Ensemble Kalman filtering with residual nudging: some recent progress

Xiaodong Luo[1] and Ibrahim Hoteit[2]

[1]International Research Institute of Stavanger (IRIS), Bergen, Norway
[2]King Adbullah University of Science and Technology, Thuwal, Saudi Arabia

Summary

The ensemble Kalman filter (EnKF) is an efficient algorithm for many data assimilation problems. In certain circumstances, however, divergence of the EnKF might be spotted. In previous studies [4, 5, 6], the authors proposed an observation-space based strategy, called residual nudging, in order to improve the stability of the EnKF when dealing with linear observation operators. The main idea behind residual nudging is to monitor, and if necessary, adjust the distances (misfits) between the real observations and the simulated ones of the state estimates, in the hope that by doing so one may obtain better estimation accuracy.

In a more recent study [3], residual nudging is extended and modified in order to handle nonlinear observation operators. Such extension and modification result in an iterative filtering framework that is able to achieve the objective of residual nudging in the presence of nonlinear observation operators, under suitable conditions. The 40 dimensional Lorenz 96 model is used to illustrate the performance of the iterative filter. Numerical results show that, while a normal EnKF may diverge in the presence of nonlinear observation operators, the proposed iterative filter performs very stably and leads to reasonable estimation accuracies under various experiment settings.

Notations

- $\mathbf{X}_k^a = \{\mathbf{x}_{k,i}^a\}_{i=1}^N$: analysis ensemble at time k, with N members
- $\hat{\mathbf{x}}_k^a$: analysis (sample) mean of $\mathbf{X}_k^a$
- $\mathbf{x}_k^{tr}$: true system state (truth)
- Observation system: $\mathbf{y}_k = \mathbf{H}_k(\mathbf{x}_k) + \mathbf{v}_k$, with $\mathbb{E}(\mathbf{v}_k) = \mathbf{0}$ and $\mathbf{cov}(\mathbf{v}_k) = \mathbf{R}_k$
- $\dim(\mathbf{x}_k) = n$ and $\dim(\mathbf{y}_k) = p$
- Residual $\mathbf{r}_k^a \equiv \mathbf{H}_k(\hat{\mathbf{x}}_k^a) - \mathbf{y}_k = [\mathbf{H}_k(\hat{\mathbf{x}}_k^a) - \mathbf{H}_k(\mathbf{x}_k^{tr})] - \mathbf{v}_k$
- Residual norm $\|\mathbf{r}_k^a\| \equiv \sqrt{(\mathbf{r}_k^a)^T \mathbf{R}_k^{-1} \mathbf{r}_k^a}$

Residual nudging with linear observation operators

The method developed in [5] is the focus of our introduction below, since it will be extended to the cases with nonlinear observation operators. Readers are referred to [4, 6] for more details of other implementations of residual nudging.

Objective

Make the residual norm of the state estimate be bounded in a specified interval $[\beta_l \sqrt{p}, \beta_u \sqrt{p}]$, with $0 \le \beta_l < \beta_u$.

Implementation

- For mean update, a family of the formula is considered

$$\hat{\mathbf{x}}_k^a = \hat{\mathbf{x}}_k^b + \mathbf{G}_k \left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^b \right) , \tag{1}$$

$$\mathbf{G}_k = \hat{\mathbf{C}}_k^b \mathbf{H}_k^T \left(\delta_k \, \mathbf{H}_k \hat{\mathbf{C}}_k^b \, \mathbf{H}_k^T + \gamma_k \, \mathbf{R}_k \right)^{-1} , \tag{2}$$

  where $\hat{\mathbf{C}}_k^b$ in general might be related, but not necessarily identical, to the background sample error covariance $\hat{\mathbf{P}}_k^b$ (e.g., $\hat{\mathbf{C}}_k^b$ can be a hybrid of $\hat{\mathbf{P}}_k^b$ and another full rank matrix); δ_k and γ_k are some positive scalars. Here we consider the case $\delta_k = 1$, so that $\mathbf{G}_k$ in Eq. (2) resembles the Kalman gain in the EnKF, with an inflation factor $1/\gamma_k$. The more general situation with $\delta_k \neq 1$ is considered in [5]
- Some sufficient conditions: with $\delta_k = 1$, it can be shown that the residual norm w.r.t $\hat{\mathbf{x}}_k^a$ is given by

$$\|\hat{\mathbf{r}}_k^a\| = \|\Phi_k \, (\mathbf{R}_k^{-1/2} \hat{\mathbf{r}}_k^b)\|_2 , \tag{3}$$

$$\Phi_k = \gamma_k \, (\mathbf{A}_k + \gamma_k \mathbf{I})^{-1} , \tag{4}$$

$$\mathbf{A}_k = \mathbf{R}_k^{-1/2} \mathbf{H}_k \hat{\mathbf{C}}_k^b \mathbf{H}_k^T \mathbf{R}_k^{-T/2} . \tag{5}$$

