

博士論文

散逸力学系に対するアンサンブル平方根フィルターの 誤差解析

京都大学数学教室

竹田航太 *

指導教員

京都大学 坂上 貴之 教授

理化学研究所 三好 建正 チームリーダー

ポツダム大学 Sebastian Reich 教授

*京都府 京都市 左京区 北白川追分町 京都大学 数学教室, takeda.kota.53r@st.kyoto-u.ac.jp

目 次

1 導入	5
2 準備	7
2.1 記法	7
2.2 関数解析	7
2.2.1 基礎理論	7
2.2.2 コンパクト作用素	9
2.2.3 摂動論	10
2.3 線形作用素に対する計算	10
2.3.1 ベクトルのアンサンブル	10
2.3.2 逆作用素	12
2.4 確率論	14
2.5 ベイズ推定	16
2.5.1 ベイズの定理	17
2.5.2 事後分布の適切性	18
2.6 力学系	20
3 状態空間モデルとデータ同化	23
3.1 有限次元問題	23
3.1.1 状態空間モデル	23
3.1.2 ベイズ的データ同化問題	24
3.2 無限次元問題	28
3.2.1 離散時間状態空間モデル	28
3.2.2 連続時間状態空間モデル	29
4 フィルタリングアルゴリズム	32
4.1 カルマンフィルタ	32
4.2 非線形力学系への拡張	34
4.3 アンサンブルカルマンフィルタ	35
4.4 EnKF に対する数値テクニック	41
4.5 連続時間アルゴリズム	44

5 散逸力学系	45
5.1 散逸力学系とその例	45
5.1.1 有限次元系の例	46
5.1.2 無限次元系の例	47
5.2 部分観測からの状態の再構成	49
5.2.1 Navier-Stokes 方程式に対する連続データ同化	49
5.2.2 有限次元系に対する連続データ同化	51
5.2.3 離散データ同化	52
5.3 確率的ダイナミクス	53
6 EnKF の数学解析	55
6.1 EnKF の基本的な性質	55
6.2 フィルタリングアルゴリズムの誤差解析	56
6.2.1 3DVar	57
6.2.2 PO 法	59
7 ESRF の誤差解析	62
7.1 ESRF の適切性	62
7.2 ESRF の時間大域的な誤差評価	66
7.3 数値計算	71
8 まとめと今後の展望	77
参考文献	78

更新履歴

- 2025 年 2 月 1 日：第一版.
- 2026 年 1 月 1 日：第二版. 不要な補題を削除.

1. 導入

現代の気象予測において、大気モデルの数値シミュレーションは極めて重要な役割を果たしている。しかしながら、初期状態が十分な解像度で推定された場合であっても、大気のカオス的特性により、シミュレーションの値は時間の経過とともに真の状態から乖離することがある。すなわち、初期状態に加えられた微小な摂動が、決定論的な時間発展の過程で指数関数的に成長しうることが知られている [50, 66]。この特性は、大気運動の長期的予測の困難性を引き起こす要因の一つである。したがって、シミュレーションの値を観測データを用いて修正することが必要となる。このようなアプローチはデータ同化 (data assimilation) として知られている。気象学および海洋学におけるデータ同化の動機や課題については、[34, 50] を参照されたい。

シミュレーションの値を補正する単純な方法として、それを観測データで置き換える手法が考えられる。しかしながら、この方法は以下の 2 つの理由により現実的ではない。第一に、観測データは数値シミュレーションを初期化するには解像度が不十分である [28]。したがって、低解像度の観測データから高解像度の状態を再構築する必要がある。第二に、観測データには常に測定ノイズが含まれている。そのため、ノイズを含む観測データから状態を推定する際の不確実性を定量化する必要がある [80]。以上の理由から、状態推定はしばしばベイズ逆問題として定式化される [26, 76]。このアプローチでは、推定は確率分布として表現される。さらに、数値モデルが事前分布を与え、それは観測データとベイズの定理を用いて事後分布へと更新される。データ同化をベイズ逆問題として数学的定式化については、[63, 76, 78] の標準的な教科書を参照されたい。

モンテカルロ法は、ベイズ逆問題における基本的なアプローチであり、無限に多くのサンプルを用いることで事後分布を復元する手法である [63, 76]。しかしながら、高次元の状態空間において事後分布全体を評価することは、計算コストの観点から現実的ではない。この課題を回避するための一つのアプローチとして、事後確率分布の最大化点である最大事後確率推定 (MAP 推定) を求める方法がある。これはデータ同化の文脈において変分法と呼ばれている [63, 76]。もう一つのアプローチは、事後分布の平均値と共に分散のみを評価する方法である。これらの 2 つのアプローチは、状態の発展と観測作用素が線形写像で表され、全てのノイズがガウス分布に従う場合、最小二乗法と等価となる。この場合に対応する逐次データ同化アルゴリズムがカルマンフィルタ (KF) である [4, 24, 49]。しかしながら、例えば地球流体の運動は通常、非線形ダイナミクスでモデル化される。このような場合には、カルマンフィルタの非線形拡張としてアンサンブルカルマンフィルタ (EnKF) が提案されている [34]。EnKF は、アンサンブルと呼ばれる有限個の状態ベクトルの組を用いて平均値と共に分散を近似する。EnKF は、モンテカルロ法と最小二乗法のハイブリッドアプローチであり、少数のアンサンブルメンバーを用い

て状態推定における不確実性を表現する。この手法は、地球物理学における高次元かつ非線形システムのデータ同化問題に幅広く適用されている [20, 34, 50]。

アンサンブルカルマンフィルタ (EnKF) には、主に 2 つの実装が知られている。確率論的手法である観測摂動法 (perturbed observation method, PO 法) [19, 33] と、決定論的手法であるアンサンブル平方根フィルタ (ensemble square root filter, ESRF) [5, 14, 86] である。数学解析により、PO 法と ESRF に共通するいくつかの基本的性質が明らかにされている [51, 58, 70, 88, 89]。例えば、アンサンブルの有界性や、線形システムかつガウスノイズの場合における大アンサンブルの極限でカルマンフィルタ (KF) に収束する性質が挙げられる。しかしながら、誤差解析に関しては、PO 法についてのみ得られており [54]、ESRF については未解明のままである。したがって、EnKF の理論的解析を発展させるためには、ESRF に関する誤差解析を行うことが必要である。

データ同化の数学解析におけるもう一つの課題は、モデルダイナミクスの定式化である。地球流体の運動は、しばしばナビエ-ストークス方程式のような散逸的な偏微分方程式 (PDE) によってモデル化される [23, 38]。これらの方程式は無限次元の状態空間上で定義される。さらに、地球流体のカオス性は、気象予測において本質的な研究対象である。モデルダイナミクスにこれらの性質を取り込むために、ヒルベルト空間上の散逸力学系を考える。典型的な解は有界な軌道を持ち、グローバルアトラクタと呼ばれるコンパクトな不变集合上でカオス的挙動を示す。モデルダイナミクスは無限次元空間で定式化され、数値計算のために離散化されることで有限次元系となる [27]。

