2023年5月26日 統計数理研究所 オープンハウス

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： 気象や海洋などの高次元・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も適用可能だが、退化の問題がある。

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。

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)

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

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

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

システムモデルは以下のように相関のある形で表される。

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

ここで

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

2-2. 等価重み粒子フィルタの実現方法

Zhu et al. によって提案された The Implicit Equal-Weights Particle Filter (IEWPF) [1] は、提案分布を以下の式(5)で与える。

$$x_i^n = x_i^{n,a} + \alpha_i^{1/2} P^{1/2} \xi_i, \qquad \xi \sim N(0, \mathbf{I}) \tag{5}$$

ここで、$x^{n,a}$, P は以下の最適提案分布の平均と分散共分散行列である。

$$x_i^{n,a} = f\left(x_i^{n-1}\right) + QH^T(HQH^T + R)^{-1}\left(y^n - Hf\left(x_i^{n-1}\right)\right), P = (Q^{-1} + H^T R^{-1} H)^{-1} \tag{6}$$

式(5)は提案分布 $q(x^n | x_{1:N}^{n-1}, y^n) = \frac{q(\xi)}{\left\|\frac{dx}{d\xi}\right\|}$ 導入による変数変換 $x_i \rightarrow \xi_i$ を表し、各粒子の重みが等しくなるために α_i が満たすべき条件は、

$$-2\log w_i^n = -2\log w_i^{n-1} - 2\log\left(\frac{p(y^n|x_i^n)\, p\left(x_i^n | x_i^{n-1}\right)}{q(\xi)}\left\|\frac{dx}{d\xi}\right\|\right) \tag{7}$$

と表される。ここで 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 を線形とすると、

$$e^{-\frac{\alpha_i \gamma_i}{2}} (\alpha_i \gamma_i)^{\frac{N_x}{2}-1} \left|\frac{d(\alpha_i \gamma_i)}{d\gamma_i}\right| = e^{-\frac{\gamma_i}{2}} \gamma_i^{\frac{N_x}{2}-1} e^{-\frac{c_i}{2}} \tag{8}$$

と変形できる[1]。ただし $\gamma_i = \xi_i^T \xi$, c は重みの定数項を示す。式(8)は、不完全ガンマ関数 Γ を用いて

$$\pm\Gamma\left(\frac{N_x}{2}, \frac{\alpha_i \gamma_i}{2}\right) = \Gamma\left(\frac{N_x}{2}, \frac{\gamma_i}{2}\right) e^{-\frac{c_i}{2}} \tag{9}$$

と表される[1]。この式(9)を満たすα_iは、高次元の仮定（変数 x の次元 $N_x \to \infty$）の下で、解析解は Lamber W function で表される。

図1に各時間ステップで算出されたα_i の分布を示す。 図1に示すように、$\alpha = 1$ を境に2つの branch をもつため、$\alpha \leq 1$ と $\alpha \geq 1$ から適切に選択する必要がある。

