Introduction to data assimilation

Stan Swierczek

University of Arizona

swierczek@math.arizona.edu

August 17, 2018

overview

1. background
   - simple examples
   - why?
2. data assimilation
   - what is it?
   - set up
3. DA methods
   - EnKF
   - 4DVar
   - hybrid methods
4. experiment

example

Suppose

- I have two measurements y_1 and y_2 of some quantity x

and

- I know the error structure of the two measuring devices

example

If the errors for the first and second measurement devices are unbiased

$$E[e_1] = E[e_2] = 0$$

and the variance is known

$$Var(e_1) = \sigma_1^2 \qquad Var(e_2) = \sigma_2^2$$

then one* can show that the

$$BLUE[x] = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\, y_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\, y_2$$

*not me

another example

Suppose we have a time dependent model $x_k = F(x_{k-1})$, an initial condition x_{k-1}, and some observations at time k.

Given

$$\text{state estimate } \widetilde{x}_k = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}_k = F\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}_{k-1} \right) \text{ \& observations } y_k = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

is there a method to use the observations to correct the current state estimate?

Lorenz 96 model

In 1996, mathematician and meteorologist Edward Lorenz introduced a dynamical system described by

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F$$

If the dimension of the system is 40, and $F = 8$, this system exhibits chaotic behavior.

Lorenz 96 model

This system is sensitive to initial conditions.

watch movies

chaotic systems

This is a problem.

How can we try to model these systems and make predictions when they are this sensitive to perturbations?

data assimilation

Data Assimilation is a process where observations are incorporated into a numerical model.

To perform DA, you need:

- a numerical model that captures the dynamics of the physical system you are trying to approximate
- some observations of some of the state variables at some points in time
- an idea of the error in the observations you are getting

DA

One step of a DA problem:

- Start with an estimate of the current state
- Push the estimate forward one time step through the model
- Collect measurements/observations
- Change your new estimate to be *between* the model estimate and any observations (if you received any)

DA

Comments:

- Much of DA framework relies on the assumption that prior and/or posterior distributions of state variables are Gaussian
- An important element is estimating and updating covariances between state variables
- Usually combined with Monte Carlo methods

problem

This is the general framework for a DA problem:

$x_k = M(x_{k-1})$	model
$y_k = Hx_k + \eta_k$	observations
$\eta_k \sim \mathcal{N}(0, R)$	measurement noise

problem

An example observation operator H:

$$y_k = Hx_k + \eta_k = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}_k = \begin{bmatrix} x_2 \\ x_5 \end{bmatrix} + \begin{bmatrix} \eta_1 \text{ (noise)} \\ \eta_2 \text{ (noise)} \end{bmatrix}$$

Ensemble Kalman Filter

One step of a SqEnKF algorithm for solving this problem:

given

y_k	observations at time k
$\{\mathbf{x_{k-1}}\}$	ensemble of size Ne at time $k-1$

forecast

$\mathbf{x_f} = M(\mathbf{x_{k-1}})$	forecast ensemble
$\mu_f = \frac{1}{Ne}\sum_{j=1}^{Ne} x_f^j$	forecast mean
$X = \frac{1}{\sqrt{Ne-1}}[\mathbf{x_f} - \mu_f]$	forecast perturbations
$P_f = XX^T$	forecast covariance

EnKF

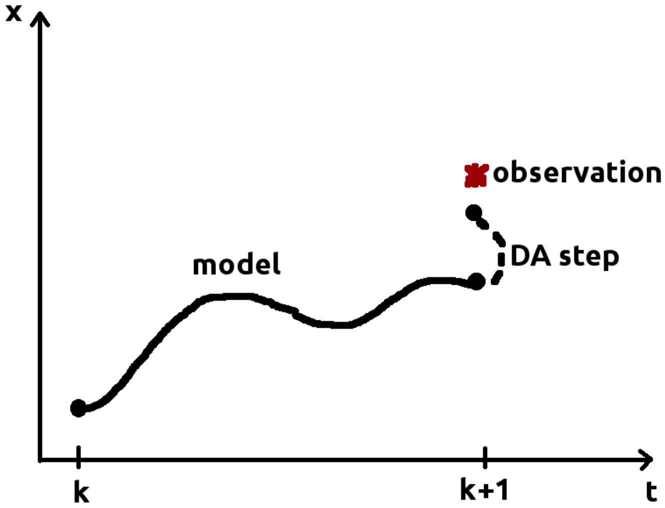
One step of a SqEnKF algorithm for solving this problem:

analysis

$K = P_f H^T (H P_f H^T + R)^{-1}$	Kalman Gain
$\mu_a = \mu_f + K(y_k - H\mu_f)$	analysis mean
$ZZ^T = I - X^T H^T (H X X^T H^T + R)^{-1} H X$	compute square root
$P_a = XZ(XZ)^T$	analysis covariance
$\widetilde{X}_a = XZ$	analysis perturbations
$\mathbf{x_a} = \mu_a + \sqrt{Ne - 1}\, \widetilde{X}_a$	analysis ensemble

