Stability of the EnKF under nested covariance estimators

Marie Turčičová [1] Jan Mandel [2] Kryštof Eben [3]

[1] Charles University, Faculty of Mathematics and Physics (Prague)

[2] University of Colorado Denver

[3] Czech Academy of Sciences, Institute of Computer Science

20th EYSM, 17th August, Uppsala

State-space model

- $\boldsymbol{X}^t$ true (hidden) system state
- $\boldsymbol{X}^f$ first guess
- $\boldsymbol{y}$ observations

Model

$$\boldsymbol{X}^f = \boldsymbol{X}^t + \boldsymbol{e}_X$$
$$\boldsymbol{y} = \boldsymbol{X}^t + \boldsymbol{e}_y$$

$\boldsymbol{e}_X$ and $\boldsymbol{e}_y$ are independent with zero mean and covariances

$$\mathrm{cov}(\boldsymbol{e}_X) = C$$
$$\mathrm{cov}(\boldsymbol{e}_y) = R,$$

i.e. $\mathrm{cov}(\boldsymbol{X}^f) = C, \mathrm{cov}(\boldsymbol{y}) = R.$

Kalman filter

- $\boldsymbol{X}_j^f$ first guess at time $j = 0, 1, 2, \ldots$
- $C_j^f = \text{cov}(\boldsymbol{X}_j^f)$
- $\boldsymbol{y}_j$ observations at time j
- assume $\dim \boldsymbol{y}_j = \dim \boldsymbol{X}_j^f$

Analysis step:

$$\begin{aligned} \boldsymbol{X}_j^a &= \boldsymbol{X}_j^f + C_j^f (C_j^f + R)^{-1} (\boldsymbol{y}_j - \boldsymbol{X}_j^f) \\ C_j^a &= (I - C_j^f (C_j^f + R)^{-1}) C_j^f \end{aligned}$$

Forecast step:

$$\boldsymbol{X}_{j+1}^f = \psi(\boldsymbol{X}_j^a) \quad \text{(dynamics model)}$$

Dimension problem

In numerical weather prediction models, the state is a random field of dimension cca 10^7.

Kalman filter update cannot be applied directly.

Solution: Ensemble Kalman filter (EnKF)[1]

- the system state is represented by an ensemble $\boldsymbol{X}_j^{(1)}, \boldsymbol{X}_j^{(2)}, \ldots, \boldsymbol{X}_j^{(N)} \qquad (N \propto 10, 100)$
- each member is updated by Kalman filter equations
- C_j^f is estimated through the ensemble

[1] Evensen (1994)

Estimating covariance for EnKF

Sample covariance

$$\hat{C}_j^f = \frac{1}{N-1}\sum_{n=1}^{N}(\boldsymbol{X}_j^{(n)f} - \bar{\boldsymbol{X}}_j^f)(\boldsymbol{X}_j^{(n)f} - \bar{\boldsymbol{X}}_j^f)^\top$$

- low rank
- spurious correlations

Improved estimators

- shrinkage estimators
- sparse approximations

In what follows we will focus on one particular time j.

Nested models for sparse approximation

- C^f is a $p \times p$ matrix
- we approximate C^f by a small number of nonzero elements
- these elements can be governed by a model with parameter $\boldsymbol{\theta}$

$$C^f(\boldsymbol{\theta}), \;\; \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^m, m << p^2$$

- consider nested parameter spaces

$$\Theta_k \subset \Theta_{k+1}, \quad \dim(\Theta_k) \leq \dim(\Theta_{k+1}), \quad k \in \mathbb{N}$$

such that the true value $\boldsymbol{\theta}_k^0$ is an interior point of $\Theta_k, \forall k \in \mathbb{N}$

$\rightarrow$ nested covariance models $C^f(\boldsymbol{\theta}_k)$

Estimating parameters for Gaussian fields

- assume $\boldsymbol{X}^f$ is Gaussian
- likelihood can be formed (for an ensemble of size N)
- parameters θ_k can be estimated by the maximum likelihood method $\rightarrow$ MLE $\hat{\theta}_k$

Using the delta method we obtain[2]

$$\sqrt{N}\big(C^f(\hat{\theta}_k) - C^f(\theta_k^0)\big) \xrightarrow{\text{d}} \mathcal{N}\left(\mathbf{0}, Q_{C^f(\theta_k^0)}\right),$$

where $Q_{C^f(\theta_k^0)}$ represents the asymptotic variance of $C^f(\hat{\theta}_k)$ and can be understood as an inverse of the Fisher information matrix in a generalized sense.