これまでに、EnKF の数学解析における 2 つの課題を紹介した。これに基づき、本論文は以下の 3 つのテーマを議論することを目的とする。第一に、モデルダイナミクスのさまざまな定式化における EnKF の解析結果を概観する。第二に、散逸力学系に適用される ESRF の誤差解析を証明する。これが本論文の主要な貢献である。最後に、さまざまな定式化における結果を比較しつつ、EnKF の数学解析における将来的な研究の方向性を示す。

本論文の構成は以下の通りである。節 2 では、記法を導入し、関数解析、測度論、ベイズ推論、およびヒルベルト空間上の力学系の理論を概観する。節 3 では、データ同化問題をさまざまな定式化に基づいて定義する。これらは、確率論的／決定論的、有限次元／無限次元、および離散時間／連続時間のシステムによって分類される。節 4 では、逐次データ同化アルゴリズムである KF と、EnKF の各手法を概観する。また、これらを実装の観点から検討する。節 5 では、散逸力学系の解析と、地球流体に関連するいくつかの具体例を示す。節 6 では、EnKF の数学解析を説明する。ここでは、PO 法の誤差解析および EnKF の基本的な解析結果を概観する。節 7 では、本論文の主たる結果を説明する。散逸力学系に対する ESRF の誤差評価を与えるとともに、数値例を用いてその解析を検証する。節 8 では、本論文の結果をまとめ、将来の研究方向について議論する。

2. 準備

2.1 記法

大文字（例: U ）は確率変数を表し、小文字（例: u ）はその実現値または決定論的変数を表すものとする。自然数 $n \in \mathbb{N}$ に対し、 $u \in \mathbb{R}^n$ は列ベクトルであると仮定する。ベクトル $u \in \mathbb{R}^n$ に対し、 u^i はその i 番目の要素を表し、 u^* は u の転置を表すものとする。また、太字（例: U や u ）はベクトルの組を表すものとする。

2.2 関数解析

ヒルベルト空間上での状態推定問題を扱うため、関数解析の理論に関するいくつかの基本事項を振り返る。この節における詳細および証明については、入門的な教科書 [25] を参照されたい。

2.2.1 基礎理論

内積 $\langle \cdot, \cdot \rangle$ とそれに対応するノルム $|\cdot|$ を備えたヒルベルト空間を \mathcal{H} とする。ヒルベルト空間 \mathcal{H} がユークリッド空間の場合も同じ記法を用いる。ここで、ヒルベルト空間は可分であると仮定し、 $I_{\mathcal{H}}$ を \mathcal{H} 上の恒等作用素とする。別のヒルベルト空間 \mathcal{G} に対し、 $\mathcal{L}(\mathcal{H}, \mathcal{G})$ は、 \mathcal{H} から \mathcal{G} への有界線形作用素の空間を表す。作用素 $A \in \mathcal{L}(\mathcal{H}) := \mathcal{L}(\mathcal{H}, \mathcal{H})$ に対し、 $|A|_{\mathcal{L}}$ は A の作用素ノルム、 $\text{Ran}(A)$ は A の像、 A^* は A の随伴作用素を表す。ヒルベルト空間の元 $u, v \in \mathcal{H}$ に対し、それらの積 $u \otimes v \in \mathcal{L}(\mathcal{H})$ を $u \otimes v : \mathcal{H} \ni w \mapsto u \langle v, w \rangle \in \mathcal{H}$ によって定義する。これは $uv^* = u \otimes v$ と等価である。作用素 $U \in \mathcal{L}(\mathcal{H})$ がユニタリであるとは、 $U^*U = UU^* = I_{\mathcal{H}}$ を満たすことをいう。 $A^* = A$ を満たす $\mathcal{L}(\mathcal{H})$ の自己共役作用素の集合を $\mathcal{L}_{sa}(\mathcal{H})$ と書く。以下では、 $\mathcal{L}_{sa}(\mathcal{H})$ における重要な概念を定義する。

定義 2.1. $A \in \mathcal{L}_{sa}(\mathcal{H})$ とする。

- (a) A が半正定値であるとは、任意の $u \in \mathcal{H}$ に対して $\langle u, Au \rangle \geq 0$ を満たすときにいい、これを $A \succeq 0$ と表す。
- (b) A が正定値であるとは、ある定数 $c > 0$ が存在して、任意の $u \in \mathcal{H}$ に対して $\langle u, Au \rangle \geq c|u|^2$ を満たすときにいい、これを $A \succ 0$ と表す。

注意 2.2. 定義 2.1 の (b) は、正定値性の通常の定義とは異なる点に注意が必要である。通常、(b) は「下に有界」と呼ばれ、それから A が可逆であることが従う。

半正定値 $A \in \mathcal{L}_{sa}(\mathcal{H})$ に対して, 平方根 $A^{\frac{1}{2}} \succeq 0$ は一意的に定義される [25]. したがって, $A \succ 0$ のとき, \mathcal{H} 上に重み付きノルム $|\cdot|_A = |A^{-1/2} \cdot|$ を定義する. また, $A, B \in \mathcal{L}_{sa}(\mathcal{H})$ に対して, 順序 $A \succ$ (または \succeq) B は $A - B \succ$ (または \succeq) 0 を意味するものとする. 作用素ノルムを評価するために次の不等式を用いることができる.

補題 2.3 ([25]). $A \succeq B$ のとき, $|A|_{\mathcal{L}} \geq |B|_{\mathcal{L}}$ が成り立つ.

線形作用素 $A : \mathcal{H} \rightarrow \mathcal{H}$ に対して, そのスペクトルを次のように表す.

$$\sigma(A) = \{\lambda \in \mathbb{C} \mid (\lambda I - A)^{-1} \notin \mathcal{L}(\mathcal{H})\}.$$

また, レゾルベントを次のように定義する.

$$\rho(A) = \mathbb{C} \setminus \sigma(A).$$

自己共役作用素 $A \in \mathcal{L}_{sa}(\mathcal{H})$ について, $\sigma(A) \subset \mathbb{R}$ が成り立つ. さらに, $A \succeq 0$ の場合は $\sigma(A) \subset [0, \infty)$ となる. 作用素 A のスペクトル半径を次のように定義する.

$$r(A) = \sup_{\lambda \in \sigma(A)} |\lambda|.$$

一般に, $r(A) \leq |A|_{\mathcal{L}}$ が成り立つ, すなわち, 任意の $\lambda \in \sigma(A)$ に対して $|\lambda| \leq |A|_{\mathcal{L}}$ である. したがって, 逆を取ることで $|\lambda| > |A|_{\mathcal{L}}$ ならば $\lambda \in \rho(A)$ が成り立つ. 正規作用素に対して, 以下の事実はよく知られている.

命題 2.4 ([25]). $A \in \mathcal{L}(\mathcal{H})$ が正規作用素, すなわち $AA^* = A^*A$ を満たす場合, 次が成り立つ.

$$r(A) = |A|_{\mathcal{L}}.$$

なお, A が自己共役またはユニタリである場合, A は正規作用素である.

