

# 等価重み粒子フィルタによる非線形モデルの時変パラメータ推定

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

## 1. 概要

**目的:** 気象や海洋等の高次元・非線形システムの予測精度向上に向け、予測精度を左右するモデルパラメータの推定手法を提案する。

**課題:** 粒子フィルタは非線形モデルに適用可能だが、退化が課題(図1)。

**アプローチ:** 全粒子の重みを等しくすることで退化を抑制する、等価重み粒子フィルタ (IEWPF) [1] をパラメータ推定に拡張する。

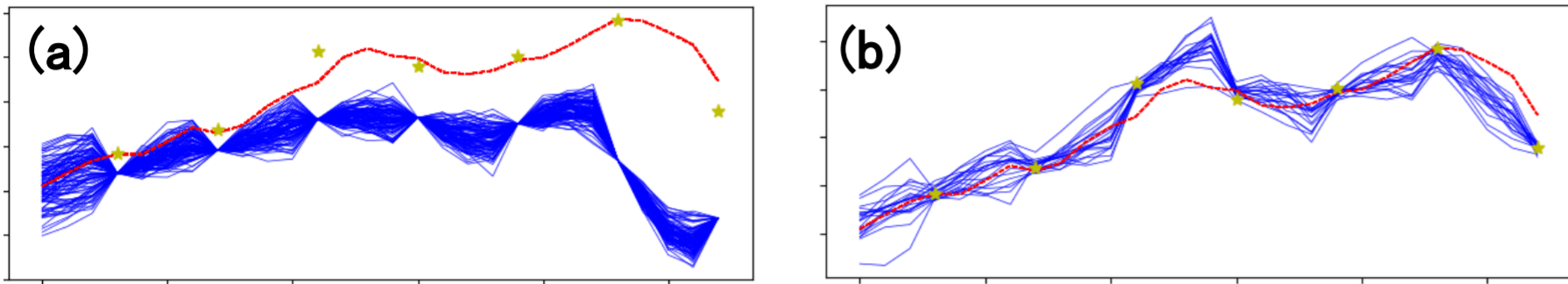


図1. (a)一般的な粒子フィルタと(b)IEWPFの軌跡の例(赤:真値、青:各粒子)

## 2. 提案手法

### 2-1. 変数・パラメータに相関のある拡大状態空間モデル

下記のような一般的な拡大状態空間モデルへの拡張では、変数とパラメータの摂動項(システムノイズ)が独立である。(n: 時間ステップ)

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

非線形モデル  $f$  へのパラメータの寄与を以下のように一次近似すると

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

以下のように相関のある形で、システムモデルを表すことができる:

$$z^n \equiv \begin{pmatrix} x^n \\ \theta^{n-1} \end{pmatrix} = \begin{pmatrix} f(x^{n-1}, \theta^{n-2}) + \frac{\partial f}{\partial \theta} \eta^{n-1} + \beta^n \\ \theta^{n-2} + \eta^{n-1} \end{pmatrix} \equiv \tilde{f}(z^{n-1}) + \tilde{\beta}^n, \quad (3)$$

ここで

$$\tilde{\beta}^n \sim N(0, \tilde{Q}^n), \quad \tilde{Q} = \begin{pmatrix} Q_\beta + \frac{\partial f}{\partial \theta} Q_\eta \frac{\partial f^T}{\partial \theta} & \frac{\partial f}{\partial \theta} Q_\eta \\ Q_\eta \frac{\partial f^T}{\partial \theta} & Q_\eta \end{pmatrix}. \quad (4)$$

### 2-2. 等価重み粒子フィルタの実現方法

The Implicit Equal-Weights Particle Filter (IEWPF) [1] は、提案分布を

$$z_i^n = \zeta_i^n + \alpha_i^{1/2} P^{1/2} \xi_i^n, \quad \xi \sim N(0, I) \quad (5)$$

で与える。ここで  $\zeta_i^n$ ,  $P$  は最適提案分布の平均と分散共分散行列である。

$$\zeta_i^n = \tilde{f}(z_i^{n-1}) + QH^T(HQH^T + R)^{-1} (y^n - Hf(z_i^{n-1})), \quad P = (Q^{-1} + H^T R^{-1} H)^{-1} \quad (6)$$

式(5)は提案分布  $q(z^n | z_{1:n}^{n-1}, y^n) = \frac{q(\xi)}{\| \frac{dz}{d\xi} \|}$  導入による変数変換  $z_i \rightarrow \xi_i$

