等価重み粒子フィルタを拡張したパラメータ推定

佐藤　峰斗　　総合研究大学院大学　統計科学専攻　博士課程（5年一貫制）5年

Introduction

Motivation: To improve robustness and efficiency for estimating the state and time-varying parameters in nonlinear high-dimensional systems.

Difficulty: Typical state augmentation techniques tend to fail to detect parameter changes unless there are enough correlations.

Solution: Introduce correlated perturbation and use an efficient particle filter "IEWPF" [1] combined with adaptive moment estimation (Adam [2]) based optimization technique from machine learning.

Methods

State augmentation with correlated perturbation

Time evolution of state in augmented state-space model can be written as

$$\begin{pmatrix} x^n \\ \theta^n \end{pmatrix} = \begin{pmatrix} f(x^{n-1}, \theta^{n-1}) \\ \theta^{n-1} \end{pmatrix} + \begin{pmatrix} \beta^n \\ \eta^n \end{pmatrix}, \quad \beta \sim N(0, Q_\beta), \quad \eta \sim N(0, Q_\eta).$$

By using Taylor series expansion:

$$f(x^{n-1}, \theta^{n-1}) \cong f(x^{n-1}, \theta^{n-2}) + \frac{\partial f}{\partial \theta}(\theta^{n-1} - \theta^{n-2}),$$

time evolution can be approximated as

$$z^n = \begin{pmatrix} x^n \\ \theta^{n-1} \end{pmatrix} = \begin{pmatrix} f(x^{n-1}, \theta^{n-2}) + \frac{\partial f}{\partial \theta}\eta^{n-1} + \beta^n \\ \theta^{n-2} + \eta^{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} f(x^{n-1}, \theta^{n-2}) \\ \theta^{n-2} \end{pmatrix} + \tilde{\beta}^n,$$

where

$$\tilde{\beta}^n \sim N(0, \tilde{Q}^n), \qquad \tilde{Q} = \begin{pmatrix} \frac{\partial f}{\partial \theta} Q_\eta \frac{\partial f}{\partial \theta}^T + Q_\beta & \frac{\partial f}{\partial \theta} Q_\eta \\ Q_\eta \frac{\partial f}{\partial \theta}^T & Q_\eta \end{pmatrix}.$$

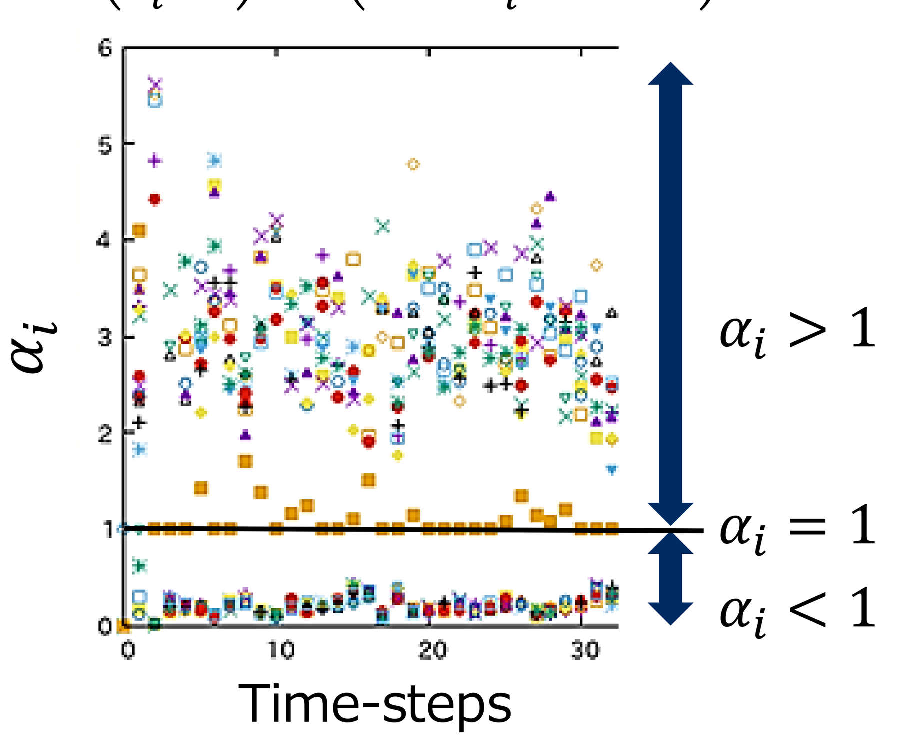
Combination of IEWPF and parameter correction

The Implicit Equal-Weights Particle Filter (IEWPF) [1] can avoid filter degeneracy by using a proposal density which the weight of each particle can be equal to the target weight. The method mainly contains two procedures dependent on the observation: (1) state and parameter update by IEWPF at the observation step, and (2) parameter correction with proposal density in the next step after observation.

