

データ同化の数学解析について

竹田航太

京都大学 理学研究科 数学教室/理研計算科学研究センター

2023/03/11



自己紹介

竹田航太

京都大学 理学研究科 数学教室 坂上研究室
D1

理化学研究所 計算科学研究センター データ
同化研究チーム JRA

個人サイト



研究内容

Uncertainty Quantification —————> 数理解体力学

数理モデルの誤差・不確実性を
定量的に評価.

流体現象を数学的に研究.

キーワード

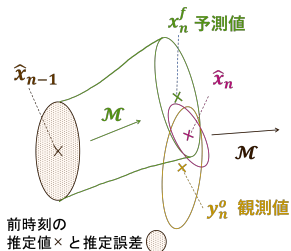
UQ, 流体力学, データ同化, 乱流, モンテカルロ, MCMC, ベイズ, 逆問題, 気象, 一般化勾配法, トポロジカルデータ解析

データ同化とは

Filtering(Smoothing) 問題の気象学・海洋学界限での方言であり、シミュレーションと観測データを融合させる理論。予測に使うための精度の良い初期値の推定，モデルのパラメータ推定，観測データの平滑化や補間などを目的とする。

気象学・海洋学に限らず宇宙科学や生命科学など応用範囲は広い。統計的なアプローチから理論が作られているため，数学的な解析がまだ少ない。

データ同化のイメージ



設定 | 状態空間モデル (離散時間)

状態空間を \mathbb{R}^m として. 関数 $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}^m$ に対して, 以下の離散力学系を考える.

$$u_j = \Psi(u_{j-1}) \quad (1)$$

初期値 $u_0^\dagger \in \mathbb{R}^m$ に対する時間発展 $(u_j^\dagger)_{j \geq 0}$ が存在するが直接観測できないとする.

系に対する情報はノイズを含む有限次元 $d \in \mathbb{N}$ の観測 $y_j \in \mathbb{R}^d$ として得られる.

$$y_j = \mathcal{H}(u_j^\dagger) + \eta_j, \quad \eta_j \sim N(0, \Gamma). \quad (2)$$

ただし, 観測関数を $\mathcal{H} : \mathbb{R}^m \rightarrow \mathbb{R}^d$ とし, 各 $\eta_j \in \mathbb{R}^d$ は互いに独立で $\Gamma \in \mathbb{R}^{d \times d}$ は半正定値とする.

設定 | データ同化の目標

問題

$J \in \mathbb{N}$ 個の観測値 $Y_J = (y_j)_{j=1}^J$ からどのようにして真の状態 $(u_j^\dagger)_{j=1}^J$ に対する推定値を得るか？

使える情報は

時間発展 $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}^m$.

観測関数 $\mathcal{H} : \mathbb{R}^m \rightarrow \mathbb{R}^d$ とそのノイズの情報 $N(0, \Gamma)$.

適当に $\hat{u}_0 \in \mathbb{R}^m$ と誤差 $\Sigma_0 : \mathbb{R}^{m \times m}$ を取り, 初期値に関する不確実性 $\mu_0(du_0^\dagger) = N(du_0^\dagger; \hat{u}_0, \Sigma_0)$ を設定すると $j = 1, \dots, J$ に対して,

不確実性 $\mu_j(du_j^\dagger)$ をどのように計算できるか？

また, 推定値 \hat{u}_j をどのように計算できるか？

という問題になる.

方法 | ベイズに基づいたデータ同化

帰納的に $\mu_j(du_j)$, $j = 1, \dots, J$ を構成する.

予測 $\mu_{j-1}(du_{j-1})$ に対して, モデル Ψ による時間発展で時刻 j での不確実性 $\hat{\mu}_j(du_j)$ を推定する.

$$\hat{\mu}_j(du_j) = \mu_{j-1}(\Psi^{-1}(du_j)). \quad (3)$$

ベイズ更新 ベイズの公式 $P(u|y) = \frac{P(y|u)P(u)}{P(y)}$ から, 事前分布 $\hat{\mu}_j(du_j)$ と尤度 $p(y|u) = N(y; \mathcal{H}(u), \Gamma)$ に対して事後分布として $\mu_j(du_j)$ を計算する.

$$\mu_j(du_j) = \frac{p(y_j|u_j)\hat{\mu}_j(du_j)}{\int p(y_j|u_j)\hat{\mu}_j(du_j)}. \quad (4)$$

Point

モデルによる予測 (3) とベイズ更新 (4) を繰り返す.