を表し、各粒子の重みが等しくなるためには  $\alpha_i$  が次式を満たせば良い。

$$-2 \log w_i^n = -2 \log w_i^{n-1} - 2 \log \left( \frac{p(y^n | z_i^n) p(z_i^n | z_i^{n-1})}{q(\xi)} \left\| \frac{dz}{d\xi} \right\| \right) \quad (7)$$

$w_i^n$  は粒子  $i$ , ステップ  $n$  における重みを表す。観測・システムノイズをガウス分布、観測モデル  $H$  を線形とし、高次元(変数  $x$  の次元  $N_x \rightarrow \infty$ )の下で、式(7)を満たす  $\alpha_i$  は、解析解を Lambert W function から得ることができる[1]。

### 2-3. パラメータ予測へのオンライン最適化手法の導入

観測値があるステップ  $n$  では、粒子  $i$  の事後分布(フィルタ分布)からのサンプリングは式(5), (6) より得られる。一方、次の観測までの、観測値が無いステップ  $n+1$  の予測は、ステップ  $n$  の観測値  $y^n$  の情報を使った提案分布  $q(z_i^{n+1} | z_i^n, y^n)$  を用いて、パラメータの予測分布を次式で得る:

$$p(z^n | z_i^{n-1}, y^{1:n}) = \frac{1}{N} \sum_{i=1}^{N_p} \frac{p(z^n | z_i^{n-1})}{q(z^n | z_i^{n-1}, y^n)} q(z^n | z_i^{n-1}, y^n). \quad (8)$$

## References

- Zhu, Mengbin, Peter Jan Van Leeuwen, and Javier Amezcua. "Implicit equal-weights particle filter." Quarterly Journal of the Royal Meteorological Society 142.698 (2016): 1904-1919.
- Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv, 1412.6980.
- Satoh, Mineto, Peter Jan van Leeuwen, and Shin'ya Nakano. "Online state and time-varying parameter estimation using the implicit equal-weights particle filter." Quarterly Journal of the Royal Meteorological Society (2024).

この提案分布  $q$  を以下のように生成する:

$$q(z_i^{n+1} | z_i^n, y^n) = N \left( \tilde{f}(z_i^n) + \left( -\lambda g(\theta_i^{n-2}, y^n) \right), \tilde{Q} \right). \quad (9)$$

$\lambda$  は Step-size factor,  $g$  はステップ  $n$ , 粒子  $i$  の nudging 項である。ここで、

$$L_i^n(\theta_i^{n-2}) = \left( y^n - Hf(x_i^{n-1}, \theta_i^{n-2}) \right)^T R^{-1} \left( y^n - Hf(x_i^{n-1}, \theta_i^{n-2}) \right) \quad (10)$$

の勾配を用いて関数  $g$  を定義:  $g \propto \nabla_\theta L_i^n$  すると、式(9)は式(10)の評価関数の勾配が小さくなる方向にパラメータ  $\theta$  を nudging する。関数  $g$  は、最適化アルゴリズムのAdam[2]を用いて次式で定義する:

$$g(\theta_i^{n-2}, y^n) = \hat{m}_i^n / \sqrt{\hat{v}_i^n}, \quad \hat{m}_i^n \propto \nabla_\theta L_i^n, \quad \hat{v}_i^n \propto (\nabla_\theta L_i^n)^2. \quad (11)$$

## 3. 数値実験: Lorenz-96 モデルによる双子実験

### 3-1. 外力項にパラメータを導入したモデル (1000次元)

Lorenz-96 モデルを用いて、高次元・非線形システムにおける時間変化するパラメータ推定の妥当性を評価した[3]。Lorenz-96 モデル:

$$\frac{d}{dt} x_j = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F_j \quad (12)$$

の外力項  $F_j$  に、3つのパラメータ ( $\theta_0, \theta_1, \theta_2$ ) を導入して以下のように表す:

$$F_j = \theta_{0\text{true}} \theta_0 + \theta_{1\text{true}} \theta_1 \sin \left( \frac{2\pi}{\theta_{2\text{true}} \theta_2} j \right) \quad (13)$$

( $\theta_{0\text{true}}, \theta_{1\text{true}}, \theta_{2\text{true}}$ ) = (8, 4, 1000) で、これらの係数となっている3つパラメータの真値を 1.0 から変化させて疑似観測データを生成した。変数と観測データの分散共分散行列  $Q_\beta$ ,  $R$  は、それぞれ対角項 0.1, 隣接(上下)対角項 0.025、 $R = 0.02 I$  とし、疑似観測データ生成時と推定時で同じ値を用いた。パラメータの分散共分散行列  $Q_\eta$  は対角とし、観測データ生成時は 0 とした。