次に, 自己共役作用素の積のスペクトルに関する評価について述べる.

命題 2.5 ([43]). $A, B \in \mathcal{L}_{sa}(\mathcal{H})$ かつ $B \succeq 0$ とする. このとき, 以下が成り立つ.

$$(1) \quad \sigma(AB) = \sigma(BA) = \sigma(B^{\frac{1}{2}}AB^{\frac{1}{2}}).$$

$$(2) \quad A \succeq 0 \text{ の場合, } \sigma(AB) \subset [m(A)m(B), M(A)M(B)] \text{ となる. ただし, } m(A) = \inf \sigma(A), M(A) = \sup \sigma(A).$$

2.2.2 コンパクト作用素

線形作用素 $K : \mathcal{H} \rightarrow \mathcal{G}$ がコンパクトであるとは、有界列 $(u_n)_{n \in \mathbb{N}} \subset \mathcal{H}$ に対して、列 $(Ku_n)_{n \in \mathbb{N}} \subset \mathcal{G}$ が収束部分列を含むときという。コンパクト作用素の空間を $\mathcal{K}(\mathcal{H}, \mathcal{G})$ 、および $\mathcal{K}(\mathcal{H}) = \mathcal{K}(\mathcal{H}, \mathcal{H})$ と表す。無限次元空間上のコンパクト作用素を考える際に次の事実が重要となる。

命題 2.6. $\dim(\mathcal{H}) = \infty$ とする。 $K \in \mathcal{K}(\mathcal{H})$ ならば $K^{-1} \notin \mathcal{L}(\mathcal{H})$ が成り立つ。

この命題から、 $A, A^{-1} \in \mathcal{L}(\mathcal{H})$ ならば $A \notin \mathcal{K}(\mathcal{H})$ が導かれる。例えば、 $I_{\mathcal{H}} \notin \mathcal{K}(\mathcal{H})$ となる。自己共役かつコンパクトな作用素はユニタリ対角化可能であり、次のようなスペクトル分解を持つ。

命題 2.7 (スペクトル定理). $K \in \mathcal{L}_{sa}(\mathcal{H}) \cap \mathcal{K}(\mathcal{H})$ とする。このとき、固有値 $(\lambda_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ と、それに対応する正規直交基底 $(\phi_n)_{n \in \mathbb{N}} \subset \mathcal{H}$ が存在し、次が成り立つ。

$$K = \sum_{n \in \mathbb{N}} \lambda_n \phi_n \otimes \phi_n.$$

さらに、次の系が得られる。

系 2.8 (特異値分解). $K \in \mathcal{K}(\mathcal{H})$ とする。このとき、 $K^*K \succeq 0$ かつ $K^K \in \mathcal{L}_{sa}(\mathcal{H}) \cap \mathcal{K}(\mathcal{H})$ が成り立つ。したがって、 $(s_n)_{n \in \mathbb{N}} \subset [0, \infty)$ と、正規直交基底 $(\phi_n)_{n \in \mathbb{N}} \subset \mathcal{H}$ が存在し、次が成り立つ。

$$K^*K = \sum_{n \in \mathbb{N}} s_n^2 \phi_n \otimes \phi_n. \quad (2.1)$$

ここで、 $(s_n)_{n \in \mathbb{N}}$ は K の特異値と呼ばれる。

コンパクト作用素の重要なクラスとして、トレースクラス作用素を導入する。

定義 2.9. $T \in \mathcal{K}(\mathcal{H})$ について、次を満たすとき、 T はトレースクラスであるという。

$$\mathrm{Tr}|T| := \sum_{n \in \mathbb{N}} s_n(T) < \infty. \quad (2.2)$$

ここで、 $(s_n(T))_{n \in \mathbb{N}}$ は T の特異値を表す。トレースクラス作用素の集合を $\mathcal{K}_1(\mathcal{H})$ と表記する。さらに、 $T \succeq 0$ の場合、 T のトレースを次で定義する。

$$\mathrm{Tr} T := \mathrm{Tr}|T|.$$

このトレースの定義は、行列におけるトレースの定義と一致する。次の補題が成り立つ。

補題 2.10 ([25]). $A \in \mathcal{L}(\mathcal{H})$, $B \in \mathcal{K}_1(\mathcal{H})$ のとき, 次が成り立つ.

$$\mathrm{Tr}|AB| \leq |A|_{\mathcal{L}} \mathrm{Tr}|B|.$$

コンパクト作用素のもう一つの重要なクラスとして, ヒルベルト-シュミットクラスが挙げられる [25].

定義 2.11. 作用素 $A \in \mathcal{L}(\mathcal{H})$ が次を満たすとき, A はヒルベルト-シュミット作用素と呼ばれる.

$$|A|_{HS} = \left(\sum_{i \in \mathbb{N}} |A\phi_i|^2 \right)^{\frac{1}{2}} < \infty.$$

ここで, $(\phi_i)_{i \in \mathbb{N}}$ は \mathcal{H} の正規直交基底である.

2.2.3 摂動論

この論文では, 行列の固有値問題に関する摂動論を用いる.

命題 2.12 ([52, 77]). 自然数 $N \in \mathbb{N}$ と区間 $I \subset \mathbb{R}$ に対し, 行列値関数 $I \ni t \mapsto S(t) \in \mathbb{R}^{N \times N}$ が自己共役であり, 連続的に微分可能であると仮定する. このとき, $S(t)$ の固有値 $\lambda_n(t)$ ($n = 1, \dots, N$) が I 上で連続的に微分可能である.

次の補題は, 節 7 におけるフィルタリングアルゴリズムの解析において用いられる.

補題 2.13 ([30, 32]). 定義 2.12 同じ条件を仮定する. このとき, 区間 I 上にユニタリ行列値関数 $U(t)$ が存在し, 次を満たす.

$$\frac{d}{dt} \lambda_n(t) = \left[U(t)^* \left(\frac{d}{dt} S(t) \right) U(t) \right]_{nn}.$$

なお, 一般には $U(t)$ は微分可能ではない [32].

2.3 線形作用素に対する計算

2.3.1 ベクトルのアンサンブル

アンサンブル数 $m \in \mathbb{N}$ とし, 状態ベクトル $v^{(k)} \in \mathcal{H}$ ($k = 1, \dots, m$) からなる集合をアンサンブルと呼ぶ. アンサンブルを表す記法として, $\mathbf{V} = [v^{(k)}]_{k=1}^m \in \mathcal{H}^m$ を用いる. もし $\mathcal{H} = \mathbb{R}^l$ ($l \in \mathbb{N}$) であれば, \mathbf{V} は $\mathbb{R}^{l \times m}$ に属する行列に等しい.

