An adaptive localization method

Yicun (Ethan) Zhen
Fuqing Zhang

Dec 18, 2013
Outline

1. Introduction
2. A cost function approach
3. Cases that the true covariance is unknown
4. Test results using Lorenz-96 system
5. Summary
The sampling error in the ensemble Kalman filters

**Sampling error**: When the state dimension is far larger than the ensemble size, the sample covariance is not accurate!
A popular treatment of the sampling error

Covariance localization

Let $B^s$ be the sample covariance, $\rho(x, y)$ be the localization function. For example, $\rho(d) = e^{-\lambda^2 d^2}$, or the Gaspari-Cohn function. Construct the localized covariance $B_{Loc}$, such that for each pair of state variables $x_i, x_j$,

$$(B_{Loc})_{i,j} = (B^s)_{i,j} \rho(d(x_i, x_j)).$$
Covariance localization

$n=120, m=30, N=31$

- True covariance
- Sample covariance
- Localized sample covariance
Questions:

- How to choose localization function? Mostly solved by Gaspari and Cohn 1999;
- How to choose the localization parameter (ROI)?

The plot of Gaspari–Cohn 4th degree rational polynomial

Covariance localization
Some notations for convenience

For a single observation $y^o \in \mathbb{R}$, and a single state variable $x_i$ which is the i-th component of the whole state vector, the update of the model forecast can be written as:

$$\Delta x_i = \frac{\text{cov}(x_i, \Delta y)}{\text{var}(\Delta y)} \Delta y = r_i \Delta y$$

- If the covariance is the true covariance, $r_i$ will be denoted by $r^t_i$;
- If the covariance is the sample covariance, $r_i$ will be denoted by $r^s_i$;
- If the covariance is the localized sample covariance, $r_i = \rho_i r^s_i$;
A cost function approach

\[ \Delta x_i = r_i^t \Delta y \quad \text{using the true covariance} \]

\[ \Delta x_i = \rho_i r_i^s \Delta y \quad \text{using the localized sample covariance} \]

This cost function directly compares the difference between the true regression coefficient with the localized sample regression coefficient.
An alternative cost function

\[ F_{10}(ROI) = \sum_{d_i \leq ROI} \{(r^t_i - \rho_i r^s_i)^2 - (r^t_i)^2\} \]
How to determine ROI using cost functions

Sketch of the algorithm:

- Compute the value of $F_{10}$ or $F_0$ for the ROI in a wide range: $1 \leq ROI \leq ROI_{max}$;
- Choose the ROI that minimizes $F_{10}$ or $F_0$.

$F_{10}$ reaches its minimum at the same place as the cost function $F_0$. 
Test 1

- Region: $n = 120$ grid points on a unit circle;
- ensemble size: $N = 11, 21, 31, 61, 121$;
- true radius of influence: $ROI = 1, 5, 10$ grid points;
- pick the true covariance matrix of the form

$$\text{cov}(x_i, x_j) = \left(\frac{1}{2}\right)^2 \left(\frac{2n \times d(x_i, x_j)}{ROI}\right)^2$$
Test 1

- draw random sample vectors from the distribution $\mathcal{N}(0, \text{cov}(x_i, x_j))$;
- find the value of $ROI$, denoted by $ROI^{(0)}$ that minimizes the cost function $F_{10}$.

But the value $ROI^{(0)}$ may depend on the sample in each experiment. Hence we do this experiment for 1000 times and find the maximum likelihood estimate of $ROI^{(0)}$. 
Test 1

ROI

ROI = 1

ROI = 5

ROI = 10

ensemble size

ROI^{(0)}
If the true covariance is unknown

We find a probability distribution of the true covariance matrix \( p(B|B^s) \) that is based on the sample. Then we find the value of \( ROI \), denoted by \( ROI^{(1)} \), that minimizes the average of the cost function \( F_{10} \) over the distribution \( p(B|B^s) \).
A cost function

\[ F(ROI) = \int \sum_{d_i < ROI} [(\rho_i r_i^s - r_i^t)^2 - (r_i^t)^2] \frac{p(B_{Loc}^s | B_{Loc})p(B_{Loc})}{p(B_{Loc}^s)} dB_{Loc} \]

