Multilevel ensemble Kalman filtering

Håkon Hoel[1] Kody Law[2] Raúl Tempone[3]

[1]Department of Mathematics, University of Oslo, Norway

[2]Oak Ridge National Laboratory, TN, USA

[3]Applied Mathematics and Computational Sciences, KAUST university, Saudi Arabia

UQAW 2016, Thuwal

Overview

Problem description

Consider the underlying and unobservable dynamics

$$u_{n+1} = \underbrace{u_n + \int_{n-1}^{n} a(u_t)\, dt + \int_{n-1}^{n} b(u_t)\, dW(t)}_{=:\Psi(u_n)}$$

with $u_n \in \mathbb{R}^d$, and Lipschitz continuous $a : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^{d\times\hat{d}}$.

And noisy observations

$$y_n = Hv_n + \gamma_n,$$

with i.i.d. $\gamma \sim N(0, \Gamma)$ and $H \in \mathbb{R}^{k\times d}$.

Objective: Let $Y_n := (y_1, y_2, \dots, y_n)$ and let Y_n^{obs} be a sequence of *fixed* observations. Construct an efficient method for tracking $u_n|(Y_n = Y_n^{\mathrm{obs}})$. That is, approximate

$$\mathbb{E}\left[\phi(u_n)|Y_n = Y_n^{\mathrm{obs}}\right]$$

for an observable $\phi : \mathbb{R}^d \to \mathbb{R}$.

Abuse of notation: will write $u_n|Y_n^{\mathrm{obs}}$ to represent $u_n|(Y_n = Y_n^{\mathrm{obs}})$.

Problem description

Consider the underlying and unobservable dynamics

$$u_{n+1} = \underbrace{u_n + \int_{n-1}^{n} a(u_t)\,dt + \int_{n-1}^{n} b(u_t)\,dW(t)}_{=:\Psi(u_n)}$$

with $u_n \in \mathbb{R}^d$, and Lipschitz continuous $a : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^{d\times\hat{d}}$.

And noisy observations

$$y_n = Hv_n + \gamma_n,$$

with i.i.d. $\gamma \sim N(0,\Gamma)$ and $H \in \mathbb{R}^{k\times d}$.

Objective: Let $Y_n := (y_1, y_2, \dots, y_n)$ and let Y_n^{obs} be a sequence of *fixed* observations. Construct an efficient method for tracking $u_n|(Y_n = Y_n^{\mathrm{obs}})$. That is, approximate

$$\mathbb{E}\left[\phi(u_n)|Y_n = Y_n^{\mathrm{obs}}\right]$$

for an observable $\phi : \mathbb{R}^d \to \mathbb{R}$.

Abuse of notation: will write $u_n|Y_n^{\mathrm{obs}}$ to represent $u_n|(Y_n = Y_n^{\mathrm{obs}})$.

Problem description

Consider the underlying and unobservable dynamics

$$u_{n+1} = \underbrace{u_n + \int_{n-1}^{n} a(u_t)\,dt + \int_{n-1}^{n} b(u_t)\,dW(t)}_{=:\Psi(u_n)}$$

with $u_n \in \mathbb{R}^d$, and Lipschitz continuous $a : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^{d\times\hat{d}}$.

And noisy observations

$$y_n = Hv_n + \gamma_n,$$

with i.i.d. $\gamma \sim N(0,\Gamma)$ and $H \in \mathbb{R}^{k\times d}$.

Objective: Let $Y_n := (y_1, y_2, \ldots, y_n)$ and let Y_n^{obs} be a sequence of *fixed* observations. Construct an efficient method for tracking $u_n|(Y_n = Y_n^{\mathrm{obs}})$. That is, approximate

$$\mathbb{E}\left[\phi(u_n)|Y_n = Y_n^{\mathrm{obs}}\right]$$

for an observable $\phi : \mathbb{R}^d \to \mathbb{R}$.

Abuse of notation: will write $u_n|Y_n^{\mathrm{obs}}$ to represent $u_n|(Y_n = Y_n^{\mathrm{obs}})$.