アンサンブル $\mathbf{U} = [u^{(k)}]_{k=1}^m$ と $\mathbf{V} = [v^{(k)}]_{k=1}^m \in \mathcal{H}^m$ について, ℓ_2 ノルム $|\mathbf{U}|_2$ を次で定義する :

$$|\mathbf{U}|_2 = \left(\frac{1}{m} \sum_{k=1}^m |u^{(k)}|^2 \right)^{\frac{1}{2}}. \quad (2.3)$$

また, 積 $\mathbf{U}\mathbf{V}^* \in \mathcal{L}(\mathbf{U})$ および $\mathbf{U}^*\mathbf{V} \in \mathbb{R}^{m \times m}$ を以下のように定義する :

$$\mathbf{U}\mathbf{V}^* = \sum_{k=1}^m u^{(k)} \otimes v^{(k)}, \quad \mathbf{U}^*\mathbf{V} = \left[\langle u^{(i)}, v^{(j)} \rangle \right]_{i,j=1}^m. \quad (2.4)$$

同じベクトル $u \in \mathcal{H}$ のみからなるアンサンブルを $u\mathbf{1} = [u, \dots, u] \in \mathcal{H}^m$ と表し, ここで $\mathbf{1} = (1, \dots, 1) \in (\mathbb{R}^m)^*$ とする. このとき, $|u\mathbf{1}|_2^2 = |u|^2$ が成り立つ.

さらに, $x \in \mathbb{R}^m$, $T \in \mathbb{R}^{m \times m}$, および $A \in \mathcal{L}(\mathcal{H})$ に対して次を定義する :

$$\begin{aligned} \mathbf{U}x &= \sum_{k=1}^m x^k u^{(k)} \in \mathcal{H}, \\ u + \mathbf{U} &= u\mathbf{1} + \mathbf{U} = [u + u^{(k)}]_{k=1}^m \in \mathcal{H}^m, \\ \mathbf{U}T &= \left[\sum_l l = 1^m u^{(l)} T_{l,k} \right]_{k=1}^m \in \mathcal{H}^m, \\ A\mathbf{U} &= \left[A u^{(k)} \right]_{k=1}^m \in \mathcal{H}^m. \end{aligned}$$

アンサンブル $\mathbf{V} = [v^{(k)}]_{k=1}^m \in \mathcal{H}^m$ に対し, $\bar{v} = \frac{1}{m} \sum_{k=1}^m v^{(k)}$ をアンサンブル平均, $d\mathbf{V} = [v^{(k)} - \bar{v}]_{k=1}^m \in \mathcal{H}^m$ をアンサンブル摂動と呼ぶ. このとき, アンサンブル \mathbf{V} は平均と摂動に分解され, 次が成り立つ :

$$\mathbf{V} = \bar{v}\mathbf{1} + d\mathbf{V}.$$

(不偏) アンサンブル共分散 $\text{Cov}_m(\mathbf{V}) \in \mathcal{L}_{sa}(\mathcal{H})$ は次で定義される :

$$\text{Cov}_m(\mathbf{V}) = \frac{1}{m-1} d\mathbf{V} d\mathbf{V}^*.$$

この共分散について, $\text{Cov}_m(\mathbf{V}) = \text{Cov}_m(d\mathbf{V})$ および $\text{Cov}_m(\mathbf{V}) \succeq 0$ が容易に確かめられる.

補題 2.14. アンサンブル $\mathbf{V} \in \mathcal{H}^m$ の任意の摂動 $d\mathbf{V}$ に対して, 次が成り立つ.

$$d\mathbf{V}\mathbf{1}^* = 0 \in \mathcal{H}. \quad (2.5)$$

Proof. (2.5) はアンサンブル平均の定義から得られる.

$$d\mathbf{V}\mathbf{1}^* = \sum_{k=1}^m (v^{(k)} - \bar{v}) \cdot 1 = m \left(\frac{1}{m} \sum_{k=1}^m v^{(k)} - \frac{1}{m} \sum_{k=1}^m \bar{v} \right) = m(\bar{v} - \bar{v}) = 0.$$

□

次に, アンサンブル $\mathbf{V} \in \mathcal{H}^m$ の ℓ_2 ノルム (2.3) の同値な表現を示す補題を示す.

補題 2.15. アンサンブル $\mathbf{V} \in \mathcal{H}^m$ の ℓ_2 ノルムについて, 次が成り立つ.

$$|\mathbf{V}|_2^2 = \frac{1}{m} \operatorname{Tr} \mathbf{V}^* \mathbf{V} = \frac{1}{m} \operatorname{Tr} \mathbf{V} \mathbf{V}^* = |\bar{v}|^2 + |d\mathbf{V}|_2^2. \quad (2.6)$$

Proof. 最初の等式は $|\mathbf{V}|_2$ の定義から導かれる.

$$|\mathbf{V}|_2^2 = \frac{1}{m} \sum_{k=1}^m |v^{(k)}|^2 = \frac{1}{m} \sum_{k=1}^m \langle v^{(k)}, v^{(k)} \rangle = \frac{1}{m} \operatorname{Tr} \mathbf{V}^* \mathbf{V}.$$

\mathcal{H} の完全正規直交基底 $(\phi_i)_{i \in \mathbb{N}}$ を取ると, 次が成り立つ.

$$\begin{aligned} |\mathbf{V}|_2^2 &= \frac{1}{m} \sum_{k=1}^m |v^{(k)}|^2 = \frac{1}{m} \sum_{k=1}^m \sum_{i \in \mathbb{N}} \langle v^{(k)}, \phi_i \rangle^2 = \frac{1}{m} \sum_{i \in \mathbb{N}} \sum_{k=1}^m \langle v^{(k)}, \phi_i \rangle^2 \\ &= \frac{1}{m} \sum_{i \in \mathbb{N}} \sum_{k=1}^m \langle \phi_i, (v^{(k)} \otimes v^{(k)}) \phi_i \rangle = \frac{1}{m} \operatorname{Tr} \mathbf{V} \mathbf{V}^*. \end{aligned}$$

定義 2.14 の関係式 $d\mathbf{V} \mathbf{1}^* = 0$ より, $\mathbf{V} \mathbf{V}^* = \bar{v} \mathbf{1} \mathbf{1}^* \bar{v}^* + d\mathbf{V} d\mathbf{V}^* = m\bar{v}, \bar{v}^* + d\mathbf{V} d\mathbf{V}^*$ が成り立つ. よって, $\frac{1}{m} \operatorname{Tr} \mathbf{V} \mathbf{V}^* = |\bar{v}|^2 + |d\mathbf{V}|_2^2$ が得られる. \square

2.3.2 逆作用素

作用素の逆行列を計算するために, 次の技術的な補題を用いる. その他の同値な式については [44, 75] を参照されたい.

補題 2.16 (Woodbury identity [39, 80]). ヒルベルト空間 $\mathcal{H}_1, \mathcal{H}_2$ と, 線形作用素 $A : \mathcal{H}_1 \rightarrow \mathcal{H}_1, B : \mathcal{H}_1 \rightarrow \mathcal{H}_2, C : \mathcal{H}_2 \rightarrow \mathcal{H}_2, D : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ を考える. $A, C, A + BCD$ が可逆であるならば, 次が成り立つ.