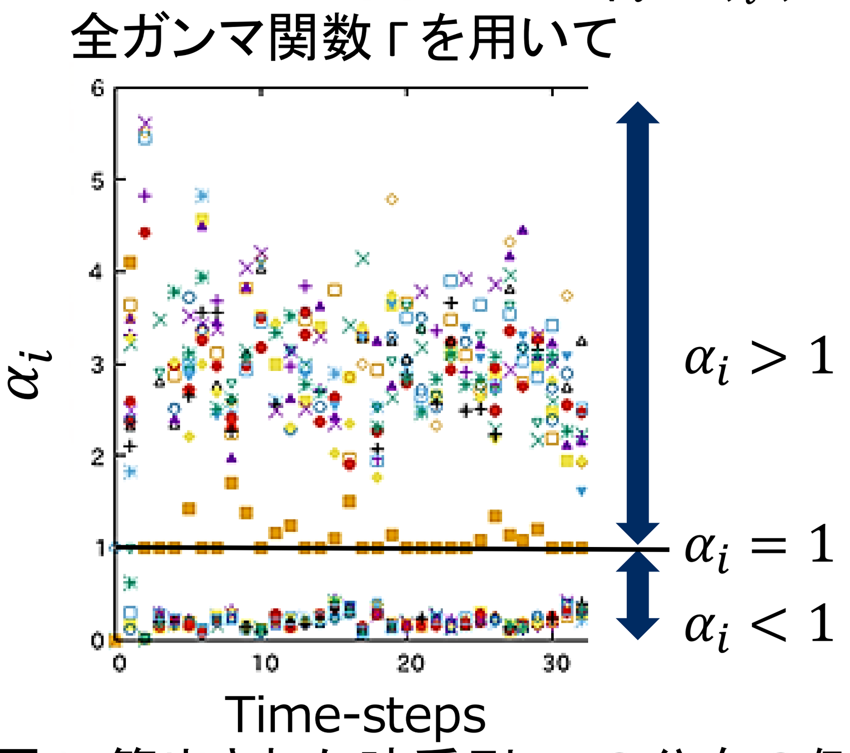

図1. 算出された時系列 α_i の分布の例

2-3. パラメータ予測へのオンライン最適化手法の導入

式(5)(6) より、観測値があるステップ n ,粒子 i の事後分布（フィルタ分布）からのサンプリングは、

$$\begin{pmatrix} x_i^n \\ \theta_i^{n-1} \end{pmatrix} = \begin{pmatrix} f\,(x_i^{n-1}, \theta_i^{n-2}) \\ \theta_i^{n-2} \end{pmatrix} + K\left(y^n - Hf(x_i^{n-1}, \theta_i^{n-2})\right) + \alpha_i^{1/2} P^{1/2} \xi_i \tag{10}$$

と表される。ただし、$\mathrm{K} = \tilde{Q}H^T\left(H\tilde{Q}H^T + R\right)^{-1}, P = \left(\tilde{Q}^{-1} + H^T R^{-1} H\right)^{-1}$.

$\tilde{Q}$は式(4)で与えられる。観測値のないステップは、観測データ y^n の情報を入れた提案分布 $q(\theta^n | \theta_i^{n-1}, y^n)$ を用いて、パラメータの予測分布を

$$p\left(\theta^n | \theta_i^{n-1}, y^{1:n}\right) = \frac{1}{N}\sum_{i=1}^{N_p} \frac{p\left(\theta^n | \theta_i^{n-1}\right)}{q\left(\theta^n | \theta_i^{n-1}, y^n\right)} q\left(\theta^n | \theta_i^{n-1}, y^n\right) \tag{11}$$

と表す。 この提案分布 q を以下のように生成し、

$$q\left(\theta^n | \theta_i^{n-1}, y^n\right) = N(\theta^{n-1} - \lambda \mathrm{g}_\mathrm{i}^\mathrm{n}, Q) \qquad \lambda : \text{Step size factor} \tag{12}$$

関数 $\mathrm{g}_\mathrm{i}^\mathrm{n}$ をステップn, 粒子 i のロス関数：

$$L_i^n(\theta) = \left(y^n - Hf\left(x_i^{n-1}, \theta_i^{n-2}\right)\right)^T (HQH^T + R)^{-1} \left(y^n - Hf\left(x_i^{n-1}, \theta_i^{n-2}\right)\right) \tag{13}$$

の勾配を用いて計算すると、式(12)は、

$$\theta^n \leftarrow \theta^{n-1} - \lambda g^n \ \ (g^n \propto \nabla L_i^n) \tag{14}$$

と、パラメータ θ をステップ毎にロス関数の勾配で更新することを示す。式(14)の更新には、最適化で用いられる各種アルゴリズムが適用できる。

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

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する。Lorenz-96 モデル：

$$\frac{d}{dt}x_j = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F_j \tag{15}$$

の外力項 F_j を、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：

$$F_j = \theta_{0\mathrm{true}}\theta_0 + \theta_{1\mathrm{true}}\theta_1 \sin\left(\frac{2\pi}{\theta_{2\mathrm{true}}\theta_2} j\right) \tag{16}$$

$(\theta_{0\mathrm{true}}, \theta_{1\mathrm{true}}, \theta_{2\mathrm{true}}) = (8, 4, 1000)$ で、パラメータ3つの真値を 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_β, R は、それぞれ対角項 0.1, 隣接（上下）対角項 0.025 、$R = 0.01\, I$ 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_η は対角とし、観測データ生成時は 0 、推定時は 1.0×10^{-5}とした。

3-2. 時変パラメータ推定の手法間比較 (1000次元)

