Feature Data Assimilation in the unstable subspace

John Maclean[1], Erik Van Vleck[2]

[1]Department of Mathematics and RENCI, University of North Carolina at Chapel Hill
[2]Department of Mathematics, Kansas University

April 16, 2018

Overview

Feature DA

AUS

Particle Filtering

Projected PF

Numerical Example

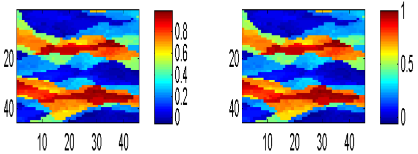
Future Directions

Feature Data Assimilation

"**A feature can be thought of as a low-dimensional representation of the data**, e.g., a principal component analysis (PCA) (Jolliffe, 2002), a Gaussian process model (Rasmussen and Williams, 2006), or a Gaussian mixture model (McLachlan and Peel, 2000).
Features are either constructed a priori, or learned from data. The same ideas carry over to data assimilation.
One can extract low-dimensional features from the data and then use the features to define a feature-based likelihood, which in turn defines a feature-based posterior distribution."

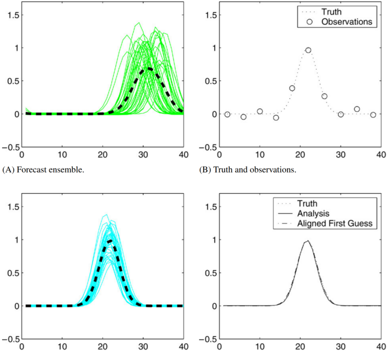
-*Morzfeld, Adams, Lunderman, Orozco, NPG 2018*

Feature DA - history



Context: history matching from geophysics. A kernel based EnKF is used to represent the image above in "feature space", and a minimization problem is solved to return an ensemble estimate in state space.
Sarma and Chen, SPE 119177, 2009

Feature DA - history



(A) Forecast ensemble. (B) Truth and observations.

(C) Aligned forecast ensemble. (D) Deterministic analysis – 1.

Context: field alignment. Rather than rely on DA to correct the magnitude of the state at different locations, one looks for a (regularized) warping of the underlying grid before carrying out the analysis step.

Ravela, Emanuel and Mclaughlin, Physica D 2006

Feature DA - recent history

M., Santitissadeekorn, Jones - feature DA using SMC-ABC

Weiss, Grooms - (?) feature DA assimilating observations only of vortex locations

Morzfeld, Adams, Lunderman, Orozco - feature DA using "perturbed observations"

Assimilation in the Unstable Subspace - AUS

Framework designed for the extended Kalman Filter, that observes only a reduced number of unstable directions.

Particle filters

The particle filter sequentially approximates the distribution of x at time n by a set $\{x_n^{(i)}, w_n^{(i)}\}$, $i = 1, \ldots, N$ of particles and weights.

The weight update for the i-th particle is

$$\begin{aligned} w_n^{(i)} &\propto w_{n-1}^{(i)}\, p(y_n^o | x_n^{(i)}) \\ &= w_{n-1}^{(i)} \exp\left(-\frac{1}{2} (y_n^o - \boldsymbol{H} x_n^{(i)})^T \mathbf{R}^{-1} (y_n^o - \boldsymbol{H} x_n^{(i)}) \right), \end{aligned}$$

where $\mathbf{R}$ is the covariance matrix for the observations.

Observe that the key quantity by which the particle filter gains information is the innovation $y_n^o - \boldsymbol{H} x_n^{(i)}$.

Goal

Initial goal to create a feature DA method where, instead of taking the full observations, one projects to a basis for the largest p local Lyapunov exponents.

That is, from another PoV, to create a "PF-AUS" method.

Identifying observation space as a subspace of model space

The state variables are $x \in \mathbb{R}^d$, the observations $y \in \mathbb{R}^m$, the observation operator $\boldsymbol{H}$ is linear.
Goal: identify the observations in a subspace of model space.
Assuming $\boldsymbol{H}$ is full rank one can identify the projection $P_{\boldsymbol{H}} = \boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{H}^T)^{-1}\boldsymbol{H}$, $P_{\boldsymbol{H}} : \mathbb{R}^d \to \mathbb{R}^d$.
Defining $\tilde{y} = \boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{H}^T)^{-1}y$ and $\tilde{\mathbf{R}} = \boldsymbol{H}^{-1}\mathbf{R}(\boldsymbol{H}^T)^{-1}$, one then has

$$(\tilde{y} - P_{\boldsymbol{H}}x)^T \, \tilde{\mathbf{R}}^{-1} \, (\tilde{y} - P_{\boldsymbol{H}}x) = (y - \boldsymbol{H}x)^T \, \mathbf{R}^{-1}(y - \boldsymbol{H}x) \, .$$

A particle filter in which the weight update step uses $\tilde{y}$ as the observations, $P_{\boldsymbol{H}}x$ as the state and $\tilde{\mathbf{R}}$ as the observation covariance matrix is identical to the 'standard' particle filter.