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}. \quad (2.7)$$

恒等作用素を $I = I_{\mathcal{H}}$ と表す. 以下の補題を示す.

補題 2.17. $A \in \mathcal{L}(\mathcal{H})$ とする. $I + A$ が可逆であるならば, 次が成り立つ.

$$(I + A)^{-1} = I - (I + A)^{-1}A. \quad (2.8)$$

さらに, $A \in \mathcal{L}_{sa}(\mathcal{H})$ かつ $A \succeq 0$ の場合, 次が成り立つ.

$$0 \preceq A(A + I)^{-1} = (A + I)^{-1}A \preceq I, \quad 0 \preceq (A + I)^{-1} \preceq I. \quad (2.9)$$

Proof. (2.8) は次のように確認される.

$$LHS = (I + A)^{-1}(I + A - A) = I - (I + A)^{-1}A.$$

(2.9) について, (2.8) から $A(A + I)^{-1} = (A + I)^{-1}A$ が得られる. 不等式はスペクトル写像定理 [25] より従う. \square

補題 2.18. $\Gamma : \mathcal{H} \rightarrow \mathcal{H}$ を可逆作用素, $V \in \mathcal{H}^m$ とする. このとき, $VV^* + \Gamma$ は可逆であり, 次が成り立つ.

$$(I + V^*\Gamma^{-1}V)^{-1}V^*\Gamma^{-1} = V^*(VV^* + \Gamma)^{-1}. \quad (2.10)$$

さらに,

$$(I + V^*\Gamma^{-1}V)^{-1} = I - V^*(VV^* + \Gamma)^{-1}V. \quad (2.11)$$

Proof. $VV^* \succeq 0$ および $\Gamma \succ 0$ より, $VV^* + \Gamma$ は可逆である. このとき,

$$V^*\Gamma^{-1}(VV^* + \Gamma) = V^*\Gamma^{-1}VV^* + V^* = (I + V^*\Gamma^{-1}V)V^*,$$

が成り立ち, (2.10) が得られる.

(2.11) について, (2.10) を用いると,

$$(I + V^*\Gamma^{-1}V)^{-1}V^*\Gamma^{-1}V = V^*(VV^* + \Gamma)^{-1}V.$$

したがって,

$$I - V^*(VV^* + \Gamma)^{-1}V = I - (I + V^*\Gamma^{-1}V)^{-1}V^*\Gamma^{-1}V = (I + V^*\Gamma^{-1}V)^{-1}$$

が成り立つ. ここで, 最後の等式は (2.8) から従う. \square

補題 2.19. $A \in \mathcal{L}(\mathcal{H})$ を可逆とする. $U \in \mathcal{L}(\mathcal{H})$ がユニタリであれば, 次が成り立つ.

$$UA^{-1}U^* = (UAU^*)^{-1}.$$

また, 対角行列 $\Sigma \succ 0$ に対して, 次が成り立つ.

$$\Sigma A^{-1}\Sigma = (\Sigma^{-1}A\Sigma^{-1})^{-1}.$$

Proof. $U^{-1} = U^*$ および Σ が可逆であることを用いると, いずれの等式も可逆作用素 $A, B \in \mathcal{L}(\mathcal{H})$ に対する $(AB)^{-1} = B^{-1}A^{-1}$ から従う. \square

2.4 確率論

状態推定問題における不確実性を定量化するため、確率論が必要である。標本空間 Ω 、 σ -代数 \mathcal{F} 、および確率測度 \mathbb{P} に対し、確率空間を $(\Omega, \mathcal{F}, \mathbb{P})$ とする。この確率空間に関する期待値を $\mathbb{E}[\cdot]$ で表す。部分集合の族 $E \subset 2^\Omega$ に対し、 E を含む最小の σ -代数を $\sigma(E)$ と表記する。ヒルベルト空間 \mathcal{H} とそのボレル σ -代数 $\mathcal{B}(\mathcal{H}) = \sigma(\{O \subset \mathcal{H} \mid O : \text{open}\})$ を持つ可測空間 $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ を考える。確率測度の集合を $\mathcal{M}_1(\mathcal{H})$ と表す。バナッハ空間 \mathcal{X} 、可測関数 $f : \mathcal{H} \rightarrow \mathcal{X}$ 、および $\mu \in \mathcal{M}_1(\mathcal{H})$ に対して、Pettis 積分の意味で

$$\mathbb{E}_\mu[f] = \int_{\mathcal{H}} f(u) d\mu(u) \quad (2.12)$$

と定義する [3, 80]。

$\mu \in \mathcal{M}_1(\mathcal{H})$ に対し、その平均を

$$\varpi_\mu = \mathbb{E}_\mu[x] = \int_{\mathcal{H}} x d\mu(x) \in \mathcal{H}, \quad (2.13)$$

共分散を

$$C_\mu = \int_{\mathcal{H}} (x - \varpi_\mu) \otimes (x - \varpi_\mu) d\mu(x) \in \mathcal{L}_{sa}(\mathcal{H}) \quad (2.14)$$

と定義する。

確率変数 U を可測写像 $U : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathcal{H}, \mathcal{B}(\mathcal{H}))$ とする。 U に対応する押し出し測度（像測度）を $\mathbb{P}^U = U_* \mathbb{P} \in \mathcal{M}_1(\mathcal{H})$ とし、 U に付随する σ -代数を

$$\sigma(U) = \sigma(U^{-1}(E) \mid E \in \mathcal{B}(\mathcal{H})) \quad (2.15)$$

と定義する。

同様に、 U の平均と共分散を次のように与える：

$$\varpi_U = \mathbb{E}[U] = \int_{\mathcal{H}} u d\mathbb{P}^U(u) = \int_{\Omega} U(\omega) d\mathbb{P}(\omega), \quad (2.16)$$

$$C_U = \mathbb{E}[(U - \varpi_U) \otimes (U - \varpi_U)]. \quad (2.17)$$

部分 σ -代数 $\mathcal{G} \subset \mathcal{F}$ に対する U の条件付き期待値を $\mathbb{E}[U \mid \mathcal{G}]$ と表す。

時間インデックス集合を $\mathcal{T} = \mathbb{N} \cup \{0\}$ または $[0, \infty) \subset \mathbb{R}$ とし、確率過程 $U : \mathcal{T} \times \Omega \rightarrow \mathcal{H}$ を考える。部分 σ -代数の族 $(\mathcal{F}_t)_{t \in \mathcal{T}}$ が $s \leq t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t$ を満たすとき、これをフィルトレーションと呼ぶ。確率過程 U に付随するフィルトレーションを次のように定義する：

$$\mathcal{F}_t^U = \sigma(\{U_s^{-1}(E) \subset \Omega \mid E \subset \mathcal{B}(\mathcal{H}), s \leq t, s \in \mathcal{T}\}). \quad (2.18)$$

さらに, $U_0 = u \in \mathcal{H}$ を条件とした期待値を

$$\mathbb{E}^u[f(U_t)] = \mathbb{E}[f(U_t) | U_0 = u] \quad (2.19)$$

と表す. ここで, $t \in \mathcal{T}$, $f : \mathcal{H} \rightarrow \mathbb{R}$ は可積分関数とする.

