

等価重み粒子フィルタによる変数・パラメータの事後分布推定

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が可能な、等価重み粒子フィルタ (IEWPF) [1] [2] 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)

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

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

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

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

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

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(1)

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

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

ここで、 $x_i^{n,a}$, 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。

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

式(5)は提案分布 $q(x^n | x_1^{n-1}, y^n) = \frac{q(\xi)}{\| \frac{dx}{d\xi} \|}$ 導入による変数変換 $x_i \rightarrow \xi_i$

を表し、このとき、各粒子の重みが等しくなるために α_i が満たす条件は、

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

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。線形・ガウスの観測・システムノイズ、及び観測モデル H を線形とすると、

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

と変形できる[1]。ただし $\gamma_i = \xi_i^T \xi_i$, 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

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

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 $N_x \rightarrow \infty$)の下で、解析解は Lambert W function で表される [3]。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。

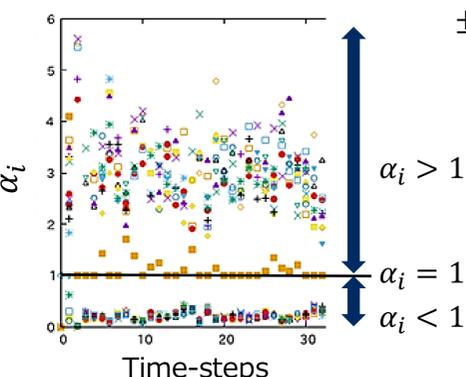


図1. 算出された時系列 α_i の分布の例

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.
- Skauvold, Jacob, et al. "A revised implicit equal - weights particle filter." Quarterly Journal of the Royal Meteorological Society 145.721 (2019): 1490-1502.
- Weisstein, Eric W. "Lambert W-function." <https://mathworld.wolfram.com/> (2002).

2-3. 等価重み粒子フィルタの実現方法(2)

その後、Skauvold et al. によって A revised implicit equal-weights particle filter [2] が提案された。提案分布を与える式は

$$x_i^n = x_i^{n,a} + \beta^{1/2} P^{1/2} \eta_i + \alpha_i^{1/2} P^{1/2} \xi_i, \quad \xi_i \sim N(0, \mathbf{I}), \eta_i \sim N(0, \mathbf{I}), \quad (10)$$

と表され、粒子に依存しない係数 β の擾乱項が追加された。これは、式(9)を満たす α_i を $\alpha \leq 1$ のみとした結果として、分散が最適提案分布より狭くなることを補償するためである。 $\alpha \leq 1$ の理由は、式(8)で $\frac{d(\alpha_i \gamma_i)}{d\gamma_i} > 0$ が仮定され、かつ $c > 0$ であるため、式(9)の不完全ガンマ関数 (lower) の性質からである。

3. 数値実験: 線形モデルを用いた事後分布の評価

3-1. モデルと理論値

推定された変数とパラメータの事後分布の妥当性を評価するために、式(11)で表される線形モデル F を用いた双子実験により理論値と比較をおこなった。

$$\begin{pmatrix} x_j^n \\ \theta_k^{n-1} \end{pmatrix} = \begin{pmatrix} F_x & F_{x\theta} \\ 0 & I \end{pmatrix} \begin{pmatrix} x_j^{n-1} \\ \theta_k^{n-2} \end{pmatrix} + \begin{pmatrix} I & F_{x\theta} \\ 0 & I \end{pmatrix} \begin{pmatrix} \beta_j^n \\ \eta_k^{n-1} \end{pmatrix} = F \begin{pmatrix} x_j^{n-1} \\ \theta_k^{n-2} \end{pmatrix} + G \begin{pmatrix} \beta_j^n \\ \eta_k^{n-1} \end{pmatrix}, \quad (11)$$

$$y_j^n = (H \quad 0) \begin{pmatrix} x_j^n \\ \theta_k^{n-1} \end{pmatrix} + w, \quad \beta \sim N(0, Q_\beta), \quad \eta \sim N(0, Q_\eta), w \sim N(0, R).$$