Overview

Kalman filtering

Consider the linear setting

$$\begin{aligned} u_{n+1} &= Au_n + \xi_{n+1}, & \xi_{n+1} &\sim N(0,\Sigma), \\ y_{n+1} &= Hu_{n+1} + \gamma_{n+1}, & \gamma_{n+1} &\sim N(0,\Gamma). \end{aligned} \tag{1}$$

Gaussianity of the filtering distribution

If $u_n | Y_n^{\text{obs}} \sim N(m_n, C_n)$, then under (1) both $u_{n+1}|Y_n^{\text{obs}}$ and $u_{n+1}|Y_{n+1}^{\text{obs}}$ will also be Gaussians.

$$u_{n+1}|Y_n^{\text{obs}} = (Au_n + \xi_{n+1})|Y_n^{\text{obs}} = A(u_n|Y_n^{\text{obs}}) + \xi_{n+1} \sim N(\underbrace{Am_n}_{=:\widehat{m}_{n+1}}, \underbrace{AC_nA^T + \Sigma}_{=:\widehat{C}_{n+1}}),$$

Can use Bayesian inference to show $u_{n+1}|Y_{n+1}^{\text{obs}} \sim N(m_{n+1}, C_{n+1})$, where

$$\begin{aligned} m_{n+1} &= (I - K_{n+1}H)\widehat{m}_{n+1} + K_{n+1}y_{n+1}^{\text{obs}}, \\ C_{n+1} &= (I - K_{n+1}H)\widehat{C}_{n+1}, \\ K_{n+1} &= \widehat{C}_{n+1}H^T(\Gamma + H\widehat{C}_{n+1}H^T)^{-1}. \end{aligned}$$

Kalman filtering

Consider the linear setting

$$\begin{aligned} u_{n+1} &= Au_n + \xi_{n+1}, & \xi_{n+1} &\sim N(0, \Sigma), \\ y_{n+1} &= Hu_{n+1} + \gamma_{n+1}, & \gamma_{n+1} &\sim N(0, \Gamma). \end{aligned} \tag{1}$$

Gaussianity of the filtering distribution

If $u_n | Y_n^{\text{obs}} \sim N(m_n, C_n)$, then under (1) both $u_{n+1} | Y_n^{\text{obs}}$ and $u_{n+1} | Y_{n+1}^{\text{obs}}$ will also be Gaussians.

$$u_{n+1} | Y_n^{\text{obs}} = (Au_n + \xi_{n+1}) | Y_n^{\text{obs}} = A(u_n | Y_n^{\text{obs}}) + \xi_{n+1} \sim N(\underbrace{Am_n}_{=:\widehat{m}_{n+1}}, \underbrace{AC_nA^T + \Sigma}_{=:\widehat{C}_{n+1}}),$$

Can use Bayesian inference to show $u_{n+1} | Y_{n+1}^{\text{obs}} \sim N(m_{n+1}, C_{n+1})$, where

$$\begin{aligned} m_{n+1} &= (I - K_{n+1}H)\widehat{m}_{n+1} + K_{n+1} y_{n+1}^{\text{obs}}, \\ C_{n+1} &= (I - K_{n+1}H)\widehat{C}_{n+1}, \\ K_{n+1} &= \widehat{C}_{n+1} H^T (\Gamma + H\widehat{C}_{n+1} H^T)^{-1}. \end{aligned}$$

Kalman filtering

Consider the linear setting

$$\begin{aligned} u_{n+1} &= Au_n + \xi_{n+1}, \qquad && \xi_{n+1} \sim N(0, \Sigma), \\ y_{n+1} &= Hu_{n+1} + \gamma_{n+1}, \qquad && \gamma_{n+1} \sim N(0, \Gamma). \end{aligned} \tag{1}$$

Gaussianity of the filtering distribution

If $u_n | Y_n^{\mathrm{obs}} \sim N(m_n, C_n)$, then under (1) both $u_{n+1} | Y_n^{\mathrm{obs}}$ and $u_{n+1} | Y_{n+1}^{\mathrm{obs}}$ will also be Gaussians.

$$u_{n+1} | Y_n^{\mathrm{obs}} = (Au_n + \xi_{n+1}) | Y_n^{\mathrm{obs}} = A(u_n | Y_n^{\mathrm{obs}}) + \xi_{n+1} \sim N(\underbrace{Am_n}_{=:\widehat{m}_{n+1}}, \underbrace{AC_nA^T + \Sigma}_{=:\widehat{C}_{n+1}}),$$

Can use Bayesian inference to show $u_{n+1} | Y_{n+1}^{\mathrm{obs}} \sim N(m_{n+1}, C_{n+1})$, where

$$\begin{aligned} m_{n+1} &= (I - K_{n+1}H)\widehat{m}_{n+1} + K_{n+1}y_{n+1}^{\mathrm{obs}}, \\ C_{n+1} &= (I - K_{n+1}H)\widehat{C}_{n+1}, \\ K_{n+1} &= \widehat{C}_{n+1}H^T(\Gamma + H\widehat{C}_{n+1}H^T)^{-1}. \end{aligned}$$