測度 μ, ν が $(\mathcal{X}, \mathcal{F})$ 上の σ 有限測度であり, $\nu \ll \mu$ ならば, 次が成り立つ [80] :

$$\int_E f d\nu = \int_E f \rho d\mu. \quad (2.20)$$

ここで $\rho = \frac{d\nu}{d\mu}$ は Radon-Nikodým 微分と呼ばれ, ベイズ推定において重要な役割を果たす.

有限次元の場合を考える. $l \in \mathbb{N}$ に対し, $\mu \in \mathcal{M}_1(\mathbb{R}^l)$ が \mathbb{R}^l 上のルベーグ測度に関して絶対連続であるならば, Radon-Nikodým 微分は確率密度関数 (PDF) と呼ばれ, p_μ と表記する. 同様に確率変数 U に対し, \mathbb{P}^U の PDF を p_U と表記する. \mathbb{R}^l 上のガウス測度はその PDF によって定義される.

定義 2.20 (\mathbb{R}^l 上のガウス測度). $\varpi \in \mathbb{R}^l$ および $C \in \mathcal{L}_{sa}(\mathbb{R}^l)$ で $C \succ 0$ を満たすものとする. 平均 ϖ および共分散 C を持つガウス測度は, その PDF を次で定義される:

$$p(x) = \frac{1}{\sqrt{\det C(2\pi)^l}} \exp\left(-\frac{1}{2}|x - \varpi|_C^2\right). \quad (2.21)$$

これを $\mathcal{N}(\varpi, C) \in \mathcal{M}_1(\mathbb{R}^l)$ と表記する.

一般の場合に戻って, 以下のように確率測度間の距離を定義する. その他の距離やそれらの関係については [37] を参照されたい.

定義 2.21 ([80]). $\mu, \nu \in \mathcal{M}_1(\mathcal{H})$ とする. 全変動距離 (total variation distance) は次で定義される:

$$d_{TV}(\mu, \nu) = \sup \{|\mu(A) - \nu(A)| \mid A \in \mathcal{B}(\mathcal{H})\}.$$

ヘリンジャー距離 (Hellinger distance) は次で与えられる:

$$d_H(\mu, \nu) = \left(\int_{\mathcal{H}} \left| \sqrt{\frac{d\mu}{d\rho}} - \sqrt{\frac{d\nu}{d\rho}} \right|^2 d\rho \right)^{\frac{1}{2}}.$$

ここで, $\rho \in \mathcal{M}_1(\mathcal{H})$ は参考測度であり, 選び方によらない.

次の不等式は, 測度やモーメントの差を評価する際に有用である.

命題 2.22 ([37, 80]). $\mu, \nu \in \mathcal{M}_1(\mathcal{H})$ に対し, 次が成り立つ:

$$d_H(\mu, \nu)^2 \leq d_{TV}(\mu, \nu) \leq 2d_H(\mu, \nu).$$

命題 2.23 ([80]). $\mu, \nu \in \mathcal{M}_1(\mathcal{H})$, \mathcal{X} をノルム $|\cdot|$ を備えたバナッハ空間とし, $f : \mathcal{H} \rightarrow \mathcal{X}$ を可測関数とする. また, $\mathbb{E}_\mu[|f|^2]$ および $\mathbb{E}_\nu[|f|^2] < \infty$ であると仮定する. このとき, 次が成り立つ:

$$|\mathbb{E}_\mu[f] - \mathbb{E}_\nu[f]| \leq 2\sqrt{\mathbb{E}_\mu[|f|^2] + \mathbb{E}_\nu[|f|^2]}d_H(\mu, \nu).$$

最後に, 無限次元ヒルベルト空間上の測度に関する重要な事実を振り返る. 詳細については [80] の第 2 章および関連文献を参照されたい.

命題 2.24 (ヒルベルト空間上のルベーグ測度 [80]). 無限次元ヒルベルト空間 \mathcal{H} 上の測度 μ が, すべての平行移動に対して不変であり, かつ局所有限である (すなわち, 任意の $u \in \mathcal{H}$ に対して, ある可測集合 O_u が存在し, $u \in O_u$ かつ $\mu(O_u) < \infty$ である) と仮定する. このとき, μ は零測度である.

この命題から, \mathcal{H} 上にルベーグ測度を定義することはできない. しかし, ガウス測度は \mathcal{H} 上において適切に定義される.

定義 2.25 (ヒルベルト空間上のガウス測度). $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ 上の測度 μ が, 任意の連続線形汎関数 $\ell : \mathcal{H} \rightarrow \mathbb{R}$ に対し, 押し出し測度 $\ell_*\mu$ が (非退化な) \mathbb{R} 上のガウス測度であるとき, μ はガウス測度と呼ばれる.

命題 2.26 (Sazonov の定理 [80]). 平均ゼロのガウス測度 $\mu \in \mathcal{M}_1(\mathcal{H})$ を考える. このとき, 共分散 $C_\mu \in \mathcal{K}_1(\mathcal{H})$ であり, 次が成り立つ:

$$\mathrm{Tr} C_\mu = \int_{\mathcal{H}} |x|^2 d\mu(x).$$

逆に, $C \in \mathcal{L}_{sa} \cap \mathcal{K}_1(\mathcal{H})$ で, 任意の $x \in \mathcal{H}$ に対して $\langle Cx, x \rangle > 0$ が成り立つならば, \mathcal{H} 上に共分散 $C_\mu = C$ を持つガウス測度 μ が存在する.

定義 2.26 は, ガウス測度の共分散がトレースクラスである必要があることを意味する. この場合, $\mu = \mathcal{N}(0, C)$ と表記する. さらに, $\varpi \in \mathcal{H}$ および $X_0 \sim \mathcal{N}(0, C)$ としたとき, 平行移動されたガウス確率変数 $X_\varpi = \varpi + X_0$ に対して, $\mathbb{P}^{X_\varpi} = \mathcal{N}(\varpi, C)$ と表記する. また, 共分散 $C \succeq 0$ を持つ退化ガウス測度についても同じ表記を用いる.

2.5 ベイズ推定

ベイズ推定は, ノイズを含む観測 y から状態 u を推定し, その不確実性をベイズの定理に基づいて定量化するための数学的枠組みを提供する. 無限次元空間における概念や定式化の詳細については [26, 79] を参照されたい.