例 | Kalman Filter(線形, Gauss ノイズ)

$m = d = 1$, モデルは $\Psi(u) = \lambda u$, $\lambda \in \mathbb{R}$. 観測は $\mathcal{H}(u) = u$, $\Gamma = \gamma^2$. $\mu_0(du) = N(du; m_0, \sigma_0^2)$, $m_0 \in \mathbb{R}$, $\sigma_0 > 0$ とする.
(推定値は推定分布の平均値 $\hat{u}_0 = m_0$ とする.)

このとき, 推定分布の列 $(\mu_j(du))_{j=1}^J$ は Gaussian $N(m_j, \sigma_j^2)$ となり以下のように予測とベイズ更新で帰納的に構成できる.

$$\hat{m}_j = \lambda m_{j-1}, \quad \hat{\sigma}_j^2 = \lambda^2 \sigma_{j-1}^2 \quad (5)$$

$$\frac{m_j}{\sigma_j^2} = \frac{\hat{m}_j}{\hat{\sigma}_j^2} + \frac{y_j}{\gamma^2}, \quad \frac{1}{\sigma_j^2} = \frac{1}{\hat{\sigma}_j^2} + \frac{1}{\gamma^2} \quad (6)$$

(Ψ と \mathcal{H} の線形性と η_m の Gauss 性による.)

例 | Particle Filter

分布を N 点 $(u_j^{(n)})_{n=1}^N$ ¹ のヒストグラムで近似する.

$$\mu_0^N \sim \frac{1}{N} \sum_{n=1}^N \delta_{u_0^{(n)}}, \quad u_0^{(n)} \sim \mu \quad \text{i.i.d..}$$

予測のステップでは各点をモデルで時間発展する.

$$\hat{\mu}_j^N \sim \frac{1}{N} \sum_{n=1}^N \delta_{\hat{u}_j^{(n)}}, \quad \hat{u}_j^{(n)} = \Psi(u_{j-1}^{(n)}). \quad (7)$$

ベイズ更新では, ベイズの公式から各点の重みを計算する.

$$\mu_j^N = \frac{1}{N} \sum_{n=1}^N w^{(n)} \delta_{\hat{u}_j^{(n)}}, \quad w^{(n)} = \frac{p(y_j | \hat{u}_j^{(n)})}{\sum_{n=1}^N p(y_j | \hat{u}_j^{(n)})} \quad (8)$$

最後に, μ_j^N に従う等しい重みの ensemble $(u_j^{(n)})_{n=1}^N$ を取り直す.

$$\mu_j^N \sim \frac{1}{N} \sum_{n=1}^N \delta_{u_j^{(n)}}, \quad u_j^{(n)} \sim \mu \quad \text{i.i.d..} \quad (9)$$

¹点の集まりをアンサンブル (ensemble) と言う場合もある.

より一般的な設定

状態空間

関数空間（無限次元 Hilbert 空間）をとることもある。

モデル

モデル Ψ は線形とは限らない。ノイズを含む力学系を考えることもある。非自励系を考える必要があるかもしれない。また、実際の現象がどんな dynamics に従うかを知ることは不可能...

観測

観測は線形とは限らない。観測ノイズは Gaussian とは限らない。全空間で観測値が得られるわけではない。観測関数が時間に依存するかもしれない...

時間

連続時間を考えることもある。（モデルだけの場合や観測も連続時間などいろいろ。）

いろいろなデータ同化アルゴリズム

Kalman Filter

線形モデル・線形観測, Gauss ノイズを仮定.

Ensemble Kalman Filter

Kalman Filter をモンテカルロ近似し, 推定の期待値と共分散を計算する. 非線型なモデルにも適用できるが3次モーメント以上は計算しない.

Particle Filter

分布をヒストグラムで近似. モデルや観測などに仮定がない.

