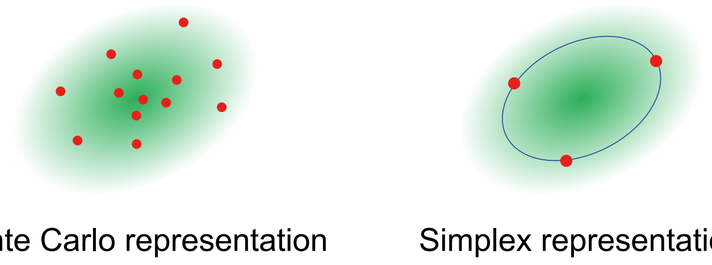
少数アンサンブルによる状態推定

中野 慎也　モデリング研究系・データ同化研究開発センター 助教

Introduction

- Many of ensemble-based data assimilation algorithm are based on a Monte Carlo representation which would provides accurate estimates with infinite particles.
- In data assimilation for high-dimensional systems, however, the number of particles is limited by computational resources. The ensemble size is typically much smaller than the state dimension.
- If the ensemble size N is smaller than the rank of the state covariance matrix, the ensemble would form a simplex in an $(N-1)$-dimensional subspace. Therefore, we can consider a spherical simplex representation [Wang et al., 2004] on the basis of a different principle from a Monte Carlo representation.



Monte Carlo representation　　Simplex representation

Simplex representation

We can represent the mean vector $\boldsymbol{x}_{k|k}$ and the covariance matrix $\mathsf{V}_{k|k}$ by using a set of particles $\{\boldsymbol{x}_{k|k}^{(1)}, \dots \boldsymbol{x}_{k|k}^{(N)}\}$ which satisfies the following two equations:

$$\boldsymbol{x}_{k|k} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_{k|k}^{(i)}, \qquad \mathsf{V}_{k|k} = \frac{1}{N}\sum_{i=1}^{N} (\boldsymbol{x}_{k|k}^{(i)} - \boldsymbol{x}_{k|k})(\boldsymbol{x}_{k|k}^{(i)} - \boldsymbol{x}_{k|k})^T.$$

If $N \geq \mathrm{rank}\mathsf{V}_{k|k} + 1$, we can obtain the ensemble satisfying the above two equations. (We do not need to rely on the law of large numbers.)

In the typical case that $N < \mathrm{rank}\mathsf{V}_{k|k} + 1$, the ensemble provides a reduced-rank approximation of the covariance matrix.

The third and higher order moments of the probability density function (PDF) are ignored. However, in the ensemble Kalman filters or the ensemble square-root filters, it does not matter.

If we define a matrix $\mathsf{X}_{k|k}$ as

$$\mathsf{X}_{k|k} = \frac{1}{\sqrt{N}}\left(\boldsymbol{x}_{k|k}^{(1)} - \boldsymbol{x}_{k|k} \quad \boldsymbol{x}_{k|k}^{(2)} - \boldsymbol{x}_{k|k} \quad \cdots \boldsymbol{x}_{k|k}^{(N)} - \boldsymbol{x}_{k|k}\right),$$

the covariance matrix can be decomposed as $\mathsf{V}_{k|k} = \mathsf{X}_{k|k}\mathsf{X}_{k|k}^T$.

If we assume $\boldsymbol{x}_k$ follows a Gaussian distribution $\mathcal{N}(\boldsymbol{x}_{k|k}, \mathsf{X}_{k|k}\mathsf{X}_{k|k}^T)$, $\boldsymbol{x}_k$ can be associated with a low-dimensional vector of latent variables, $\boldsymbol{z}$ as follows:

$$\boldsymbol{x}_k|\boldsymbol{y}_{1:k} \sim \boldsymbol{x}_{k|k} + \mathsf{X}_{k|k}\boldsymbol{z}_k, \left(\boldsymbol{z}_k \sim \mathcal{N}(\boldsymbol{z}_k; \mathbf{0}, \mathbf{I})\right).$$

The uncertanty of the high-dimensional vector $\boldsymbol{x}_k$ can thus be associated with that of the low-dimensional vector $\boldsymbol{z}_k$.