Kalman filtering formulas

1. The *prediction* step for $u_{n+1}|Y_n^{\mathrm{obs}} \sim N(\widehat{m}_{n+1}, \widehat{C}_{n+1})$:

Prediction

$$\widehat{m}_{n+1} = Am_n$$
$$\widehat{C}_{n+1} = AC_nA^T + \Sigma.$$

2. The *update* step. Update the prediction with the latest observation y_{n+1} into $u_{n+1}|Y_{n+1} \sim N(m_{n+1}, C_{n+1})$:

Update

$$m_{n+1} = (I - K_{n+1}H)\widehat{m}_{n+1} + K_{n+1}y_{n+1}^{\mathrm{obs}},$$
$$C_{n+1} = (I - K_{n+1}H)\widehat{C}_{n+1},$$
$$K_{n+1} = \widehat{C}_{n+1}H^T(\Gamma + H\widehat{C}_{n+1}H^T)^{-1}.$$

Overview

Ensemble Kalman filtering (Evensen 94)

In more complex settings, Kalman filtering has major drawbacks:

- In high dimensions, $d \gg 1$, storage and evolution of the full $d \times d$ covariance matrix C_n is very costly.
- Kalman filtering is not designed for nonlinear dynamics $u_{n+1} = \Psi(u_n)$; linearization is then needed, and it is hard to estimate the errors from linearization.

Ensemble Kalman filtering (EnKF) may overcome these drawbacks by propagating an ensemble of particles $\{v_n(\omega_i)\}_{i=1}^M \mapsto \{v_{n+1}(\omega_i)\}_{i=1}^M$ and using their sample moments in the predict and update steps.

Ensemble Kalman Filtering II

Predict

1. Compute (numerical solutions of) M particle paths one step forward

$$\widehat{v}_{n+1,i} = \Psi(v_{n,i}, \omega_i) \quad \text{for } i = 1, 2, \ldots, M.$$

2. Compute sample mean and covariance

$$\widehat{m}^{\mathrm{MC}}_{n+1} = E_M[\widehat{v}_{n+1}]$$
$$\widehat{C}^{\mathrm{MC}}_{n+1} = \mathrm{Cov}_M[\widehat{v}_{n+1}]$$

$$\text{where} \quad E_M[\widehat{v}_{n+1}] := \frac{1}{M}\sum_{i=1}^{M} \widehat{v}_{n+1,i}$$

$$\text{and} \quad \mathrm{Cov}_M[\widehat{v}_{n+1}] := E_M[\widehat{v}_{n+1}\widehat{v}^T_{n+1}] - E_M[\widehat{v}_{n+1}](E_M[\widehat{v}_{n+1}])^T.$$

Ensemble Kalman Filtering III

Update

1. Generate signal observations for the ensemble of particles

$$\tilde{y}_{n+1,i} = y_{n+1}^{\mathrm{obs}} + \gamma_{n+1,i} \quad \text{for } i = 1, 2 \ldots, M,$$

   with i.i.d. $\gamma_{n+1,1} \sim N(0, \Gamma)$.

2. Use signal observations to update particle paths

$$v_{n+1,i} = (I - K_{n+1}^{\mathrm{MC}} H)\widehat{v}_{n+1,i} + K_{n+1}^{\mathrm{MC}} \tilde{y}_{n+1,i},$$

$$\text{where} \quad K_{n+1}^{\mathrm{MC}} = \widehat{C}_{n+1}^{\mathrm{MC}} H^T (H \widehat{C}_{n+1}^{\mathrm{MC}} H^T + \Gamma)^{-1}.$$

Note: After the first step, all particles are correlated due to K_{n+1}^{MC}.

From EnKF to mean field EnKF

For studying convergence properties of EnKF it is useful to introduce the mean field EnKF (MFEnKF)

$$\text{Pr}\begin{cases}\widehat{v}^{\mathrm{MF}}_{n+1,i} &= \Psi(v^{\mathrm{MF}}_{n,i}, \omega_i)\\ \widehat{m}^{\mathrm{MF}}_{n+1} &= \mathbb{E}\left[\widehat{v}^{\mathrm{MF}}_{n+1,i}\right]\\ \widehat{C}^{\mathrm{MF}}_{n+1} &= \mathrm{Cov}[\widehat{v}^{\mathrm{MF}}_{n+1,i}],\end{cases} \quad \text{Up}\begin{cases}K^{\mathrm{MF}}_{n+1} &= \widehat{C}^{\mathrm{MF}}_{n+1}H^T(H\widehat{C}^{\mathrm{MF}}_{n+1}H^T+\Gamma)^{-1}\\ \tilde{y}_{n+1,i} &= y_{n+1}+\gamma_{n+1,i}\\ v^{\mathrm{MF}}_{n+1,i} &= (I-K^{\mathrm{MF}}_{n+1}H)v^{\mathrm{MF}}_{n+1,i}+K^{\mathrm{MF}}_{n+1}\tilde{y}_{n+1,i}.\end{cases}$$