(変分法)

予測値と観測値の両方にフィットするように最小2乗推定を行う. 分布は推定せず事後確率最大化点のみ計算する. 3次元変分法, 4次元変分法など.

数学的に示すべきこと I

$d(\cdot, \cdot)$ は確率測度間の適当な距離を表すものとする。上の方が基礎的、下に行くほど応用上で必要とされる評価。

A. 適切性 (Wellposedness)

事後分布推定問題における解の存在と一意性とパラメータ (観測値) に対する連続性。

$$d(\mu^y, \mu^{y'}) \leq C|y - y'|. \quad (10)$$

B. 安定性 (Stability)

推定誤差が時間無限大で有界か？

C. 事後分布の一貫性 (Posterior consistency)

十分な (abundant and accurate) 観測値が得られれば、真の状態 u^\dagger を recover できること。

$$d(\mu_j^{(Y_J, \Gamma)}, \delta_{u_j^\dagger}) \rightarrow 0 \quad (J \rightarrow \infty, tr(\Gamma) \rightarrow 0). \quad (11)$$

数学的に示すべきこと II

D. 分布近似誤差評価 (Bayesian Quality Assessment)

理論的な事後分布 μ のアルゴリズムによる近似 μ_{approx} がどれくらい近いかを評価すること. (離散化パラメータ連続性とも言える)

$$d(\mu_{approx}, \mu) \leq ?. \quad (12)$$

E. 信号推定誤差評価 (Signal Quality Assessment/Accuracy)

信号推定 (点推定) 問題において推定値 \hat{u} と真の状態 u^\dagger がどれくらい近いかを評価すること. 各時刻での誤差や経路としての誤差を評価する.

$$\limsup_{j \rightarrow \infty} \mathbb{E}[|\hat{u}_j - u_j^\dagger|^2] \leq ?. \quad (13)$$

数学的に示すべきこと III

F. エルゴード性 (Ergodicity)

Ensemble Filter の推定分布が一意的な invariant measure に収束するか，それがどれくらい速いかを評価する．

$$d(P^n \mu, P^n \nu) \leq C_{\mu, \nu} \lambda^n \quad (14)$$

(G. モデルの解析)

モデルの誤差発達の評価がデータ同化の解析に必要．そのために，モデル方程式の解や global attractor の存在などについても調べる必要がある．

データ同化ステップの書き換え

解析の見通しをよくするために分布の変化を作用素的に書く。

予測 $P : \mathcal{M}_1(\mathbb{R}^m) \rightarrow \mathcal{M}_1(\mathbb{R}^m)$ を以下で定め²,

$$(P\mu)(du) = \mu(\Psi^{-1}(A)) = \int_{\mathbb{R}^m} \delta(u - \Psi(v)) du \mu(dv), \quad (15)$$

ベイズ更新 $y_j \in \mathbb{R}^d$ に対して, $L_j : \mathcal{M}_1(\mathbb{R}^m) \rightarrow \mathcal{M}_1(\mathbb{R}^m)$ を