(1) At observation step:

- Compute parameter differentiation using model and parameter perturbation, and generate augmented covariance matrix.
- Perform forecast using model transition and apply IEWPF to both variables and parameters. The augmented state sampled from posterior distribution can be described as

$$\begin{pmatrix} x_i^n \\ \theta_i^{n-1} \end{pmatrix} = \begin{pmatrix} f(x_i^{n-1}, \theta_i^{n-2}) \\ \theta_i^{n-2} \end{pmatrix} + K(y^n - Hf(x_i^{n-1}, \theta_i^{n-2})) + \alpha_i^{1/2} P^{1/2} \xi_i,$$



where K is Kalman gain and α is a scalar. Note that α need to be chosen to satisfy equal weights for all particles.

Fig.1. An example of an analytical solution[1] of α during sequential estimation using a linear model. (As detailed in Case study.) Different colors or marks indicate each particle (20 in total).

(2) The next step after observation:

- Draw samples from the proposal transition density q, instead of the original transition density p.

$$q(\theta^n | \theta_i^{n-1}, y^n) = N(\theta_i^{n-1} - \lambda g_i^n, Q_\eta), \quad g_i^n \in \nabla L_i^n,$$

where λ is a scalar factor and g represents the term that corrects the parameter θ so that the loss function L becomes smaller.

Adam-based parameter correction term

- From the analogy of Adam (Adaptive Moment Estimation) [2], the parameter correction term g can be chosen as follows:

$$g^n = \widehat{m_i^n} / \sqrt{\widehat{v_i^n}}, \quad \widehat{m_i^n} = m_i^n/(1-\mu), \ \widehat{v_i^n} = v_i^n/(1-\rho),$$

where m and v are momentum and norm of the gradient of the loss function as following equation, respectively:

$$m_i^n = \mu m_i^{n-1} + (1-\mu)\nabla L_i^n, \quad v_i^n = \rho v_i^{n-1} + (1-\rho)(\nabla L_i^n)^2.$$

Note that the hyper-parameters μ and ρ control the decay rates of these moving averages. The loss function L can be expressed using the log-likelihood at the last observation y as follows:

$$L_i^n(\theta) = -2\ln\left[p(y^n | x_i^{n-1}, \theta_i^{n-2})\right]$$

$$\propto \left(y^n - Hf(x_i^{n-1}, \theta_i^{n-2})\right)^T (HQH^T + R)^{-1} \left(y^n - Hf(x_i^{n-1}, \theta_i^{n-2})\right).$$

Case study : Linear model (N_j independent systems)

Linear models are useful for comparison with analytical solutions before validating with nonlinear high-dimensional models. Consider N_j independent linear model equation with additive parameters, and observations for twin experiments:

$$\begin{pmatrix} x_j^n \\ \theta_m^{n-1} \end{pmatrix} = \begin{pmatrix} x_j^{n-1} \\ \theta_m^{n-2} \end{pmatrix} + \begin{pmatrix} \theta_m^{n-2} \\ 0 \end{pmatrix} + \tilde{\beta}^n,$$

$$y^n = x_{\text{truth}}^n + w^n, \qquad \tilde{\beta}^n \sim N(0, \tilde{Q}^n), \quad w \sim N(0, R),$$

where x_j (j=1,…,N_j) and θ_m (m=1,…,3) are the state variables and parameters, respectively. In the following experiments, the true model error is set to 0 and observations y is generated with observation error whose covariance matrix is R=0.01. For the assimilation, we choose model error whose covariance matrix is Qη=0.04 and observation error whose covariance matrix is R=0.01. The parameter perturbations are sampled from pdf N(0, 1E-4).

Evaluation of the shape of the posterior pdf (N_j =100)

We conduct 100 independent data assimilation runs including 3 parameters for 5000 time steps with 20 particles, and compare them with the results of using the Kalman filter (KF), in order to evaluate the shape of the posterior pdf.