and in comparison, EnKF

$$\text{Pr}\begin{cases}\widehat{v}_{n+1,i} &= \Psi(v_{n,i}, \omega_i)\\ \widehat{m}^{\mathrm{MC}}_{n+1} &= E_M[\widehat{v}_{n+1}]\\ \widehat{C}^{\mathrm{MC}}_{n+1} &= \mathrm{Cov}_M[\widehat{v}_{n+1}]\end{cases} \quad \text{Up}\begin{cases}K^{\mathrm{MC}}_{n+1} &= \widehat{C}^{\mathrm{MC}}_{n+1}H^T(H\widehat{C}^{\mathrm{MC}}_{n+1}H^T+\Gamma)^{-1}\\ \tilde{y}_{n+1,i} &= y_{n+1}+\gamma_{n+1,i}\\ v_{n+1,i} &= (I-K^{\mathrm{MC}}_{n+1}H)\widehat{v}_{n+1,i}+K^{\mathrm{MC}}_{n+1}\tilde{y}_{n+1,i}.\end{cases}$$

- When underlying dynamics is linear with Gaussian additive noise and u_0 Gaussian, it holds that $\mu^{\mathrm{MF}}_n(dx) = \mathbb{P}\left(u_n \in dx | Y^{\mathrm{obs}}_n\right)$, where $\mu^{\mathrm{MF}}_n = \mathrm{Law}(v^{\mathrm{MF}}_{n,i})$.
- In nonlinear settings, we use as approximation goal

$$\int_{\mathbb{R}^d} \phi(x)\mu^{\mathrm{MF}}_n(dx) (\approx \mathbb{E}\left[\phi(u_n) | Y^{\mathrm{obs}}_n\right])$$

From EnKF to mean field EnKF

For studying convergence properties of EnKF it is useful to introduce the mean field EnKF (MFEnKF)

$$\text{Pr}\begin{cases}\widehat{v}^{\mathrm{MF}}_{n+1,i} &= \Psi(v^{\mathrm{MF}}_{n,i}, \omega_i)\\ \widehat{m}^{\mathrm{MF}}_{n+1} &= \mathbb{E}\left[\widehat{v}^{\mathrm{MF}}_{n+1,i}\right]\\ \widehat{C}^{\mathrm{MF}}_{n+1} &= \mathrm{Cov}[\widehat{v}^{\mathrm{MF}}_{n+1,i}],\end{cases} \quad \text{Up}\begin{cases}K^{\mathrm{MF}}_{n+1} &= \widehat{C}^{\mathrm{MF}}_{n+1}H^T(H\widehat{C}^{\mathrm{MF}}_{n+1}H^T+\Gamma)^{-1}\\ \tilde{y}_{n+1,i} &= y_{n+1}+\gamma_{n+1,i}\\ v^{\mathrm{MF}}_{n+1,i} &= (I-K^{\mathrm{MF}}_{n+1}H)v^{\mathrm{MF}}_{n+1,i}+K^{\mathrm{MF}}_{n+1}\tilde{y}_{n+1,i}.\end{cases}$$

and in comparison, EnKF

$$\text{Pr}\begin{cases}\widehat{v}_{n+1,i} &= \Psi(v_{n,i}, \omega_i)\\ \widehat{m}^{\mathrm{MC}}_{n+1} &= E_M[\widehat{v}_{n+1}]\\ \widehat{C}^{\mathrm{MC}}_{n+1} &= \mathrm{Cov}_M[\widehat{v}_{n+1}]\end{cases} \quad \text{Up}\begin{cases}K^{\mathrm{MC}}_{n+1} &= \widehat{C}^{\mathrm{MC}}_{n+1}H^T(H\widehat{C}^{\mathrm{MC}}_{n+1}H^T+\Gamma)^{-1}\\ \tilde{y}_{n+1,i} &= y_{n+1}+\gamma_{n+1,i}\\ v_{n+1,i} &= (I-K^{\mathrm{MC}}_{n+1}H)\widehat{v}_{n+1,i}+K^{\mathrm{MC}}_{n+1}\tilde{y}_{n+1,i}.\end{cases}$$