$$(L_j\mu)(du) = \frac{p(y_j|u)\mu(du)}{\int p(y_j|u)\mu(du)} \quad (16)$$

とおくと, データ同化での分布の更新は以下のように書ける。

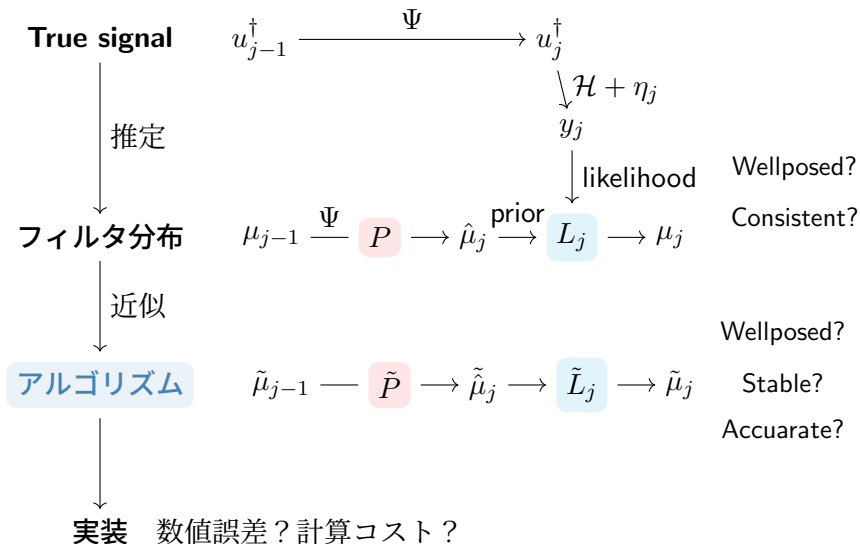
$$\mu_j = L_j P \mu_{j-1}. \quad (17)$$

時刻に依存した Markov 遷移 $\mu_{j-1} \rightarrow \mu_j$ になっている³。

²stochastic なモデルの場合は P がマルコフ遷移となる。

³真の状態 u_j も合わせると時間に依存しない Markov 遷移で表せる。

データ同化における推定・近似の階層



Particle Filter の近似誤差 [1, Law-Stuart-Zygalakis 2015]

分布の N 点のヒストグラムによる近似を

$S^N : \mathcal{M}_1(\mathbb{R}^m) \rightarrow \mathcal{M}_1(\mathbb{R}^m)$ として以下のように書く.

$$S^N \mu = \frac{1}{N} \sum_{n=1}^N \delta_{u^{(n)}}, \quad u^{(n)} \sim \mu \quad \text{i.i.d.} \quad (18)$$

Particle Filter による近似分布 $(\mu_j^N)_{j=1}^J$ は以下のように構成される.

$$\mu_j^N = L_j P S^N \mu_{j-1}^N, \quad \mu_0^N = \mu_0. \quad (19)$$

尤度に関する適切な条件のもとで以下が成り立つ. (評価 D)

$$d(\mu_J^N, \mu_J) \leq \exists C_J \frac{1}{\sqrt{N}}. \quad (20)$$

Kalman Filter についても Gaussian への射影を使って同様の記述ができるが, 非 Gaussian な分布に収束することはないので分布の近似誤差に関する解析は知られていない.

モデルの定式化の例 [2, Kelly-Law-Stuart2014]

Hilbert 空間 $(\mathcal{H}, |\cdot|)$ 上の力学系を考える.

$$\frac{du}{dt} + \mathcal{A}(u) + \mathcal{B}(u, u) = f. \quad (21)$$

ただし, 線形項 \mathcal{A} と非線形項 \mathcal{B} には適切な仮定を設ける.

(21) の解が 1-パラメータ半群 $\Psi_t: \mathcal{V} \rightarrow \mathcal{V}$ で書けるとすると, 時間刻み $h > 0$ に対して $\Psi = \Psi_h$ とおくことで離散時間の設定に帰着できる.

[2, Kelly-Law-Stuart2014] は Lorenz63, 96, トーラス上 2 次元 Navier-Stokes が含まれるように仮定を設定し, 誤差発達などについての評価を与えた. (評価 G)

$$|\Psi_t(v_0) - \Psi_t(w_0)| \leq e^{\exists\beta t} |v_0 - w_0|. \quad (22)$$

$\beta > 0$ のとき, 誤差が指数的に発達しうる. \rightarrow カオス性.

既存の結果

ベイズ推定の観測値連続性 (A)

事後分布は観測値に対して局所リプシッツ⁴ [1, Law-Stuart-Zygalakis 2015], [3, Sullivan 2015].

3次元変分法 (E)

線形部分観測 ($d < m$) の場合でも, 推定誤差が観測ノイズで評価. Lorenz63 や Lorenz96 を対象 [1, Law-Stuart-Zygalakis 2015], [4, Law-Alonso-Shukla-Stuart2016].

Ensemble Kalman Filter (B, F)

PO 法や ETKF に対する線形部分観測の下で安定性やエルゴード性 [2, Kelly-Law-Stuart 2014], [5, Tong-Majda-Kelly 2016].