  Let $\xi_\bullet^k = \beta_\bullet \sqrt{p} / \|\hat{\mathbf{r}}_k^b\| (\bullet = u \text{ or } l)$, λ_{max}^k and λ_{min}^k be the maximum and minimum eigenvalues of $\mathbf{A}_k$, respectively, then a sufficient condition for $\|\hat{\mathbf{r}}_k^a\| \in [\beta_l \sqrt{p}, \beta_u \sqrt{p}]$ is [5]

$$\frac{\xi_l^k}{1 - \xi_l^k} \lambda_{max}^k \le \gamma^k \le \frac{\xi_u^k}{1 - \xi_u^k} \lambda_{min}^k \tag{6}$$

- The covariance update formula is the same as the ensemble square root filter (e.g., the ensemble transform Kalman filter, ETKF [1])

Residual nudging with nonlinear observation operators

Note that the analysis $\hat{\mathbf{x}}^a$ in Eq. (1) (with $\delta_k = 1$) is actually the solution of the following least-squares problem

$$\underset{\mathbf{x}}{\text{argmin}} \, \|\mathbf{y}^o - \mathbf{H}\mathbf{x}\|_{\mathbf{R}}^2 + \gamma \|\mathbf{x} - \hat{\mathbf{x}}^b\|_{\hat{\mathbf{C}}^b}^2 . \tag{7}$$

Inspired by this point of view and the iterative regularization methods in inverse problems theory (see, for example, [2]), in cases of nonlinear observation operators, we solve a sequence of (local) minimization problems, in terms of

$$\underset{\mathbf{x}^{i+1}}{\text{argmin}} \|\mathbf{y}^o - \mathcal{H}(\hat{\mathbf{x}}^i) - \mathbf{J}^i(\mathbf{x}^{i+1} - \hat{\mathbf{x}}^i)\|_{\mathbf{R}}^2 + \gamma^i \|\mathbf{x}^{i+1} - \hat{\mathbf{x}}^i\|_{\hat{\mathbf{C}}^b}^2 , \tag{8}$$

where $\mathcal{H}$ is the nonlinear observation operator and $\mathbf{J}^i$ is the Jacobian matrix (approximated by SPSA [7] here) of $\mathcal{H}$ at $\hat{\mathbf{x}}^i$ ($\hat{\mathbf{x}}^0 = \hat{\mathbf{x}}^b$), such that $\mathcal{H}(\hat{\mathbf{x}}^i) + \mathbf{J}^i(\mathbf{x}^{i+1} - \hat{\mathbf{x}}^i)$ is the first order Taylor approximation of $\mathcal{H}(\mathbf{x}^{i+1})$. The solution $\hat{\mathbf{x}}^{i+1}$ of the minimization problem (8) is given by

$$\hat{\mathbf{x}}^{i+1} = \hat{\mathbf{x}}^i + \mathbf{G}^i \left(\mathbf{y}^o - \mathcal{H}(\hat{\mathbf{x}}^i) \right) , \tag{9}$$

$$\mathbf{G}^i = \hat{\mathbf{C}}^b (\mathbf{J}^i)^T \left(\mathbf{J}^i \hat{\mathbf{C}}^b (\mathbf{J}^i)^T + \gamma^i \, \mathbf{R} \right)^{-1} , \tag{10}$$

which is similar to the mean update formula in the ETKF. For this reason, we call our algorithm iterative ETKF (iETKF). Following [2], the parameter rule of γ^i is given by

$$\gamma^0 > 0 , \tag{11}$$

$$\gamma^{i+1} = \rho^i \gamma^i, \text{ with } 1/r < \rho^i < 1 \text{ for some scalar } r > 1 , \tag{12}$$

$$\lim_{i \to +\infty} \gamma^i = 0 , \tag{13}$$

which ensues a local convergence of the residual norm under suitable conditions. To prevent over-fitting the observations, however, the iteration process (9) stops if the residual norm is lower than $2\sqrt{p}$ for the first time, or if the maximum iteration number is reached.

Experiment results with the 40-dimensional Lorenz 96 model

The observation function is $f(x) = x^3/5$ (applied to all or partial state variables) plus certain noise. The ETKF diverged in all cases below.

Performance of the iETKF in two different observation scenarios

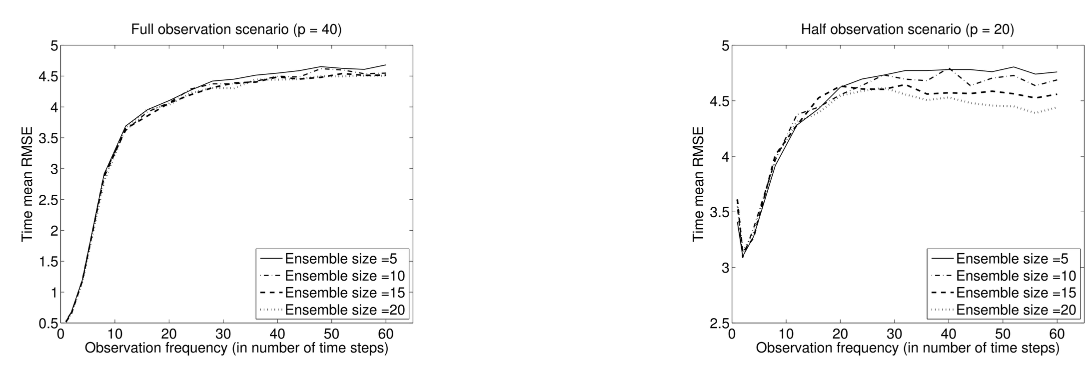

Fig. 1: : Performance of iETKF in full (p=40, all state variables observed) and half (p=20, odd number state variables observed) observation scenarios.