- When underlying dynamics is linear with Gaussian additive noise and u_0 Gaussian, it holds that $\mu^{\mathrm{MF}}_n(dx) = \mathbb{P}\left(u_n \in dx | Y^{\mathrm{obs}}_n\right)$, where $\mu^{\mathrm{MF}}_n = \mathrm{Law}(v^{\mathrm{MF}}_{n,i})$.
- In nonlinear settings, we use as approximation goal

$$\int_{\mathbb{R}^d} \phi(x)\mu^{\mathrm{MF}}_n(dx) (\approx \mathbb{E}\left[\phi(u_n) | Y^{\mathrm{obs}}_n\right])$$

Convergence of EnKF

Theorem 1 (Le Gland et al. (2009))

Consider the dynamics and observations,

$$u_{n+1} = f(u_n) + \xi_{n+1}, \quad \xi_{n+1} \sim N(0, \Sigma),$$
$$y_{n+1} = Hu_{n+1} + \gamma_{n+1}, \quad \gamma_{n+1} \sim N(0, \Gamma),$$

and assume $u_0 \in L^p(\Omega)$ for any $p \geq 1$, and that

$$\max(|f(x) - f(x')|, |\phi(x) - \phi(x')|) \leq C|x - x'|(1 + |x|^s + |x'|^s), \text{ for an } s \geq 0.$$

Then, for the EnKF update ensemble $\{v_{n,i}\}_{i=1}^M$,

$$\sup_{M \geq 1} \sqrt{M} \left(\mathbb{E}\left[\left| \sum_{i=1}^{M} \frac{\phi(v_{n,i})}{M} - \int_{\mathbb{R}^d} \phi(x) \mu_n^{\mathrm{MF}}(dx) \right|^p \right] \right)^{1/p} < \infty.$$

for any order $p \geq 1$ and finite n.

Extension to further nonlinear settings in [Law et al. (2014)].

Central step in proof:

$$\begin{aligned}\left\|\sum_{i=1}^{M}\frac{\phi(v_{n,i})}{M}-\int_{\mathbb{R}^d}\phi(x)\mu_n^{\mathrm{MF}}(dx)\right\|_p &\leq \left\|\sum_{i=1}^{M}\frac{\phi(v_{n,i})}{M}-\sum_{i=1}^{M}\frac{\phi(v_{n,i}^{\mathrm{MF}})}{M}\right\|_p \\ &+\left\|\sum_{i=1}^{M}\frac{\phi(v_{n,i}^{\mathrm{MF}})}{M}-\int_{\mathbb{R}^d}\phi(x)\mu_n^{\mathrm{MF}}(dx)\right\|_p \\ &=\mathcal{O}(\|v_{n,i}-v_{n,i}^{\mathrm{MF}}\|_{\hat{p}}+M^{-1/2})\end{aligned}$$

Computational cost of EnKF

- To meet the constraint

$$\left(\mathbb{E}\left[\left|\sum_{i=1}^{M}\frac{\phi(v_{n,i})}{M}-\int_{\mathbb{R}^d}\phi(x)\mu_n^{\mathrm{MF}}(dx)\right|^p\right]\right)^{1/p}=\mathcal{O}(\epsilon),$$

  one thus needs ensemble of size $M=\mathcal{O}(\epsilon^{-2})$.

- How does the computational cost increase if the EnKF dynamics has to be sampled using a numerical solver for which $\left|\mathbb{E}\left[\overline{\Psi}-\Psi\right]\right|=\mathcal{O}(\Delta t^{\alpha})$?
- Short answer (under additional assumptions): the cost increases to $\mathcal{O}(\epsilon^{-(2+1/\alpha)})$.

Overview

Multilevel EnKF (MLEnKF)

Prediction

- Compute an ensemble of particle paths on a hierarchy of accuracy levels

$$\widehat{v}^{\ell-1}_{n+1,i} = \Psi^{\ell-1}(v^{\ell-1}_{n,i}, \omega_{\ell,i}), \quad \widehat{v}^{\ell}_{n+1,i} = \Psi^{\ell}(v^{\ell}_{n,i}, \omega_{\ell,i}),$$

for the levels $\ell = 0, 1, \ldots, L$ and $i = 1, 2, \ldots, M_\ell$.

