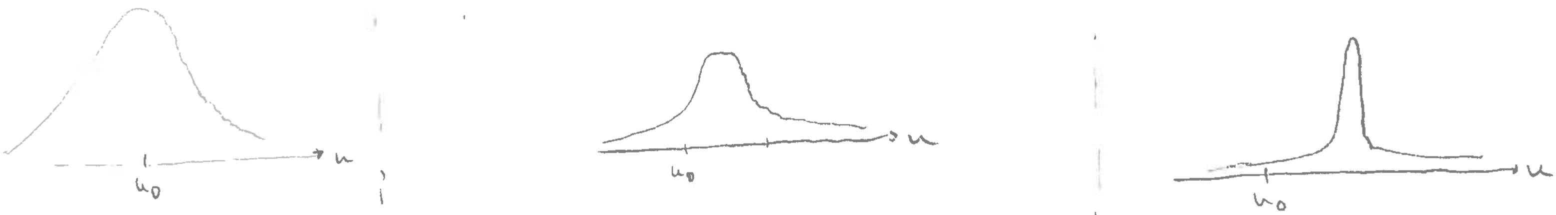
Examples

EnKF Approximate a partially, noisy observed dynamical system ? by an evolving ensemble of particles $\{u_n^{(j)}\}_{j=1}^{J}$

Idea. Interpolate between μ_0 and μ.

$t = 0 \longrightarrow t \longrightarrow t = 1$

μ_0 | $\mu_t \propto \exp(-t\,\Phi(u;y))\,\mu_0(du)$, | $\mu_1 = \mu$



Define a step size h, and approximate this dynamics using the EnKF for the first $N = 1/h$ steps.

Plug in the artificial dynamics.

$$u_{n+1}^{(j)} = u_n^{(j)} + C^{up}(u_n)\left[C^{pp}(u_n) + h^{-1}\Gamma\right]^{-1}\left(y_{n+1}^{(j)} - G(u_n^{(j)})\right) \qquad j=1,\dots,J$$

where
- $y_{n+1}^{(j)} = y + \Upsilon_{n+1}^{(j)}$ $\qquad \{\Upsilon_n^{(j)}\}$ iid $N(0,\Sigma)$
- $C^{up}(u) = \widehat{\mathrm{cov}}(u, G(u)) = \frac{1}{J}\sum_{j=1}^{J}(u^{(j)}-\bar{u})(G(u^{(j)})-\bar{G})^T$
- $C^{pp}(u) = \widehat{\mathrm{cov}}(G(u), G(u))$

How good is it?
- linear + gaussian noise: convergence as $J\to\infty$
- otherwise: complicated.

What about letting J fix (realistic) and $h\to 0$?

Continuous limit

Rewriting update rule, and letting $h\to 0$ we obtain an Euler-Maruyama discretization of

$$\frac{du^{(j)}}{dt} = \underbrace{C^{up}(u)\Gamma^{-1}(y - G(u^{(j)}))}_{\text{deterministic}} + \underbrace{C^{up}(u)\Gamma^{-1}\sqrt{\Sigma}\,\frac{dW^{(j)}}{dt}}_{\text{random}} \qquad \text{(SDE)}$$

Assumptions! $G(\cdot) = A$ is linear, $Y = \mathbb{R}^k$, $\Sigma = 0$

Then (SDE) becomes

$$\frac{du^{(j)}}{dt} = C^{up}(u)\Gamma^{-1}(y - Au^{(j)}) = \widehat{\mathrm{cov}}(u, Au)\Gamma^{-1}(y - Au^{(j)})$$

$$= \widehat{\mathrm{cov}}(u,u)A^T\Gamma^{-1}(y - Au^{(j)}) = \underbrace{\boxed{-\widehat{\mathrm{cov}}(u,u)D_u\Phi(u;y)}}_{\text{(the twist)}} \qquad \text{(ODE)}$$

Each particle performs a preconditioned gradient descent for Φ! (Demo)

$$\mathrm{IP} \to \mathrm{Opt}$$
$$\mathrm{IP} \to \mathrm{Bayes} \to \dots\ \mathrm{EnKF,\ Limit,\ Assumptions} \to \mathrm{opt,\ with\ benefits!}$$

The Analysis (noise free: $y = Au^{\dagger}$)

I) (ODE) has an unique solution

II) The ensemble collapses to its mean for $t \to \infty$

} We will prove jointly

III) The EnKF converges to the best approx of the sol } we will not prove

Defs $e^{(j)} = u^{(j)} - \bar{u}$, $\quad r^{(j)} = u^{(j)} - u^{\dagger}$, $\quad (E)_{ij} = \langle Ae^{(i)}, Ae^{(j)} \rangle_{\Gamma}$

Lemma: $E \xrightarrow{t \to \infty} 0$

Proof Messaging: $\frac{dE}{dt} = -\frac{2}{J} E^2$ with $E(0) = Q^T \Lambda(0) Q$ $\quad \Lambda(t) = \mathrm{diag}(\lambda^{(1)}(t), \dots, \lambda^{(J)}(t))$

$\Rightarrow E(t) = Q^T \Lambda(t) Q$ and $\frac{d\lambda^{(j)}}{dt} = -\frac{2}{J} \lambda^{(j)}$

$$\Rightarrow \lambda^{(j)}(t) = \left\{ \begin{array}{l} \left(\frac{2}{J} t + \frac{1}{\lambda^{(j)}(0)} \right)^{-1} \text{ for } \lambda^{(j)}(0) \neq 0 \\ 0, \text{ otherwise} \end{array} \right\} = \mathcal{O}(t^{-1}) \text{ as } \underbrace{t \to \infty}_{\text{we're not allowed to do this, yet!}}$$

Existence (I) One can show $\frac{du^{(j)}}{dt} = -\frac{1}{J} \sum_{k=1}^{J} \langle Ar^{(j)}, Ae^{(k)} \rangle_{\Gamma} e^{(k)}$

rhs is locally Lipschitz in u (polynomial) $\Rightarrow$ sol exists on some $[0, T)$.

$e^{(k)}$ is bounded? ✓

$r^{(j)}$ is bounded? Yes, with similar proof as $e^{(k)}$ ✓

$\Rightarrow$ rhs is bounded, no blow up, ✓ sol exists in $\mathbb{R}_+$.

Collapse to mean (II) Trivially follows from existence (I) and Lemma.

We note $\lambda^{(j)}(t) = \mathcal{O}(t^{-1}) \Rightarrow \|E\|_2 = \mathcal{O}(t^{-1/2})$

Best approx (III). One can show that $u^{(j)}(t) \in \mathrm{span}\{u^{(j)}(0)\}_{j=1}^{J} =: X_0$ with

. We can decompose $r^{(j)}(t)$ as $r_0^{(j)} \in AX_0$ and $r_{\perp}^{(j)} \in AX_0^{\perp}$

Then $\dim AX_0 = \min\{J-1, \dim Y\}$

and $r_0^{(j)}(t) = \mathcal{O}(t^{-1})$ for $t \to \infty$.