### 3-2. パラメータの分散共分散と Step-size factor 依存性

変数  $x$  の次元は1000、全変数が4ステップ毎に観測され、粒子数  $N_p = 20$  とした。図2に(9)式の Step-size factor  $\lambda$  を 0.001 としてパラメータの分散共分散行列  $Q_\eta$  の対角値を  $1.0 \times 10^{-6}, 5.0 \times 10^{-6}, 1.0 \times 10^{-5}, 5.0 \times 10^{-5}$  と変えた場合 (a) と、 $Q_\eta$  の対角値を  $5.0 \times 10^{-6}$  として  $\lambda$  を 0.0005, 0.001, 0.002, 0.004 と変えた場合 (b) の軌跡を示す。各パラメータの真値は 200 ステップで +30% 変化したと仮定した。図3にそれぞれの場合の時間平均 RMSE 及び粒子広がり (Spread) を予測とフィルタリングステップに分けて示す。 $Q_\eta$  の対角値が大きくなると RMSE が小さくなり、RMSE と Spread の比が 1 に近づくが Spread が大きくなる。一方、Step-size factor  $\lambda$  が大きくなると Spread の変化は少なく、RMSE が小さくなる。本結果からパラメータの時間変化を推定できることがわかるが、実用上は  $Q_\eta$  と  $\lambda$  を問題に応じて適切に設定する必要がある。