- Multilevel approximation of mean and covariance matrices:

$$\widehat{m}^{\mathrm{ML}}_{n+1} = \sum_{\ell=0}^{L} E_{M_\ell}[\widehat{v}^{\ell}_{n+1} - \widehat{v}^{\ell-1}_{n+1}],$$

$$\widehat{C}^{\mathrm{ML}}_{n+1} = \sum_{\ell=0}^{L} \mathrm{Cov}_{M_\ell}[\widehat{v}^{\ell}_{n+1}] - \mathrm{Cov}_{M_\ell}[\widehat{v}^{\ell-1}_{n+1}]$$

Notice the telescoping properties $\mathbb{E}\left[\widehat{m}^{\mathrm{ML}}_{n+1}\right] = \mathbb{E}\left[\widehat{v}^{L}_{n+1}\right]$ and $\mathbb{E}\left[\widehat{C}^{\mathrm{ML}}_{n+1}\right] = \mathrm{Cov}(\widehat{v}^{L}_{n+1}) + O(1/M_L)$.

MLEnKF update step

Update

For $\ell = 0, 1, \ldots, L$ and $i = 1, 2, \ldots, M_\ell$,

$$
\begin{aligned}
\tilde{y}^{\ell}_{n+1,i} &= y^{\mathrm{obs}}_{n+1} + \gamma^{\ell}_{n+1,i}, \qquad \text{i.i.d. } \gamma^{\ell}_{n+1,i} \sim N(0, \Gamma) \\
v^{\ell-1}_{n+1} &= (I - K^{\mathrm{ML}}_{n+1} H)\widehat{v}^{\ell-1}_{n+1,i} + K^{\mathrm{ML}}_{n+1}\tilde{y}^{\ell}_{n+1,i}, \\
v^{\ell}_{n+1,i} &= (I - K^{\mathrm{ML}}_{n+1} H)\widehat{v}^{\ell}_{n+1,i} + K^{\mathrm{ML}}_{n+1}\tilde{y}^{\ell}_{n+1,i},
\end{aligned}
$$

$$
\text{where} \quad K^{\mathrm{ML}}_{n+1} = \widehat{C}^{\mathrm{ML}}_{n+1} H^{\mathsf{T}} (H \widehat{C}^{\mathrm{ML}}_{n+1} H^{\mathsf{T}} + \Gamma)^{-1}.
$$

Convergence of MLEnKF

For observables $\phi : \mathbb{R}^d \to \mathbb{R}$, introduce notation

$$\mu_n^{\mathrm{ML}}(\phi) := \sum_{\ell=0}^{L} \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \phi(v_{n,i}^{\ell}) - \phi(v_{n,i}^{\ell-1}).$$

and

$$\mu_n^{\mathrm{MF}}(\phi) := \int_{\mathbb{R}^d} \phi(x) \mu_n^{\mathrm{MF}}(dx).$$

Question: Under what assumptions and at what cost can one achieve

$$\|\mu_n^{\mathrm{ML}}(\phi) - \mu_n^{\mathrm{MF}}(\phi)\|_{L^p(\Omega)} = \mathcal{O}(\epsilon)?$$

Assumption 1

Consider the dynamics

$$u_{n+1} = \Psi(u_n) = u_n + \int_n^{n+1} a(u_t)dt + \int_n^{n+1} b(u_t)dW(t), \quad n = 0, 1, \ldots$$

with $u_0 \in L^p(\Omega)$ *for all* $p \geq 1$ *and a hierarchy of numerical solvers* $\{\Psi^\ell\}_{\ell=0}^L$.

For any $p \geq 2$, $\ell \in \mathbb{N}$, *and* $u, v \in L^p(\Omega)$, *we assume that*

$$\|\Psi^\ell(u) - \Psi^\ell(v)\|_p \leq C\|u - v\|_p,$$
$$\|\Psi^\ell(u)\|_p \leq C(1 + \|u\|_p).$$

Assumption 2

Assume the observable $\phi : \mathbb{R}^d \to \mathbb{R}$ *satisfies*

$$|\phi(x) - \phi(x')| \leq C|x - x'|(1 + |x|^s + |x'|^s), \text{ for an } s \geq 0.$$

and there exists constants $\alpha, \beta > 0$ *such that for* $u, v \in L^p(\Omega)$*,*

(i) $\left|\mathbb{E}\left[\phi(\Psi^\ell(u)) - \phi(\Psi(v))\right]\right| \lesssim N_\ell^{-\alpha}$, *provided that* $|\mathbb{E}[u - v]| \lesssim N_\ell^{-\alpha}$,

(ii) $\|\phi(\Psi^\ell(v)) - \phi(\Psi^{\ell-1}(v))\|_p \lesssim N_\ell^{-\beta/2}$, *for all* $p \geq 2$,

(iii) $\mathrm{Cost}\left(\Psi^\ell(v)\right) \lesssim N_\ell$.