- For technical reasons we only consider \( ROI \) that is not larger than \( ROI_{\text{max}} = \frac{N-3}{2} \);
- \( B_{Loc} \) is the true covariance matrix of the variables that are within distance \( ROI \) from the observation;
- \( p(B_{Loc}^s | B_{Loc}) \) is the Wishart distribution since we assume all variables are Gaussian;
- \( p(B_{Loc}) \) is an uninformative prior which we choose to be the Jeffreys prior.
A cost function

However this cost function is monotonically decreasing in almost all cases.
An alternative cost function

\[ F(ROI) = \int \left\{ \sum_{d_i < ROI} \left[ (\rho_i r_i^s - r_i^t)^2 - (r_i^t)^2 \right] + \sum_{d_i < ROI} \rho_i^2 (r_i^s - r_i^t)^2 \right\} \]

\[ \frac{p(B_{Loc}^s | B_{Loc}) p(B_{Loc})}{p(B_{Loc}^s)} dB_{Loc} \]
An alternative cost function
Test 2

- Region: $n = 120$ grid points on a unit circle;
- ensemble size: $N = 11, 21, 31, 61, 121$;
- true radius of influence: $ROI = 1, 5, 10$ grid points;
- pick the true covariance matrix of the form
  \[ \text{cov}(x_i, x_j) = \left(\frac{1}{2}\right)(\frac{2n \times d(x_i, x_j)}{ROI})^2 \]
- draw random sample vectors from the distribution $\mathcal{N}(0, \text{cov}(x_i, x_j))$;
- find the value of $ROI$, denoted by $ROI^{(1)}$ that minimizes the cost function $F$.

We do 1000 tests to find the maximum likelihood of $ROI^{(1)}$. 
Test 2

ROI vs ensemble size for different ROI values:

- Red: $ROI^{(0)}$
- Black: $ROI^{(1)}$
- $ROI_{\text{max}}$

Values for ROI:
- ROI=1
- ROI=5
- ROI=10

Graph shows the relationship between ROI and ensemble size for each ROI value.
Test results using Lorenz-96 system

\[
\frac{dX_k}{dt} = -X_{k-2}X_{k-1} + X_{k-1}X_{k+1} - X_k + F
\]

Set up
- Region: \( n = 120 \) uniformly distributed grid points on a circle;
- observations: \( m = 30 \); uniformly distributed; error variance \( R = 0.04 \); appears every two time steps;
- ensemble size: \( N = 61 \);
- Runge-Kutta 4th order method, \( dt = 0.05 \), \( F = 8 \);
- relaxation coefficient \( \alpha = 0.5 \);
- sequential ensemble square foot filter.
Test results using Lorenz-96 system

$n=120, m=30, N=61, \alpha=0.5$

(a) RMSE

(b) RMSE

(c) ROI(1)
Summary

- In the case of known true covariance, a cost function is designed to find the ROI that minimizes the RMSE of the Kalman update; the larger the true ROI, or the larger the ensemble size, the larger the ROI is obtained by the cost function approach;

- In the case of unknown, we use a probabilistic approach to define the cost function. Similarly, with larger ensemble size or larger true ROI, the larger ROI\(^{(1)}\) is obtained;

- we need to add a penalty term to make it work, but introduce the arbitrariness. We will try to find better prior, or better penalty term;

- preliminary application to Lorenz-96 model shows promising results, though not necessarily the absolute optimum ROI.
Test results using Lorenz-96 system

n=120, m=60, N=61, $\alpha=0.5$

(a) RMSE

(b) RMSE

(c) ROI

Adaptive ROI

ROI=30
ROI=24
ROI=16
ROI=8

RMSE

10 20 30 40 50

1
2
3

1000 1050 1100 1150 1200

0.08 0.1 0.12 0.14 0.16 0.18

1000 2000 3000 4000 5000

10 15 20

1000 2000 3000 4000 5000

20 15 10
Test results using Lorenz-96 system

\[ n=120, \quad m=120, \quad N=61, \quad \alpha=0.5 \]

(a)

RMSE

ROI=30
ROI=24
ROI=16
ROI=8
Adaptive ROI

(b)

time

(c)

ROI

RMSE

0.06
0.08
0.1

1000 1050 1100 1150 1200

1200

1000 1050 1100 1150 1200

12

14

16

18

12

14

16

18