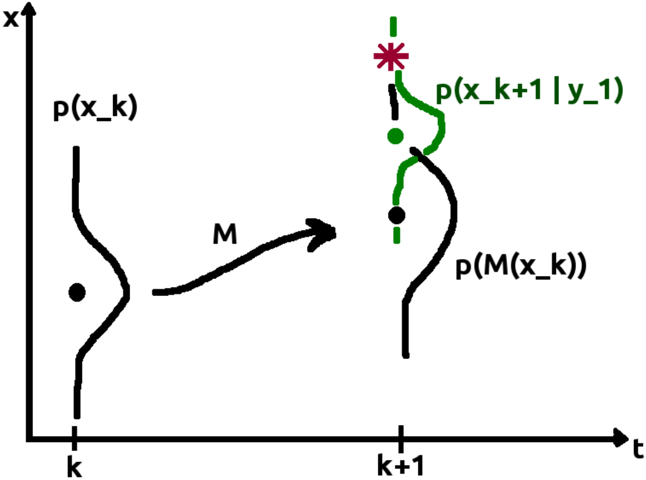
EnKF

Crude schematic of an EnKF algorithm:



EnKF

Why ensemble methods?



4DVar

Idea for 4DVar:

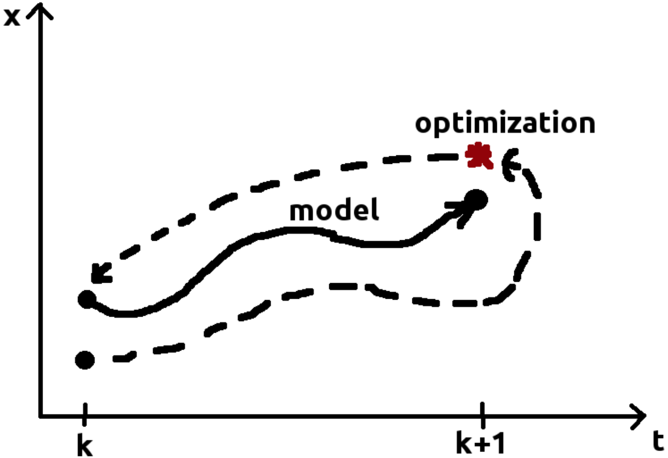
Find x_0 that maximizes $p(x_0|y_1)$

$$
\begin{aligned}
\Rightarrow \text{Minimize } F(x_0) &= -\log p(x_0|y_1) \\
p(x_0|y_1) &\propto p(x_0)p(y_1|x_0) \\
\Rightarrow F(x_0) &= \frac{1}{2}(x_0-\mu)^T B^{-1}(x_0-\mu) \\
&\quad + \frac{1}{2}(HM(x_0)-y)^T R^{-1}(HM(x_0)-y)
\end{aligned}
$$

Use a Gauss-Newton nonlinear least squares solver to minimize $F(x_0)$. This requires supplying the derivative (Jacobian) of your model.

4DVar

Crude schematic of a 4DVar algorithm:



hybrid methods

In practice, 4DVar performs poorly in sequential data assimilation because of limitations in updating state covariances.

Luckily, there exist hybrid methods that combine the strength of 4DVar (better performance in nonlinear models) with EnKF (better estimate of spread and updating covariances).

- En4DVar: Same as EnKF, but 4DVar is performed on ensemble mean
- EDA: 4DVar is performed on an ensemble of perturbed states

experiment

Let's try to estimate the Lorenz 96 system described earlier.