Assume further (i) and (ii) hold for that all monomials ϕ *of degree* ≤ 2.

Theorem 2 (MLEnKF accuracy vs. cost)

Suppose Assumptions 1 and 2 hold. Then, for any $\epsilon > 0$ and $p \geq 2$, there exists an $L > 0$ and $\{M_\ell\}_{\ell=0}^{L}$ such that

$$\|\mu_n^{\mathrm{ML}}(\phi; (M_\ell)) - \mu_n^{\mathrm{MF}}(\phi)\|_p \lesssim \epsilon.$$

And

$$\mathrm{Cost}\,(\mathrm{MLEnKF}) \lesssim \begin{cases} \left(|\log(\epsilon)|^{1-n}\epsilon\right)^{-2}, & \text{if} \quad \beta > 1, \\ \left(|\log(\epsilon)|^{1-n}\epsilon\right)^{-2} |\log(\epsilon)|^3, & \text{if} \quad \beta = 1, \\ \left(|\log(\epsilon)|^{1-n}\epsilon\right)^{-\left(2+\frac{1-\beta}{\alpha}\right)}, & \text{if} \quad \beta < 1. \end{cases} \quad (2)$$

In comparison,

$$\|\mu_n^{\mathrm{EnKF}}(\phi) - \mu_n^{\mathrm{MF}}(\phi)\|_p \lesssim \epsilon,$$

is achieved at cost $\mathcal{O}\left(\epsilon^{-\left(2+\frac{1}{\alpha}\right)}\right)$.

Theorem 2 (MLEnKF accuracy vs. cost)

Suppose Assumptions 1 and 2 hold. Then, for any $\epsilon > 0$ and $p \geq 2$, there exists an $L > 0$ and $\{M_\ell\}_{\ell=0}^{L}$ such that

$$\|\mu_n^{\mathrm{ML}}(\phi; (M_\ell)) - \mu_n^{\mathrm{MF}}(\phi)\|_p \lesssim \epsilon.$$

And

$$\mathrm{Cost}\,(\mathrm{MLEnKF}) \lesssim \begin{cases} (|\log(\epsilon)|^{1-n}\epsilon)^{-2}, & \text{if} \quad \beta > 1, \\ (|\log(\epsilon)|^{1-n}\epsilon)^{-2} \left|\log(\epsilon)\right|^{3}, & \text{if} \quad \beta = 1, \\ (|\log(\epsilon)|^{1-n}\epsilon)^{-\left(2+\frac{1-\beta}{\alpha}\right)}, & \text{if} \quad \beta < 1. \end{cases} \tag{2}$$

In comparison,

$$\|\mu_n^{\mathrm{EnKF}}(\phi) - \mu_n^{\mathrm{MF}}(\phi)\|_p \lesssim \epsilon,$$

is achieved at cost $\mathcal{O}\left(\epsilon^{-\left(2+\frac{1}{\alpha}\right)}\right)$.

Central idea in the proof

Introduce

$$\mu_n^{\mathrm{MLMF}}(\phi) := \sum_{\ell=0}^{L} \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \phi(v_n^{\mathrm{MF},\ell}(\omega_{i,\ell})) - \phi(v_n^{\mathrm{MF},\ell-1}(\omega_{i,\ell}))$$

$$\mu_n^{\mathrm{MF,L}}(\phi) := \mathbb{E}\left[\phi(v_n^{\mathrm{MF,L}}) | Y_n^{\mathrm{obs}}\right],$$

and bound MLEnKF error by

$$\begin{aligned}
\|\mu_n^{\mathrm{ML}}(\phi) - \mu_n^{\mathrm{MF}}(\phi)\|_p &\leq \|\mu_n^{\mathrm{ML}}(\phi) - \mu_n^{\mathrm{MLMF}}(\phi)\|_p \\
&\quad + \|\mu_n^{\mathrm{MLMF}}(\phi) - \mu_n^{\mathrm{MF,L}}(\phi)\|_p + \|\mu_n^{\mathrm{MF,L}}(\phi) - \mu_n^{\mathrm{MF}}(\phi)\|_p \\
&\leq c \sum_{\ell=0}^{L} \left[\|v_n^\ell - v_n^{\mathrm{MF},\ell}\|_{\hat{p}} + \frac{\|v_n^{\mathrm{MF},\ell} - v_n^{\mathrm{MF},\ell-1}\|_{\hat{p}}}{M_\ell^{1/2}} \right] + \left| \mathbb{E}\left[\phi(v_n^{\mathrm{MF,L}}) - \phi(v_n^{\mathrm{MF}}) | Y_n^{\mathrm{obs}}\right] \right| \\
&\leq c \left(\epsilon + \sum_{\ell=0}^{L} M_\ell^{-1/2} N_\ell^{-\beta/2} + N_L^{-\alpha} \right)
\end{aligned}$$