In the generative model $\boldsymbol{x}_k|\boldsymbol{y}_{1:k} \sim \boldsymbol{x}_{k|k} + \mathsf{X}_{k|k}\boldsymbol{z}_k$, each particle $\boldsymbol{x}_{k|k}^{(i)}$ can be generated when each $\boldsymbol{z}^{(i)}$ is set as

$$\left(\boldsymbol{z}_k^{(1)} \quad \cdots \quad \boldsymbol{z}_k^{(N)}\right) = \begin{pmatrix} \sqrt{N} & & 0 \\ & \ddots & \\ 0 & & \sqrt{N} \end{pmatrix}.$$

This indicates that all of the particles $\boldsymbol{x}_{k|k}^{(1)}, \dots \boldsymbol{x}_{k|k}^{(N)}$ have the same probability density.

Ensemble transform Kalman filter

The ensemble transform Kalman filter [Bishop et al., 2001] obtains the mean of the filtered distribution according to the Kalman filter algorithm:

$$\overline{\boldsymbol{x}}_{k|k} = \overline{\boldsymbol{x}}_{k|k-1} + \mathsf{K}_k\left(\boldsymbol{y}_k - \mathsf{H}_k\overline{\boldsymbol{x}}_{k|k-1}\right)$$

The square root of the covariance matrix is calculated by $\mathsf{X}_{k|k} = \mathsf{X}_{k|k-1}\mathsf{T}_k$.

If we consider the eigenvalue decomposition of the following $N \times N$ small-sized matrix:

$$\mathsf{X}_{k|k-1}^T\mathsf{H}_k^T\mathsf{R}_k^{-1}\mathsf{H}_k\mathsf{X}_{k|k-1} = \mathsf{U}\Lambda\mathsf{U}^T,$$

the matrices K_k and T_k can be obtained as follows:

$$\mathsf{K}_k = \mathsf{X}_{k|k-1}\mathsf{U}(\mathbf{I}+\Lambda)^{-1}\mathsf{U}^T\mathsf{X}_{k|k-1}^T\mathsf{H}_k^T\mathsf{R}_k^{-1},$$
$$\mathsf{T}_k = \mathsf{U}(\mathbf{I}+\Lambda)^{-\frac{1}{2}}\mathsf{U}^T.$$

Prediction with a small ensemble

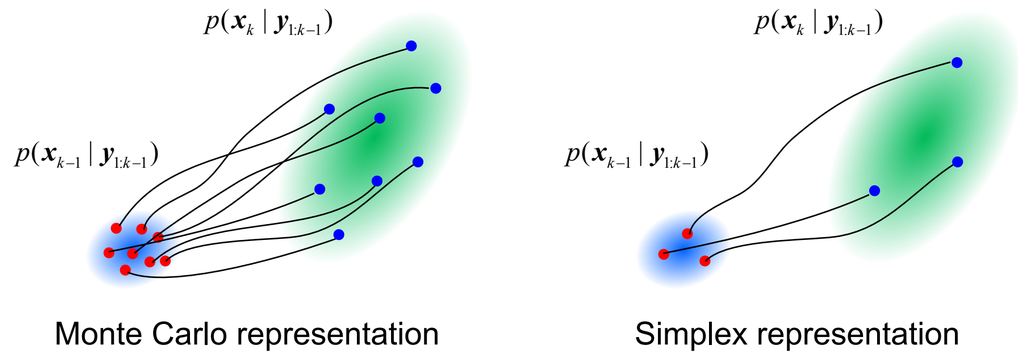
For both the representations, if we apply the dynamical system model for each particle:

$$\boldsymbol{x}_{k|k-1}^{(i)} = \mathcal{M}(\boldsymbol{x}_{k-1|k-1}^{(i)}),$$

we obtain an ensemble representing the forecast distribution $p(\boldsymbol{x}_k|\boldsymbol{y}_k)$.