[2] Turčičová et al. (2017)

Comparison of asymptotic variances

More importantly, it holds[3]

$$Q_{C^f(\theta_k^0)} \leq Q_{C^f(\theta_{k+1}^0)}, \ \ \forall k \in \mathbb{N}$$

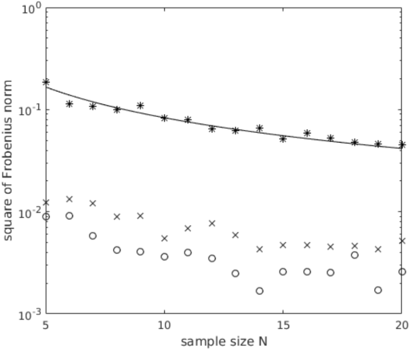
$\rightarrow$ estimator in a smaller parametric subspace (containing the true parameter θ_k^0) is asymptotically more (or equally) precise

---

[3] Turčičová et al. (2017)

Simulated distance from true covariance

Suppose C^f can be diagonalized by a known transformation (e.g. FFT) $\rightarrow$ matrix D. We simulate distance from the true covariance (model $D^{(2)}$), sampling from the true model, from two less parsimonious models and different sample sizes.



$$* \quad D^{(p)} = \text{diag}\left(\{d_i\}_{i=1}^{p}\right)$$

$$\times \quad D^{(3)} = \text{diag}\left(\left\{(c_1 - c_2\lambda_i)^{-1}(-\lambda_i)^{-\alpha}\right\}_{i=1}^{p}\right)$$

$$\circ \quad D^{(2)} = \text{diag}\left(\left\{c(-\lambda_i)^{-\alpha}\right\}_{i=1}^{p}\right)$$

Back to the Ensemble Kalman filter (EnKF)

- $\boldsymbol{X}_j^{(1)}, \boldsymbol{X}_j^{(2)}, \ldots, \boldsymbol{X}_j^{(N)}$ ensemble of system states
- $\hat{C}_j^f = \frac{1}{N-1} \sum_{n=1}^{N} (\boldsymbol{X}_j^{(n)f} - \bar{\boldsymbol{X}}_j)(\boldsymbol{X}_j^{(n)f} - \bar{\boldsymbol{X}}_j)^\top$ sample covariance
- $\boldsymbol{y}_j^{(n)} = \boldsymbol{X}_j^t + \xi_j + \xi_j^{(n)}$, where $\xi_j, \xi_j^{(n)} \sim N(\mathbf{0}, R)$, $n = 1, \ldots, N$ perturbed observations

Analysis step:

$$\boldsymbol{X}_j^{(n)a} = \boldsymbol{X}_j^{(n)f} + \hat{C}_j^f (\hat{C}_j^f + R)^{-1} \left(\boldsymbol{y}_j^{(n)} - \boldsymbol{X}_j^{(n)f} \right)$$

$C_j^{(n)a}$can be estimated from the ensemble

Forecast step:

$$\boldsymbol{X}_{j+1}^{(n)f} = \psi \left(\boldsymbol{X}_j^{(n)a} \right) \quad \text{(dynamics model)}$$

Behaviour of EnKF under covariances models

Long time behaviour of the EnKF

- under different covariance estimators
- with a fixed ensemble size.

Objective:

- stability (get an upper bound for the filter error)

Behaviour of EnKF under sample covariance

Kelly et al. (2015)

Assume

- $R = \gamma^2 I$
- model ψ is a representation of Navier-Stokes equations
- fixed ensemble size N
- filter error $\boldsymbol{e}_j^{(n)} = \boldsymbol{X}_j^{(n)a} - \boldsymbol{X}_j^t$, $j \geq 0$

Then $\exists \beta \in \mathbb{R}$ such that $\forall h > 0$ it holds

$$\mathrm{E}\,|\boldsymbol{e}_j^{(n)}|^2 \leq e^{2\beta hj}\,\mathrm{E}\,|\boldsymbol{e}_0^{(n)}|^2 + 2N\gamma^2 \frac{e^{2\beta hj} - 1}{e^{2\beta h} - 1}$$

for any $j \geq 1$.