Numerical example

Underlying dynamics Ornstein–Uhlenbeck SDE

$$du = -udt + 0.5dW(t),$$

with a set of observations

$$y_n = u_n + \gamma_n, \qquad i.i.d.\ \gamma_n \sim N(0, 0.04)$$

Solvers: Hierarchy of Milstein solution operators $\{\Psi_\ell\}_{\ell=0}^{L}$ with $\Delta t^\ell = \mathcal{O}\left(2^{-\ell}\right)$.

Compare the approximation errors for the observable $\phi(x) = x$ in terms of the RMSE

$$\sqrt{\sum_{n=1}^{100} |\mu_n^{\mathrm{ML}}(\phi) - \mu_n^{\mathrm{MF}}(\phi)|^2}.$$

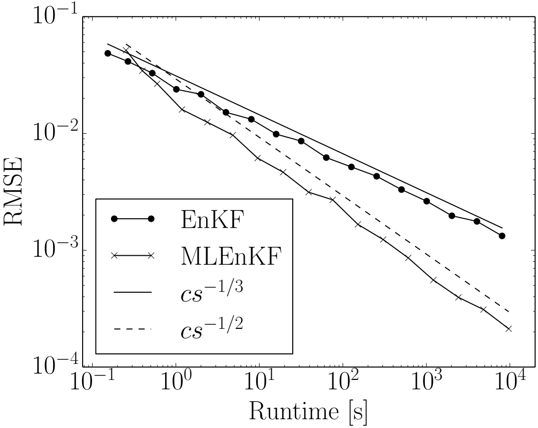
Numerical example

Underlying dynamics Ornstein–Uhlenbeck SDE

$$du = -udt + 0.5dW(t),$$

with a set of observations

$$y_n = u_n + \gamma_n, \qquad i.i.d.\ \gamma_n \sim N(0, 0.04)$$



Overview

Extension of MLEnKF to infinite dimensional state spaces

- Work in progress, together with Alexey Chernov, Law, Nobile and Tempone.
- Infinite dimensional stochastic dynamics:

$$u_{n+1} = \Psi(u_n)$$

where $u_n \in L^p(\Omega; \mathcal{H})$ with $\mathcal{H} = \mathrm{Span}(\{\nu_i\}_{i=1}^{\infty})$, and $\Psi : L^p(\Omega; \mathcal{H}) \to L^p(\Omega; \mathcal{H})$.

- And finite dimensional observations

$$y_n = H u_n + \gamma_n,$$

with linear $H : \mathcal{H} \to \mathbb{R}^m$

- Introduce nested hierarchy of Hilbert spaces

$$\mathcal{H}_0 \subset \mathcal{H}_1 \subset \ldots \subset \mathcal{H}_\infty = \mathcal{H},$$

where $\mathcal{H}_\ell = \mathrm{Span}(\{\nu_i\}_{i=1}^{N_\ell})$ and work with a hierarchy of solvers

$$\Psi^\ell : L^p(\Omega; \mathcal{H}_\ell) \to L^p(\Omega; \mathcal{H}_\ell).$$

Conclusion

- Extended EnKF to multilevel EnKF.
- Verified asymptotic efficiency gain for approximations of expectation of observables. We hope to improve result further!
- Further extension of MLEnKF to infinite dimensional state space is work in progress.
- Preprint available on Arxiv

Thank you!

1. Hoel, Håkon, Kody JH Law, and Raul Tempone, *Multilevel ensemble Kalman filtering*, arXiv preprint arXiv:1502.06069 (2015).
2. François Le Gland, Valérie Monbet, Vu-Duc Tran, et al., *Large sample asymptotics for the ensemble kalman filter*, The Oxford Handbook of Nonlinear Filtering, (2011), pp. 598–631.
3. Kody JH Law, Hamidou Tembine, and Raul Tempone, *Deterministic methods for nonlinear filtering, part i: Mean-field ensemble kalman filtering*, arXiv preprint arXiv:1409.0628, (2014).