Identifying observation space as a subspace of model space

The state variables are $x \in \mathbb{R}^d$, the observations $y \in \mathbb{R}^m$, the observation operator $\boldsymbol{H}$ is linear.
Goal: identify the observations in a subspace of model space.
Assuming $\boldsymbol{H}$ is full rank one can identify the projection $P_{\boldsymbol{H}} = \boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{H}^T)^{-1}\boldsymbol{H}$, $P_{\boldsymbol{H}} : \mathbb{R}^d \to \mathbb{R}^d$.
Defining $\tilde{y} = \boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{H}^T)^{-1}y$ and $\tilde{\mathbf{R}} = \boldsymbol{H}^{-1}\mathbf{R}(\boldsymbol{H}^T)^{-1}$, one then has

$$(\tilde{y} - P_{\boldsymbol{H}}x)^T \, \tilde{\mathbf{R}}^{-1} \, (\tilde{y} - P_{\boldsymbol{H}}x) = (y - \boldsymbol{H}x)^T \, \mathbf{R}^{-1}(y - \boldsymbol{H}x) \, .$$

A particle filter in which the weight update step uses $\tilde{y}$ as the observations, $P_{\boldsymbol{H}}x$ as the state and $\tilde{\mathbf{R}}$ as the observation covariance matrix is identical to the 'standard' particle filter. This is a precursor for the conversion of the problem into "feature space".

Projected Particle Filtering - naive approach

Consider an orthogonal projection $P_n = Q_n Q_n^T$, where $P_n \in \mathbb{R}^{d\times d}$ and $Q_n \in \mathbb{R}^{d\times p}$. Construct an approximation of the particle filter update step, but employing approximations of the innovation and obs covariance in $\mathbb{R}^p$.
First attempt: define

$$\begin{aligned}\tilde{R}(x,y) =& \exp\left(\left[P_n\left(\tilde{y} - P_{\boldsymbol{H}}x\right)\right]^T \tilde{\mathbf{R}}^{-1}\left[P_n\left(\tilde{y} - P_{\boldsymbol{H}}x\right)\right]\right)\\ =& \exp\left(\left[Q_n^T\left(\tilde{y} - P_{\boldsymbol{H}}x\right)\right]^T Q_n^T \tilde{\mathbf{R}}^{-1} Q_n \left[Q_n^T\left(\tilde{y} - P_{\boldsymbol{H}}x\right)\right]\right)\ .\end{aligned}$$

Success: the state vector has been expressed as $Q_n^T\left(\tilde{y} - P_{\boldsymbol{H}}x\right) \in \mathbb{R}^p$ and the covariance matrix as $Q_n^T \tilde{\mathbf{R}}^{-1} Q_n \in \mathbb{R}^{p\times p}$.
This approach, with a suitable projection P_n, should reduce the effective dimension of the data assimilation problem.
But: While P_n is a projection, and $P_{\boldsymbol{H}}$ is a projection, $P_n P_{\boldsymbol{H}}$ is not a projection.

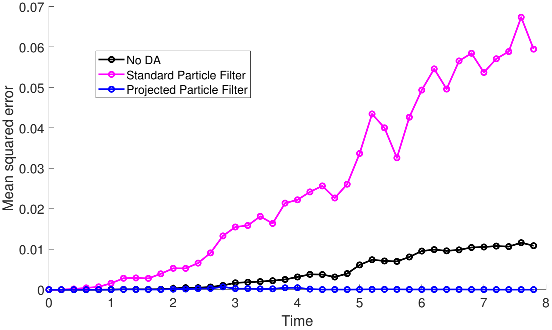
Projected Particle Filtering

Define $P_n^{\boldsymbol{H}}$ as the orthogonal projection onto the intersection of the subspaces spanned by the columns of Q_n and the rows of $\boldsymbol{H}$.
$P_n^{\boldsymbol{H}}$ may be approximated by e.g. POCS, Dykstra's projection algorithm. Then, the innovation has been successively replaced by

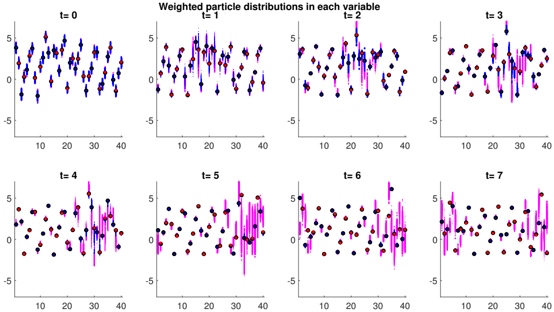
$$y - \boldsymbol{H}x \to \tilde{y} - P_{\boldsymbol{H}}x \to P_n P_n^{\boldsymbol{H}} \left(\tilde{y} - P_{\boldsymbol{H}}x\right)$$

Numerical Example

We take L96 with $F = 4$, 40 state variables.



Numerical Example



Future Work

State Space Model $u_{n+1} = F_n(u_n) + \xi_n$
Data Model $y_{n+1} = Hu_{n+1} + \eta_{n+1}$

State Space Model (perturbed form) $u_{n+1}^{(0)} + \delta_{n+1} = F_n(u_n^{(0)} + \delta_n) + \xi_n$
Data Model (perturbed form) $y_{n+1} = H(u_{n+1}^{(0)} + \delta_{n+1}) + \eta_{n+1}$

Projected State Space Model

$$u_{n+1}^{(0)} + P_{n+1}\delta_{n+1} = P_{n+1}[F_n(u_n^{(0)} + P_n\delta_n) + \xi_n]$$

Projected Data Model

$$P_n^H H^\dagger y_{n+1} = P_n^H(u_{n+1}^{(0)} + \delta_{n+1}) + P_n^H H^\dagger \eta_{n+1}$$

where $H^\dagger = H^T(HH^T)^{-1}$ and P_n^H is the orthogonal projection onto the intersection of the subspaces spanned by the columns of Q_n and the rows of H.
Note: $P_H = H^T(HH^T)^{-1}H = H^\dagger H$ (we are assuming H full rank).

Assimilation in both the $P-$ and $(I-P)-$spaces!