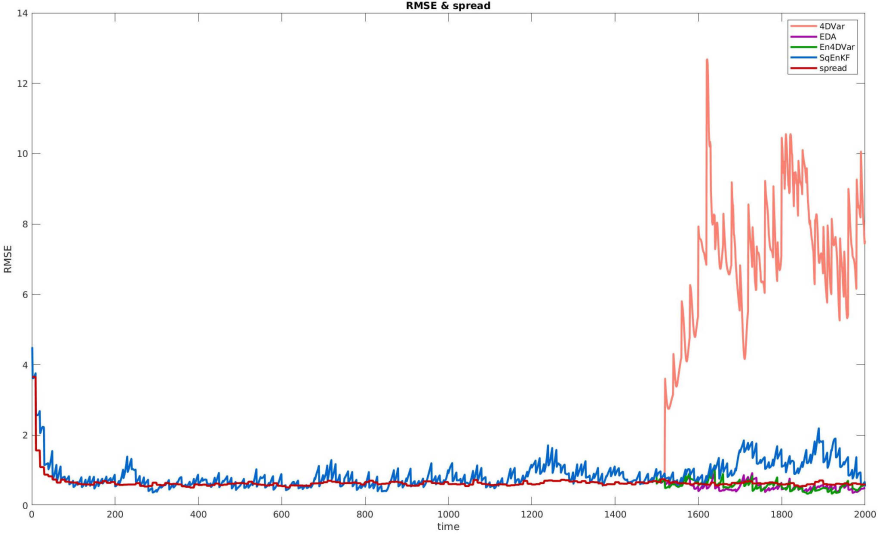
We will observe every other variable (i.e., 20 out of the 40), with $\mathcal{N}(0, 1)$ noise added to each, every tenth time step.

Our initial state will be the mean of a number of randomly drawn states from a long simulation.

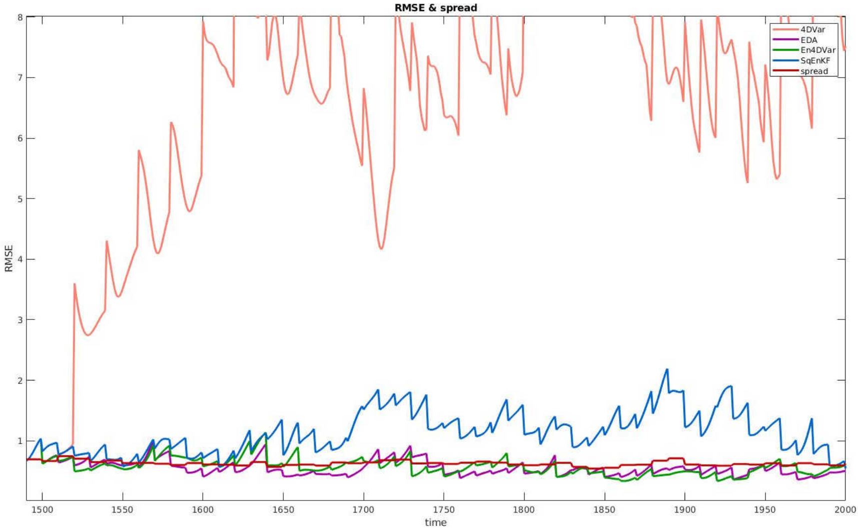
Results

watch movies

Results


RMSE & spread
RMSE
time
0
2
4
6
8
10
12
14
0
200
400
600
800
1000
1200
1400
1600
1800
2000
4DVar
EDA
En4DVar
SqEnKF
spread

Results


RMSE & spread
RMSE
time
4DVar
EDA
En4DVar
SqEnKF
spread
1500
1550
1600
1650
1700
1750
1800
1850
1900
1950
2000
1
2
3
4
5
6
7
8

Thank You

References

Asch, M., Bocquet, M., & Nodet, M. (2016). *Data Assimilation: Methods, Algorithms, and Applications.* Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics.

Hunt, B., Kostelich, E., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D: Nonlinear Phenomena*, **230**(1-2), 112-126.

Law, K., Stuart, A., & Zygalakis, K. (2015). *Data Assimilation: A Mathematical Introduction.* Cham, Switzerland: Springer International Publishing.

Reich, S., & Cotter, C. (2015) *Probabilistic Forecasting and Bayesian Data Assimilation.* Cambridge: Cambridge University Press.

Tippett, M., Anderson, J.,Bishop, C., Hamill, T., & Whitaker, J., (2003). Ensemble square root filters. *Monthly Weather Review*, **131**, 1485–1490.

Program in Applied Mathematics

University of Arizona

swierczek@math.arizona.edu