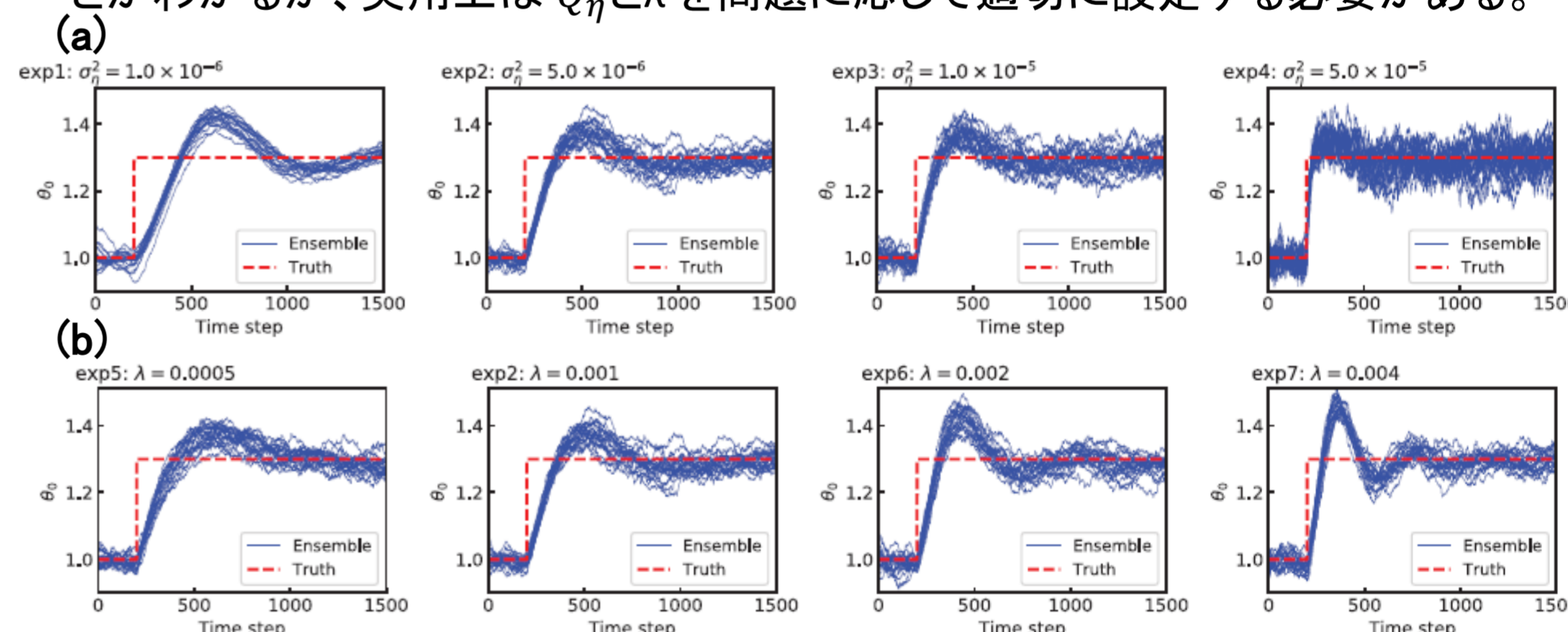


図2. パラメータ  $\theta_0$  の真値(赤点線)と粒子(青線)の軌跡。(a)パラメータの分散共分散行列  $Q_\eta$  依存性、(b) Step-size factor  $\lambda$  依存性

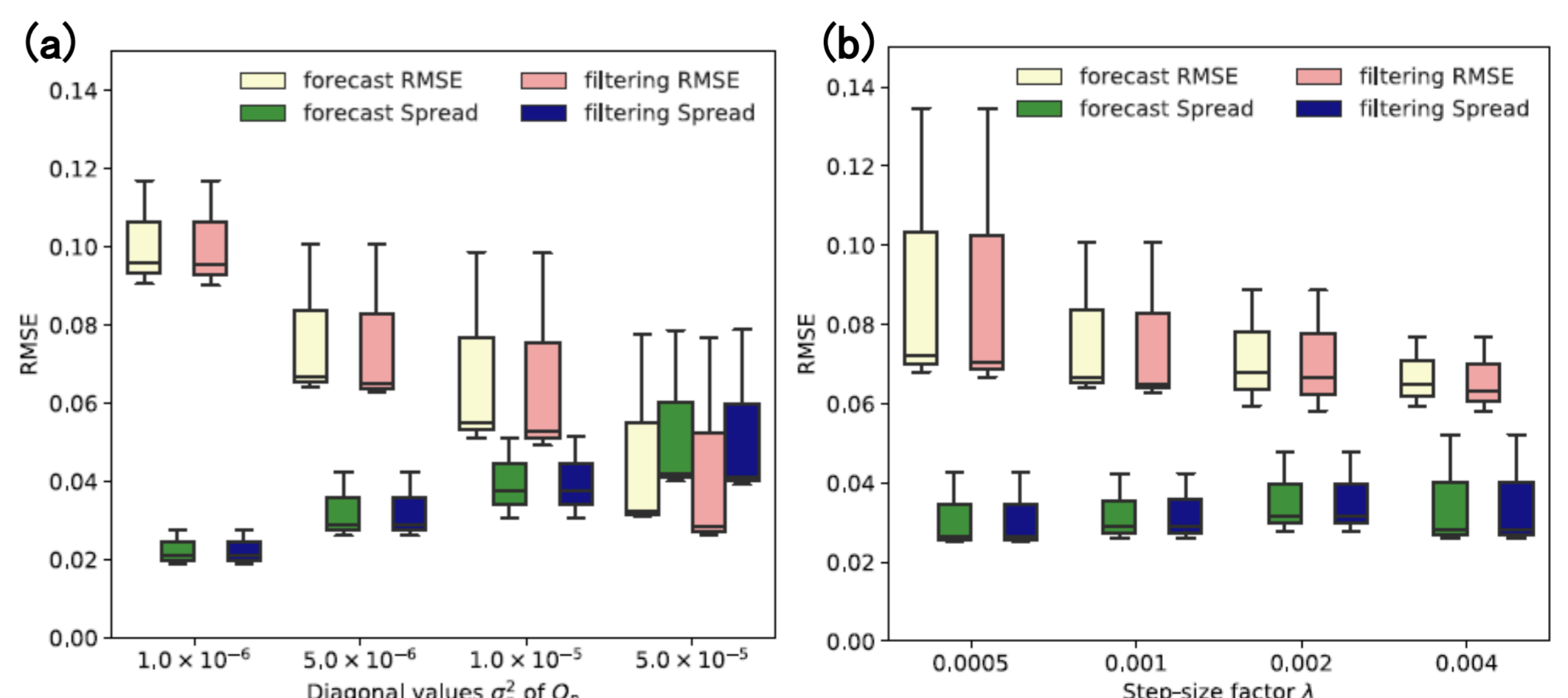


図3. 100-1500 ステップで時間平均した RMSE 及び粒子広がり (Spread) の (a)パラメータの分散共分散行列  $Q_\eta$  と、(b) Step-size factor  $\lambda$  依存性