各パラメータの真値と粒子軌跡の比較を図2に示す。変数 x の次元は1000で、全ての変数が 4 ステップ毎に観測される条件とし、粒子数 N_p =20 とした。

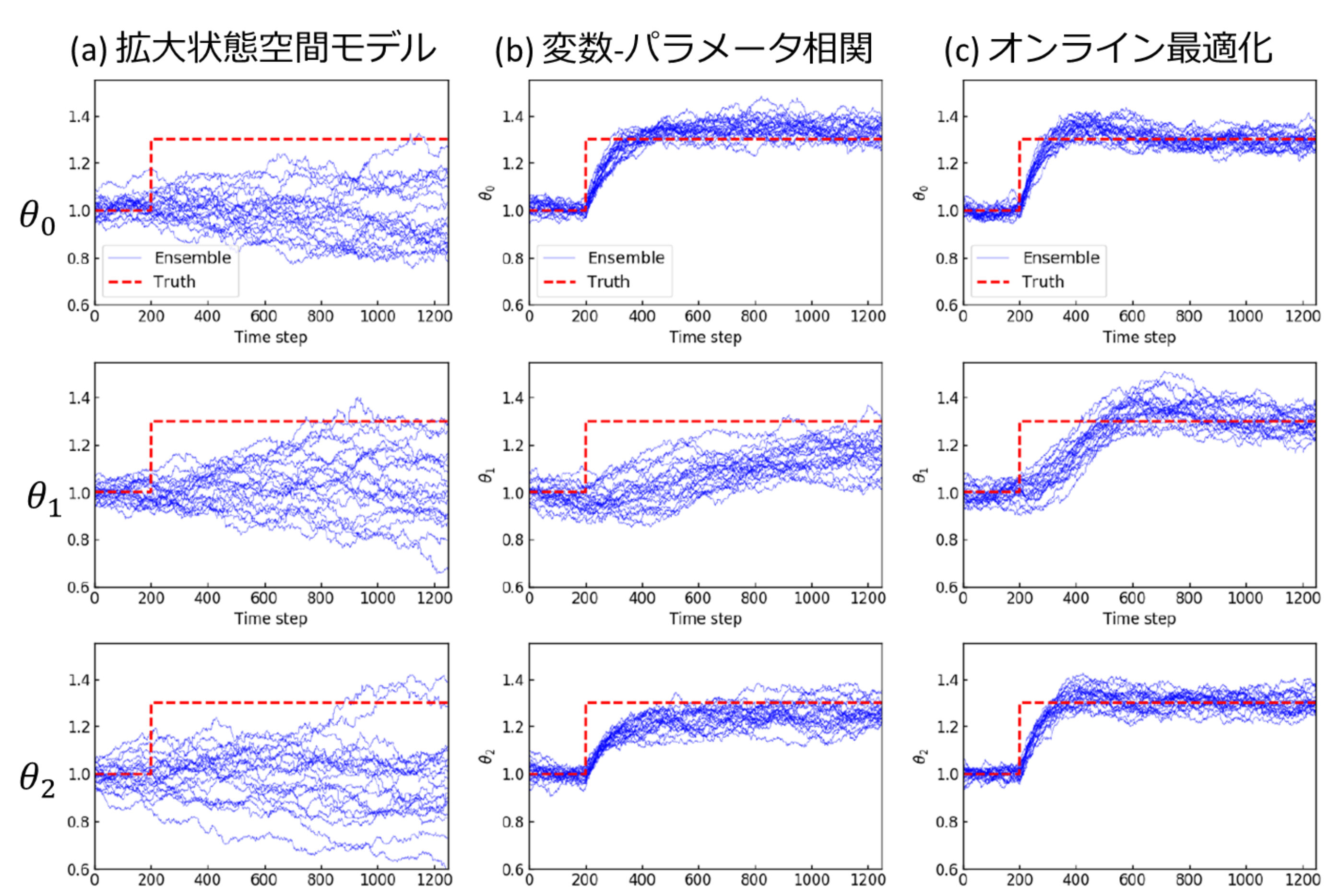

図2. 各パラメータの真値(200 ステップで +30%) と粒子軌跡の手法間比較

図1に示した式(9)を満たすα の選び方は、$\alpha \leq 1$ と $\alpha \geq 1$ から50%ずつサンプリングした[1]。図2 (c) では、式(14) の更新アルゴリズムに Adam [2] を用いて λ =0.001 とした。図2 の比較より、(c) はパラメータ変化後の真値への追従が早いことがわかる。

4. 結論

等価重み粒子フィルタを拡張し、高次元・非線形モデルのパラメータ推定における提案手法の有効性を、1000次元・20粒子の条件で確認した。特にオンライン最適化の導入は、パラメータ変動への追従性向上に効果的であった。

References

1. 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.
2. Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. arXivpreprint arXiv, 1412.(6980).