Behaviour of EnKF under nested cov. estimators

Can the upper bound for the filter error be improved by using a covariance estimator that comes from a smaller parameter subspace?

Simulations suggest a positive answer.

Simulation settings:

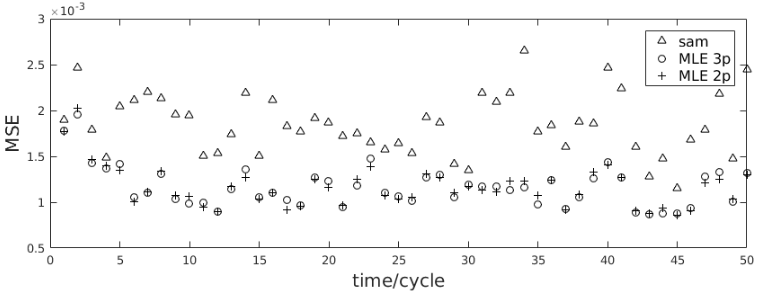
- ensemble size $N = 5$
- state dimension $p = 100$
- true covariance $C^t = FDF^\top$, where F is the discrete Fourier transform and $D = \text{diag}\{c(-\lambda_i)^{-\alpha}, i = 1, \ldots, p\}$ with λ_i's being eigenvalues of two-dimensional Laplace operator
- dynamics model: $\psi\left(\boldsymbol{X}_j^{(n)a}\right) = A\boldsymbol{X}_j^{(n)a} + \boldsymbol{b}$, where $A = a \cdot I, \boldsymbol{b} \sim \mathcal{N}(\mathbf{0}, C^t), n = 1, \ldots, N$

Time series of MSE of EnKF, nested cov. estimators

Nested covariance models:

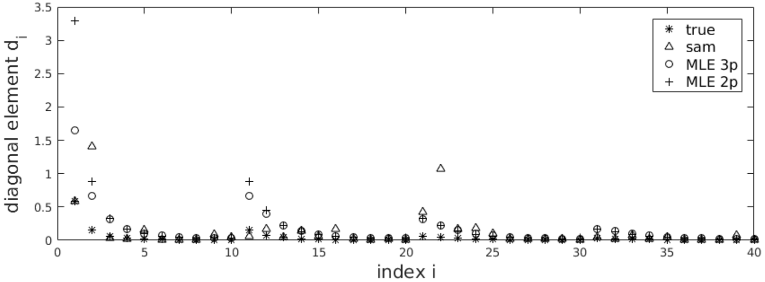
$$D^{(p)} = \mathrm{diag}\{d_1, \ldots, d_p\}$$
$$D^{(3)} = \mathrm{diag}\{(c_1 - c_2\lambda_i)^{-1}(-\lambda_i)^{-\alpha}, i = 1, \ldots, p\}$$
$$D^{(2)} = \mathrm{diag}\{c(-\lambda_i)^{-\alpha}, i = 1, \ldots, p\}$$

Mean square errors of the analysis ensemble mean:



Behaviour of EnKF under nested cov. estimators

Spectral representations of the true filtering covariance C^a and its different estimates $\hat{C}^a$ (the first 40 elements) after 50 cycles:



Conclusion

- covariance estimator that comes from a smaller (but proper) subspace does not have larger asymptotic variance
- simulation results suggest that in such a case, keeping the covariance model parsimonious makes EnKF perform significantly better
- derive the theoretical form of the upper bound for the filter error

Thank you for your attention!

turcic@karlin.mff.cuni.cz

This work was supported by the grant SVV 2017 No. 260454 and by the U.S. National Science Foundation under grant DMS-1216481.

References

G. Evensen.
Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics.
Journal of Geophysical Research, 99 (C5)(10):143–162, 1994.

M. Turčičová, J. Mandel, and K. Eben.
Multilevel maximum likelihood estimation with application to covariance matrices.
Submitted in January 2017.

D. T. B. Kelly, K. J. H. Law, and A. M. Stuart.
Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time.
Nonlinearity, 27(10):2579–2603, 2014.