Time series of residual norms and RMSEs

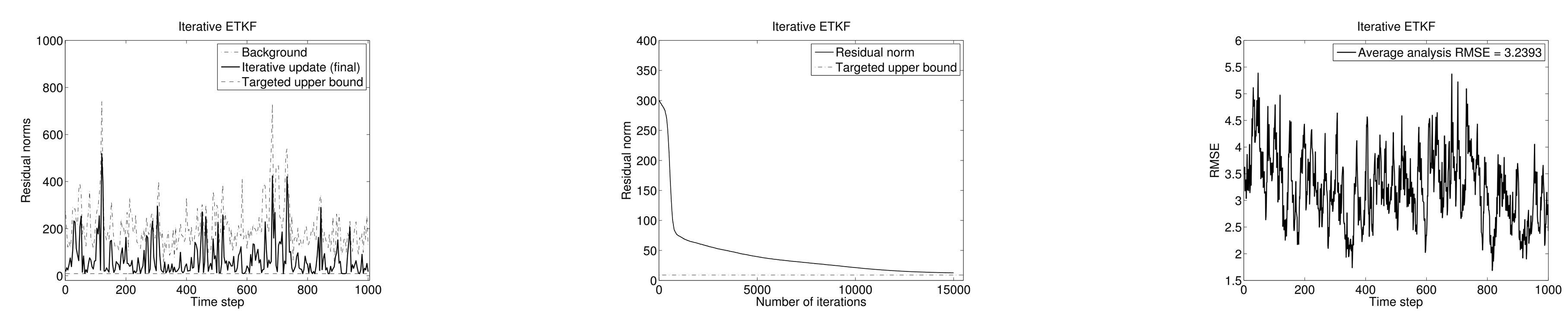

Fig. 2: : Left: time series of residual norms over the assimilation time window; Middle: residual norm (solid) reduction of the iteration process at time step 500, an example used to illustrate how the residual norms of the iterative estimates are gradually reduced toward the targeted upper bound (dash-dotted); Right: time series of RMSEs of the final analysis estimates over the assimilation time window. Experimental settings: half observation scenario, with the observations arriving every 4 steps. The ensemble size is 20.

Performance of the iETKF with model and observation variance errors

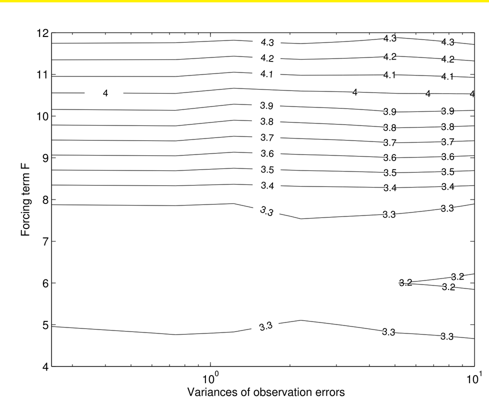

Fig. 3: : Time mean RMSE of the iterative ETKF as functions of the potentially mis-specified forcing term F and the variances of observation errors. For visualization, here the logarithmic scale is also used for the horizontal axis. Experimental settings: half observation scenario, with the observations arriving every 4 steps. The ensemble size is 20. The true $F = 8$, and the true observation error variance = 1.

References

[1] C. H. Bishop, B. J. Etherton, S. J. Majumdar, Adaptive sampling with ensemble transform Kalman filter. Part I: theoretical aspects, Mon. Wea. Rev. 129 (2001) 420–436.

[2] H. W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Springer, 2000.

[3] X. Luo, H. Hoteit, Ensemble kalman filtering with residual nudging: an extension to the state estimation problems with nonlinear observations, submitted.

[4] X. Luo, I. Hoteit, Ensemble Kalman filtering with residual nudging, Tellus A 64 (2012) 17130, open access.

[5] X. Luo, I. Hoteit, Covariance inflation in the ensemble Kalman filter: a residual nudging perspective and some implications, Mon. Wea. Rev. 141 (2013) 3360–3368.

[6] X. Luo, I. Hoteit, Efficient particle filtering through residual nudging, Quart. J. Roy. Meteor. Soc. 140 (2014) 557–572.

[7] J. C. Spall, Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, Automatic Control, IEEE Transactions on 37 (3) (1992) 332–341.

E-mail: xiaodong.luo@iris.no