Ensemble Kalman Bucy Filter⁵ (A, B, E)

完全観測 small noise の設定での解の存在証明や誤差解析 [6, Wiljes-Reich-Stannat 2018]. また, Localization という手法でサンプルサイズを減らす [7, Wiljes-Tong 2020].

⁴Hellinder 距離

⁵連続時間版の Ensemble Kalman Filter

その他

解析手法など

Gronwall の不等式, エネルギー不等式, リヤプノフ関数, エルゴード理論, マルコフ連鎖, SDE, 行列の固有値/スペクトル解析, Random Matrix...

課題

観測の非線形性, 非 Gauss 分布の近似. 4次元変分法は解析解がないので評価が難しい. ad hoc な方法を減らしたい. 計算コスト (速度, メモリ). 大規模システムとの親和性.

Reference I

- [1] K. J. H. Law, A. M. Stuart, and K. C. Zygalakis.
Data Assimilation: A Mathematical Introduction.
Springer, 2015.
- [2] D. T. B. Kelly, K. J. H. Law, and A. M. Stuart.
Well-posedness and accuracy of the ensemble kalman filter in discrete and continuous time.
NONLINEARITY, 27(10):2579–2603, OCT 2014.
- [3] Timothy John Sullivan.
Introduction to uncertainty quantification, volume 63 of *Texts in Applied Mathematics*.
Springer, 2015.
- [4] K. J. H. Law, D. Sanz-Alonso, A. Shukla, and A. M. Stuart.
Filter accuracy for the lorenz 96 model: Fixed versus adaptive observation operators.
PHYSICA D-NONLINEAR PHENOMENA, 325:1–13, JUN 15 2016.
- [5] Xin T Tong, Andrew J Majda, and David Kelly.
Nonlinear stability and ergodicity of ensemble based kalman filters.
Nonlinearity, 29(2):657, jan 2016.

Reference II

- [6] Jana de Wiljes, Sebastian Reich, and Wilhelm Stannat.
Long-time stability and accuracy of the ensemble kalman–bucy filter for fully observed processes and small measurement noise.
SIAM Journal on Applied Dynamical Systems, 17(2):1152–1181, 2018.
- [7] Jana de Wiljes and Xin T Tong.
Analysis of a localised nonlinear ensemble kalman bucy filter with complete and accurate observations.
Nonlinearity, 33(9):4752–4782, jul 2020.
- [8] 知之 樋口, 玄太 上野, 慎也 中野, 和幸 中村, and 亮 吉田.
データ同化入門 : 次世代のシミュレーション技術.
Number 6 in シリーズ予測と発見の科学. 朝倉書店, 2011.
- [9] Kody Law, Abhishek Shukla, and Andrew Stuart.
Analysis of the 3dvar filter for the partially observed lorenz'63 model.
DISCRETE AND CONTINUOUS DYNAMICAL SYSTEMS, 34(3, SI):1061–1078, MAR 2014.
- [10] Marco A Iglesias, Kody J H Law, and Andrew M Stuart.
Ensemble kalman methods for inverse problems.
Inverse Problems, 29(4):045001, mar 2013.

Reference III

- [11] Kevin Hayden, Eric Olson, and Edriss S. Titi.
Discrete data assimilation in the lorenz and 2d navier-stokes equations.
PHYSICA D-NONLINEAR PHENOMENA, 240(18):1416–1425, SEP 1 2011.
- [12] C. Foias, D. D. Holm, and E. S. Titi.
The three dimensional viscous camassa-holm equations, and their relation to the navier-stokes equations and turbulence theory.
Journal of Dynamics and Differential Equations, 14(1), 2001.
- [13] David Kelly, Andrew J. Majda, and Xin T. Tong.
Concrete ensemble kalman filters with rigorous catastrophic filter divergence.
PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, 112(34):10589–10594, AUG 25 2015.
- [14] David Kelly and Andrew M. Stuart.
Ergodicity and accuracy of optimal particle filters for bayesian data assimilation.
CHINESE ANNALS OF MATHEMATICS SERIES B, 40(5):811–842, SEP 2019.
- [15] Andrew J. Majda and Xin T. Tong.
Rigorous accuracy and robustness analysis for two-scale reduced random kalman filters in high dimensions, 2016.