Data Assimilation: Hands-On

Claudia Schillings, Matei Hanu

CEMRACS Data Assimilation and Reduced Modeling for High Dimensional Problems

Outline

Mathematical Formulation of the Problem

We assume a model of the unknown z in the form of

$$\begin{aligned} z_{n+1} &= \Psi(z_n) + \zeta_n\,, \qquad n \in \mathbb{N} \\ z_0 &\sim \mathcal{N}(m_0, C_0) \end{aligned}$$

with $\Psi \in \mathcal{C}(\mathbb{R}^{n_z}, \mathbb{R}^{n_z})$, $\zeta = (\zeta)_n$ an iid sequence with $\zeta_0 \sim \mathcal{N}(0, \Sigma)$, $\Sigma > 0$, z_0 and ζ are assumed to be independent.

There is a true trajectory of z that produces noisy observations

$$y_{n+1} = H z_{n+1} + \eta_{n+1}\,, \qquad n \in \mathbb{N}$$

with $H \in \mathcal{L}(\mathbb{R}^{n_z}, \mathbb{R}^{n_y})$ and $\eta = (\eta)_n$ an iid sequence, independent of (z_0, ζ) with $\eta_1 \sim \mathcal{N}(0, \Gamma)$, $\Gamma > 0$.

The aim of **data assimilation** is to characterize the conditional distribution of z_n given the observations.

Mathematical Formulation of the Problem

The function Ψ is often the solution operator for an ordinary differential equation of the form

$$\frac{\mathrm{d}z}{\mathrm{d}t} = f(z)\,, \quad z(0) = z_0$$

for $t \in (0, \infty)$. Under suitable assumptions on the rhs., we set

$$\Psi(z) = \Psi(z; h)$$

as the solution operator over h time units, where h is the time between the observations. For simplicity, we assume equally spaced time points, i.e. $z_j = z(jh)$. Further, we denote by $\Psi^{(j)}(\cdot)$ the j-fold composition of Ψ with itself.

One-dimensional Linear Dynamics

For some scalar $\lambda \in \mathbb{R}$, we consider

$$\Psi(z) = \lambda z \quad H(z) = z\,.$$

Discuss for $|\lambda| \neq 1$ and $|\lambda| = 1$ as well as $\sigma = 0$ and $\sigma > 0$ the dynamics of the covariance c_j (see part 1, slide 69).

One obtains for the covariance the following recursive formula:

$$c_{n+1} = \frac{\gamma^2(\lambda^2 c_n + \sigma^2)}{\gamma^2 + \lambda^2 c_n + \sigma^2} := g(c_n).$$

To obtain the fixed points one has to solve the equation:

$$c^\star = g(c^\star).$$

The equation yields a second degree polynomial which has one positive and one negative root, if $\lambda\sigma\gamma \neq 0$, which are given by

$$c^\star_\pm = \frac{-(\gamma^2 + \sigma^2 - \gamma^2\lambda^2) \pm \sqrt{(\gamma^2 + \sigma^2 - \gamma^2\lambda^2)^2 + 4\lambda^2\gamma^2\sigma^2}}{2\lambda^2}.$$

One can show that $c_0 \geq 0$ implies $c_j \geq 0$ and $c_j \leq \gamma^2$ for all $j \in \mathbb{N}_0$. Furthermore, one obtains the following convergence results for the covariance.

		$\sigma^2 = 0$	$\sigma^2 > 0$
$\lvert\lambda\rvert < 1$		$\mathsf{c}_n \to 0 (exponentially)$	$\mathsf{c}_n \to c_+^\star (exponentially)$
$\lvert\lambda\rvert = 1$		$\mathsf{c}_n \to 0 (algebraically)$	$\mathsf{c}_n \to c_+^\star (exponentially)$
$\lvert\lambda\rvert > 1$		$\mathsf{c}_n \to c_-^\star (exponentially)$	$\mathsf{c}_n \to c_+^\star (exponentially)$

Ensemble Kalman filter (EnKf)

Implement the one-dimensional EnKf for an arbitrary operator Ψ and a linear operator H (see part 2, slide 33).
Discuss dynamics for the following three models:

1. one-dimensional nonlinear dynamics:
$$\Psi(z) = \alpha \sin(z)\,,$$
for some scalar $\alpha \in \mathbb{R}$.
2. one-dimensional linear autonomous dynamical system, where we consider
$$\frac{\mathrm{d}z}{\mathrm{d}t} = lz\,, \quad z(0) = z_0\,.$$
Then,
$$\Psi(z) = az \quad \text{with} \quad a = \exp(lh)\,,$$
for some scalar $l \in \mathbb{R}$.

3. Logistic dynamics:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = rz(k - z), \quad z(0) = z_0$$

Then,

$$\Psi(z) = \frac{k}{1 + \exp(-rkh)\left(\frac{k}{z} - 1\right)}$$

for some scalars $r, k \in \mathbb{R}$.

Sequential Importance Resamling (SIR)

Implement the one-dimensional EnKf for an arbitrary operator Ψ and a linear operator H (see part 2, slide 15).
Discuss dynamics for the same models as in EnKf.

Ensemble Kalman Inversion(EKI)

1. Implement the EKI for a linear forward operator G.
   Also apply variance inflation, i.e replace $C(u_n)$ with $C(u_n) + \delta$ for some scalar $\delta > 0$ (see part 2, slide 42). What can you observe?
2. Implement the EKI for an arbitrary forward operator G.

Consider in both cases unperturbed observations.