The Monte Carlo representation would converge to the exact solution of the Chapman-Kolmogorov integral with the infinite ensemble size.

The simplex representation provides an estimate with second-order accuracy for the first and second moments of the solution of the Chapman-Kolmogorov integral.

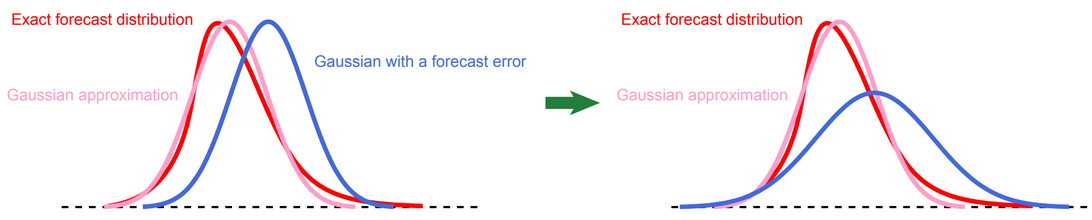


Monte Carlo representation　　Simplex representation

Even if the system dynamics is deterministic and fully known, a forecast with a limited-sized ensemble may have some errors due to the nonlinearity of the dynamical system of higher order than second.

Covariance inflation

- If the number of particles is limited, the ensemble mean can deviate from the exact forecast mean.
- The inflation of the covariance matrix is widely used to avoid this problem.



The following rough argument would suggest that the discrepancy from the exact forecast can be reduced by inflating the covariance.

- We want to attain the Gaussian forecast distribution, $\mathcal{N}(\boldsymbol{x}_{k|k-1}, \mathsf{V}_{k|k-1})$, where $\boldsymbol{x}_{k|k-1}$ and $\mathsf{V}_{k|k-1}$ are the exact mean vector and the exact covariance matrix, respectively.
- We estimate the forecast distribution as a Gaussian $\mathcal{N}(\bar{\boldsymbol{x}}_{k|k-1}, \alpha^2\bar{\mathsf{V}}_{k|k-1})$, where $\bar{\boldsymbol{x}}_{k|k-1}$ and $\bar{\mathsf{V}}_{k|k-1}$ are the sample mean and the sample covariance of the forecast ensemble.
- The cross entropy

$$-\int \mathcal{N}(\boldsymbol{x}_{k|k-1}, \mathsf{V}_{k|k-1}) \log \mathcal{N}(\bar{\boldsymbol{x}}_{k|k-1}, \alpha^2\bar{\mathsf{V}}_{k|k-1})\, d\boldsymbol{x}$$

is minimized when

$$\alpha^2 = \frac{1}{\dim \boldsymbol{x}_k}\mathrm{tr}\left[\bar{\mathsf{V}}_{k|k-1}^{-1}\mathsf{V}_{k|k-1} + \bar{\mathsf{V}}_{k|k-1}^{-1}(\bar{\boldsymbol{x}}_{k|k-1} - \boldsymbol{x}_{k|k-1})(\bar{\boldsymbol{x}}_{k|k-1} - \boldsymbol{x}_{k|k-1})^T\right].$$

- If $E\left[\bar{\mathsf{V}}_{k|k-1}^{-1}\right] = \mathsf{V}_{k|k-1}^{-1}$,

$$E[\alpha^2] = 1 + \frac{1}{\dim \boldsymbol{x}_k}\mathrm{tr}\left[\mathsf{V}_{k|k-1}^{-1}(\bar{\boldsymbol{x}}_{k|k-1} - \boldsymbol{x}_{k|k-1})(\bar{\boldsymbol{x}}_{k|k-1} - \boldsymbol{x}_{k|k-1})^T\right] \geq 1.$$

References

Bishop, C. H., Etherton, R. J. and Majumdar, S. J. (2001): Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. Mon. Wea. Rev. 129, 420-436.

Wang, X., Bishop, C. H. and Julier, S. J. 2004. Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble?. Mon. Wea. Rev. 132, 1590-1605.