以下では状態変数 $x \in \mathbb{R}^m$, $\theta \in \mathbb{R}^q$ の次元はそれぞれ $m=100$, $q=3$ とした。観測は $H = I$ としたため、 $y \in \mathbb{R}^r$ の次元は $r=m$ である。線形モデル F は $F_x = I$, パラメータの変数への影響を表す $F_{x\theta} \in \mathbb{R}^{m \times q}$ の要素は I ($j \leq q$), 0 ($j > q$), すなわちパラメータの次元 q と等しい要素までは線形に加算され q より大きい要素はパラメータ無しとなる。式(11)は行列 F, G が時間非依存であるため、カルマンフィルタの定常状態として、以下の離散時間代数 Riccati 方程式を解くことで、分散共分散行列 P を求めることができる。

$$P = FPF^T - FPH^T(HPH^T + R)^{-1}HPF^T + \tilde{Q}, \quad (12)$$

ここで、 \tilde{Q} は式(4)のモデル f を線形モデル F に置き換えたものである。以下では、変数のシステムノイズと観測ノイズの分散共分散行列は対角とし、疑似観測データ生成時と推定時で同じ値 $Q_\beta=0.04$, $R=0.06$, パラメータは推定時に $Q_\theta=0.04$, 粒子数 $N_p=20$, 観測データは全時間ステップ有とした。

3-2. 事後分布分散の比較・パラメータ推定の評価

式(5)及び式(10)により推定された変数とパラメータの事後分布の分散を、それぞれ式(12)を解いた理論値(KF)と比較した結果を示す。図2, 3は以下の式(13)により求められる時間ステップ n における変数とパラメータの分散の平均値 $\overline{s_x^2}, \overline{s_\theta^2}$ の時間ステップ100~2000における値をそれぞれヒストグラムで表した。

$$\overline{s_x^2} = \frac{1}{q} \frac{1}{N_p} \sum_{j=1}^q \sum_{i=1}^{N_p} (x_i^n - \overline{x_i^n})_j^2, \quad \overline{s_\theta^2} = \frac{1}{q} \frac{1}{N_p} \sum_{k=1}^q \sum_{i=1}^{N_p} (\theta_i^n - \overline{\theta_i^n})_k^2. \quad (13)$$

図2から式(5)によるパラメータの推定(右図)は、 α を $\alpha \leq 1$ と $\alpha > 1$ から50%ずつサンプリングした結果(青)が理論値(KF)に最も近い。図3から式(10)による推定では係数 $\beta = 0.5$ (緑)が変数・パラメータいずれも理論値に最も近く、式(5)と式(10)の比較から式(10)の方が分散値のバラツキが少ないことがわかる。

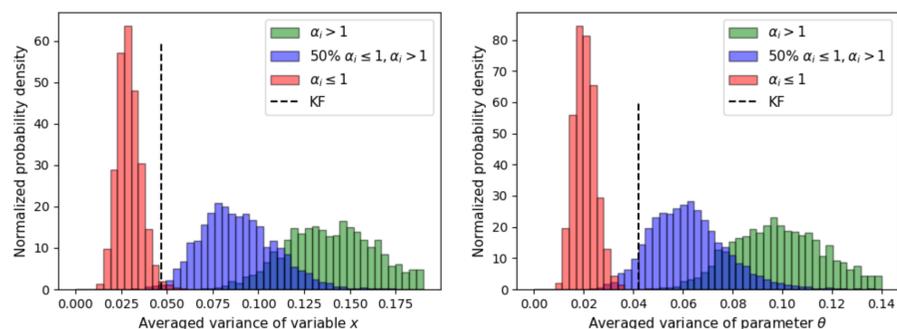


図2. 式(5)の係数 α 依存性比較 (左: 変数, 右: パラメータ)

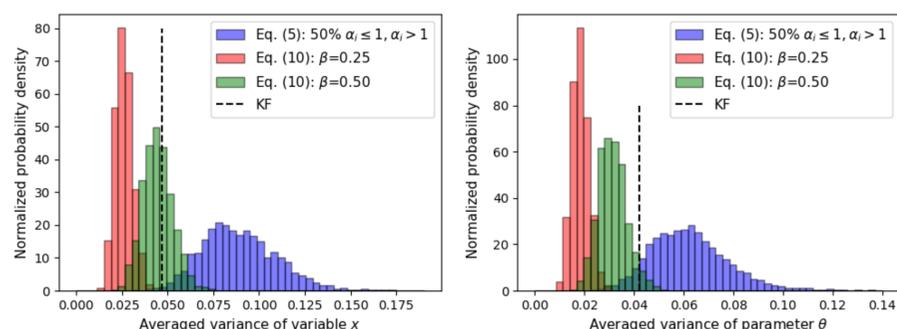


図3. 式(5) 50% α と式(10)の係数 β 依存性の比較 (左: 変数, 右: パラメータ)