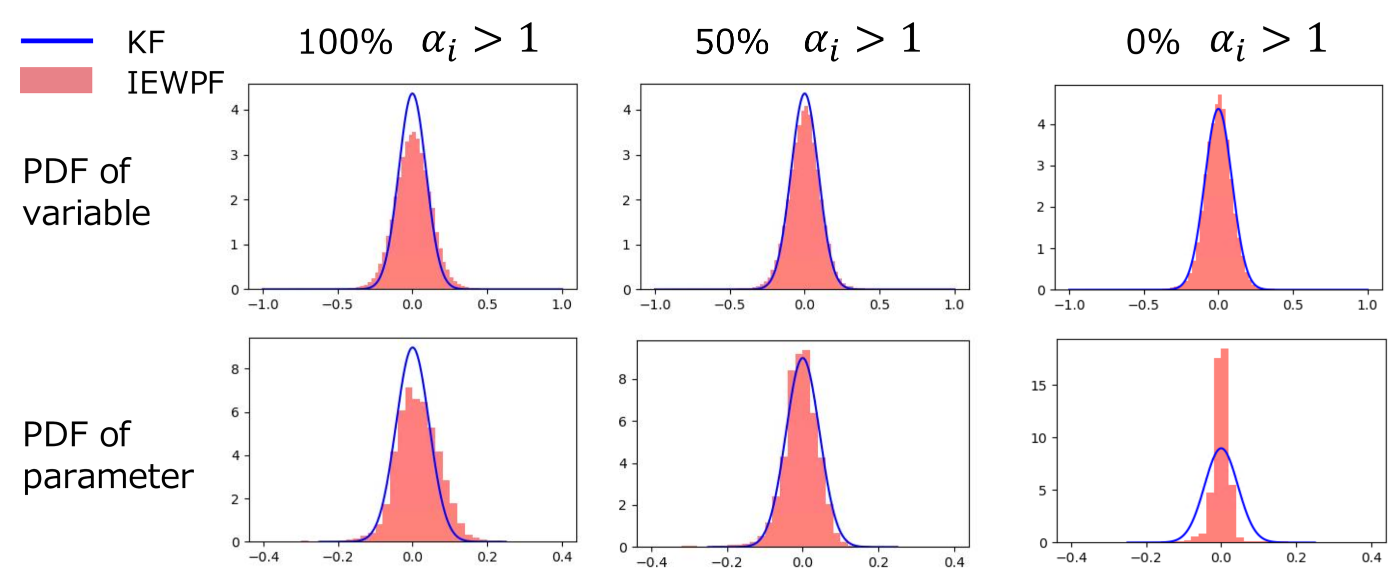


Fig.2. Comparison of the posterior pdf between KF and IEWPF using selected scalar α at different percentages.

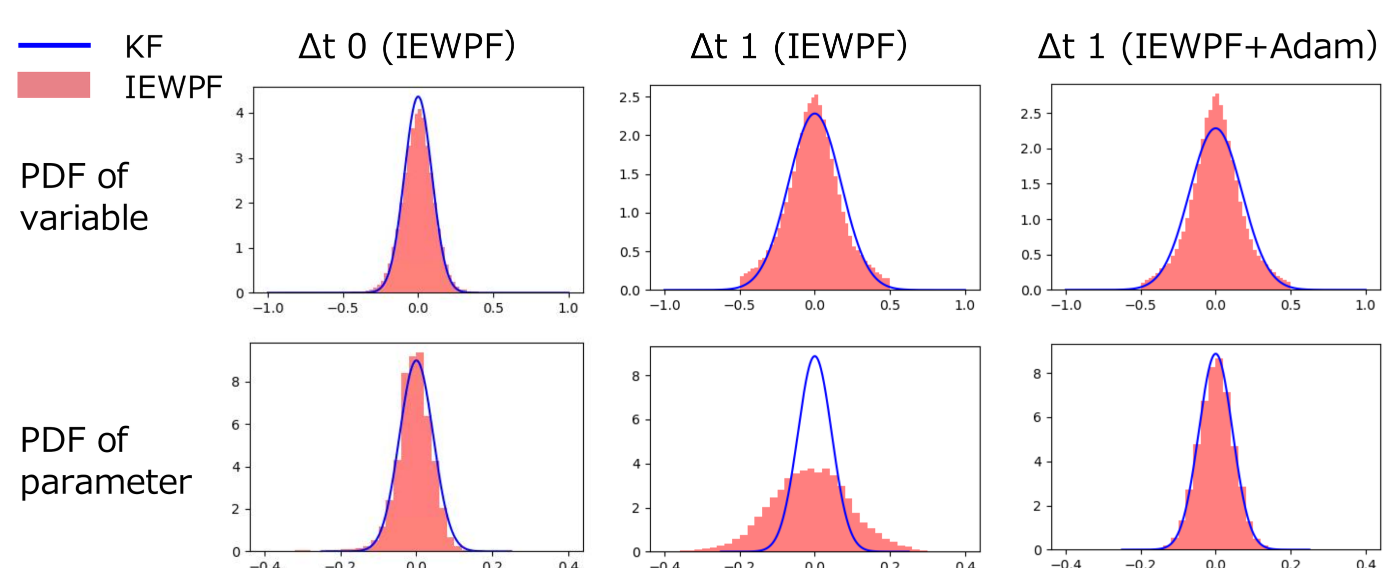


Fig. 3. Dependency of different assimilation interval: Δt 0 (every step), Δt 1 (every other step).

Conclusion

The proposed method successfully avoid filter degeneracy and the posterior pdf is consistent with that of KF at appropriate selection of α. Furthermore, the parameter correction has capability to adjust the distribution of parameters.

References

1. Zhu M., van Leeuwen PJ, Amezcua J., 2016, 'Implicit equal-weights particle filter', Q. J. R. Meteorol. Soc. 142: 1904–1919.
2. P. Kingma D. and Lei Ba J., 'Adam: A method for stochastic optimization', conference paper at ICLR 2015