2.5.1 ベイズの定理

まず、ユークリッド空間 $\mathcal{H} = \mathbb{R}^{N_u}$ と $\mathcal{Y} = \mathbb{R}^{N_y}$ におけるベイズの定理を紹介する。条件付き PDF $p_Y(y|u)$ が既知であると仮定する。たとえば、ノイズを含む観測 y が $y \sim \mathcal{N}(u, R)$ (ここで $R \in \mathcal{L}_{sa}(\mathbb{R}^{N_y})$ かつ $R \succ 0$) に従う場合、 $p_Y(y|u)$ は (2.21) によって与えられる。

命題 2.27 (ベイズの定理 [63, 76])。観測 $y \in \mathbb{R}^{N_y}$ に対し、 U の条件付き PDF はベイズの公式によって次のように与えられる：

$$p_U(u|y) = \frac{p_Y(y|u)p_U(u)}{p_Y(y)}, \quad (2.22)$$

ここで、 $p_U(u)$ は U の PDF であり、 $p_Y(y) = \int_{\mathbb{R}^{N_y}} p_Y(y|u)p_U(u) du$ である。

ベイズ推定の文脈では、 $p_U(u)$ および $p_U(u|y)$ はそれぞれ事前分布 (prior) と事後分布 (posterior) と呼ばれる。

事前分布 $p_U(u)$ は、状態 u の初期推定における不確実性を表す。観測 y が与えられた場合、事前分布 $p_U(u)$ はベイズの公式 (定義 2.27) に基づき、尤度 $p_Y(y|u)$ を掛けることで事後分布 $p_U(u|y)$ へ更新される：

$$p_U(u) \rightarrow p_Y(y|u) \propto p_Y(y|u)p_U(u).$$

事後分布 $p_U(u|y)$ は、観測 y から得られる情報を事前知識に組み込んだ後の u の推定における不確実性を反映する。

図 1 は、1 次元の状態空間と観測空間におけるベイズの公式を図示している。

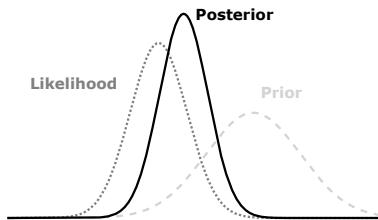


図 1: ベイズの公式

次の補題は、ガウス事後分布の明示的な公式を与える。

補題 2.28 (ガウス分布の条件付け [80])。 $(U, Y) \sim \mathcal{N}(\varpi, C)$ を、平均

$$\varpi = \begin{bmatrix} \varpi_1 \\ \varpi_2 \end{bmatrix} \in \mathbb{R}^{N_u \times N_y}$$

および共分散

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^* & C_{22} \end{bmatrix} \succ 0$$

を持つ $\mathbb{R}^{N_u \times N_y}$ 上の結合ガウス分布とする。条件 $Y = y$ の下での U の条件付き分布は、次のようなガウス分布である：

$$\mathbb{P}^U(\cdot | Y = y) \sim \mathcal{N}(\varpi_1 + C_{12}C_{22}^{-1}(y - \varpi_2), C_{11} - C_{12}C_{22}^{-1}C_{12}^*).$$

ガウス分布の条件付けは、データ同化におけるガウス近似アルゴリズムを構築する際に重要である。

無限次元ヒルベルト空間上において、ルベーグ測度が存在しないことは定義 2.24 により示されているため、それに対する密度関数を考えることはできない。そのため、(2.22) を事後分布が事前分布に対して持つ Radon-Nikodým 微分の観点から一般化する。もし $\dim(\mathcal{Y}) < \infty$ であれば、事後分布を次のように定義できる。

命題 2.29 (一般化されたベイズの公式 [26, 29, 80]). $\mathcal{Y} = \mathbb{R}^{N_y}$ とし、 $h : \mathcal{H} \rightarrow \mathcal{Y}$ を連続写像、 $\mu \in \mathcal{M}_1(\mathcal{Y})$ を観測ノイズの分布でありその PDF を p_μ とする。このとき、事後分布 $\mu^y(du) = \mathbb{P}(du | y)$ は事前分布 $\mu_0 \in \mathcal{M}_1(\mathcal{H})$ に対して絶対連続であり、その Radon-Nikodým 微分は次で与えられる：

$$\frac{d\mu^y}{d\mu_0}(u) \propto \exp(-\Phi(u; y)), \quad (2.23)$$

ここで、 $\Phi(u; y) = -\log(p_\mu(y - h(u)))$ とする。

2.5.2 事後分布の適切性

ベイズ逆問題は、観測データに対する連続的な事後分布を与える。これは事後分布の適切性 (well-posedness) として知られている。 \mathcal{H} および \mathcal{Y} をそれぞれノルム $|\cdot|_{\mathcal{H}}$ および $|\cdot|_{\mathcal{Y}}$ を備えた 2 つのヒルベルト空間とする。ここで、ポテンシャル関数を用いて事後分布を定義し、いくつかの仮定を課す。

仮定 2.30. $\Phi(\cdot; \cdot) : \mathcal{H} \times \mathcal{Y} \rightarrow \mathbb{R}$ とする。

(1) 任意の $\epsilon, r > 0$ に対して、 $M = M(\epsilon, r) \in \mathbb{R}$ が存在し、次が成り立つ：

$$\Phi(u; y) \geq M - \epsilon|u|_{\mathcal{H}}^2, \quad u \in \mathcal{H}, |y|_{\mathcal{Y}} < r.$$

(2) 任意の $r > 0$ に対して、 $K = K(r) > 0$ が存在し、次が成り立つ：

$$\Phi(u; y) \leq K, \quad |u|_{\mathcal{H}}, |y|_{\mathcal{Y}} < r.$$

(3) 任意の $r > 0$ に対して, $L = L(r) > 0$ が存在し, 次が成り立つ:

$$|\Phi(u_1; y) - \Phi(u_2; y)| \leq L|u_1 - u_2|_{\mathcal{H}}, \quad |u_1|_{\mathcal{H}}, |u_2|_{\mathcal{H}}, |y|_{\mathcal{Y}} < r.$$

(4) 任意の $\epsilon, r > 0$ に対して, $C = C(\epsilon, r) > 0$ が存在し, 次が成り立つ:

$$|\Phi(u; y_1) - \Phi(u; y_2)| \leq \exp(\epsilon|u|_{\mathcal{H}}^2 + C)|y_1 - y_2|_{\mathcal{Y}}, \quad |u|_{\mathcal{H}}, |y_1|_{\mathcal{Y}}, |y_2|_{\mathcal{Y}} < r.$$

例 2.31. $\mathcal{Y} = \mathbb{R}^{N_y}$, $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$, および $R \in \mathbb{R}^{N_y \times N_y}$ で $R \succ 0$ とする. ポテンシャル関数を次で定義する:

$$\Phi(u; y) = |y - Hu|_R^2.$$

このとき, $\Phi(u; y)$ は定義 2.30 を満たす. この例はガウス尤度から導かれる. より一般的な例については, [80] の第 6 章を参照されたい.

命題 2.32 (ベイズ逆問題の適切性 [26, 80]). Φ が定義 2.30 を満たし, μ_0 が \mathcal{H} 上のガウス分布であると仮定する. このとき, 任意の $y \in \mathcal{Y}$ に対して, 事後分布

$$\frac{d\mu^y}{d\mu_0} = Z(y)^{-1} \exp(-\Phi(u; y)), \quad Z(y) = \int_{\mathcal{H}} \exp(-\Phi(u; y)) d\mu_0(u),$$

は well-defined である. さらに, 任意の $r > 0$ に対して, ある $C = C(r) > 0$ が存在し, 次が成り立つ:

$$d_H(\mu^{y_1}, \mu^{y_2}) \leq C(r)|y_1 - y_2|_{\mathcal{Y}}, \quad |y_1|_{\mathcal{Y}}, |y_2|_{\mathcal{Y}} \leq r.$$

定義 2.30 の条件 (1) および (2) は, それぞれ定義 2.32 における正規化定数 $Z(y)$ の有限性と正値性を保証する. 定義 2.32 は, 事後分布が観測データに関して局所的なりپシツ連続性を持つことを意味する. 言い換えれば, ベイズ逆問題は不確実なデータから状態をロバストに推定する枠組みを与える. 定義 2.23 および定義 2.32 から, 次の系が得られる.

系 2.33. \mathcal{X} をノルム $|\cdot|_{\mathcal{X}}$ を持つバナッハ空間, $f : \mathcal{H} \rightarrow \mathcal{X}$ を関数とする. また, $\mathbb{E}_{\mu_0}[|f|_{\mathcal{X}}^2] < \infty$ と仮定する. このとき, 任意の $r > 0$ に対して, ある $C = C(r) > 0$ が存在し, 次が成り立つ:

$$|\mathbb{E}_{\mu^{y_1}}[f] - \mathbb{E}_{\mu^{y_2}}[f]|_{\mathcal{X}} \leq C|y_1 - y_2|_{\mathcal{Y}}, \quad |y_1|_{\mathcal{Y}}, |y_2|_{\mathcal{Y}} \leq r.$$

定義 2.33 において, $\mathcal{X} = \mathcal{H}$, $|\cdot|_{\mathcal{X}} = |\cdot|$, および $f(u) = u$ とすると, 事後分布の平均は y に関して連続であることが示される. 一方, 事後分布のモードは一般には y に関して連続ではない [63].

2.6 力学系

データ同化における数学モデルを記述するために力学系の理論を用いる。特に、偏微分方程式（PDE）を無限次元の力学系として考えることが重要である [59, 85]。有限次元の力学系の理論については [53] の包括的な教科書も参照されたい。

\mathcal{H} をヒルベルト空間とする。力学系は \mathcal{H} 上の半群として定義される。

定義 2.34 (半群). \mathcal{H} 上の半群 (*semigroup*) は、次の条件を満たす \mathcal{H} からそれ自体への写像の族 $\{\Psi_t \mid t \geq 0\}$ である：

- (1) $\Psi_0 = id_{\mathcal{H}}$ (恒等写像)；
- (2) $\Psi_{t+s} = \Psi_t \circ \Psi_s, \forall t, s \geq 0$ ；
- (3) $\Psi_t(u_0)$ は t および u_0 に関して連続である。

半群は、常微分方程式（ODE）または偏微分方程式（PDE）によって生成されることが多い。たとえば、初期条件 $u(0, x) = u_0(x)$ を持つ PDE の解 $u(t, x)$ が与えられているとする。もし $u(t, \cdot) \in \mathcal{H}$ であるならば、 $\Psi_t : \mathcal{H} \rightarrow \mathcal{H}$ を次のように定義できる：

$$\Psi_t(u_0)(\cdot) = u(t, \cdot).$$

したがって、原則として、解が半群を生成するようにモデル方程式が well-defined である（すなわち、解の存在、唯一性、および初期条件に対する連続依存性が成り立つ）ことを仮定する。ODE の場合、半群の存在に関する十分条件が次のように与えられる。

命題 2.35 (Picard-Lindelöf の定理 [59]). $\mathcal{F} : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_u}$ が局所リップシツ連続であるとする。すなわち、任意の $r > 0$ に対して、 $L = L(r) > 0$ が存在し、次が成り立つ：

$$|\mathcal{F}(u) - \mathcal{F}(v)| \leq L|u - v|, \quad \forall |u|, |v| \leq r.$$

このとき、次の ODE

$$\frac{du}{dt} = \mathcal{F}(u), \quad u(0) = u_0 \in \mathbb{R}^{N_u},$$

は $[0, T]$ ($T = T(u_0) > 0$) 上で一意な解を持つ。

力学系の長期挙動を特徴づけるための基本的な概念を次に示す。

定義 2.36. Ψ_t を \mathcal{H} 上の力学系とする。

- (1) $X \subset \mathcal{H}$ がすべての $t \geq 0$ に対して $\Psi_t(X) = X$ を満たすとき、 X は不变集合 (*invariant*) と呼ばれる。

(2) $X, B \subset \mathcal{H}$ が

$$\text{dist}(\Psi_t(B), X) \rightarrow 0 \quad (t \rightarrow \infty)$$

を満たすとき, X は B を吸引する (*attracts*) と呼ばれる. ここで, $\text{dist}(A, B) = \sup_{a \in A} \inf_{b \in B} |a - b|$. さらに, X がすべての有界集合 $B \subset \mathcal{H}$ を吸引するとき, X は吸引集合 (*attracting*) と呼ばれる.

(3) $\mathcal{A} \subset \mathcal{H}$ がコンパクト, 不変, かつ吸引集合であるとき, \mathcal{A} は大域的アトラクター (*global attractor*) と呼ばれる.

大域的アトラクターは次の性質を満たす.

命題 2.37 ([59]). Ψ_t を \mathcal{H} 上の力学系とする.

(1) 大域的アトラクター \mathcal{A} は一意である.

(2) 大域的アトラクター \mathcal{A} は最大のコンパクト不変集合であり, 最小のアトラクターである.

(3) 大域的アトラクター \mathcal{A} が存在するのは, コンパクトなアトラクターが存在するとき, かつそのときに限る.

アトラクターの存在に代えて, 多くの応用ではより強い結果を示すことが可能である.

定義 2.38 (吸収集合). $X \subset \mathcal{H}$ に対して, 任意の有界集合 $B \subset \mathcal{H}$ について, ある $T = T(B) \geq 0$ が存在し, 次が成り立つとき, X を吸収集合 (*absorbing set*) と呼ぶ:

$$\Psi_t(B) \subset X, \quad \forall t \geq T.$$

注意 2.39. 定義 2.37 より, コンパクトな吸収集合の存在は, 大域的アトラクターの存在を意味する.

吸収集合の存在を示すために有用な不等式として, 運動エネルギー原理 (kinetic energy principle) と呼ばれる次の式がある:

$$\frac{d}{dt} |\Psi_t(u_0)|^2 \leq -\lambda |u_0|^2 + K, \quad (2.24)$$

ここで, $t \geq 0$ かつ $u_0 \in \mathcal{H}$ に対して, $\lambda, K > 0$ が存在する.

Gronwall の補題の結果として, 次が成り立つ:

$$|\Psi_t(u_0)|^2 \leq e^{-\lambda t} |u_0|^2 + \frac{K}{\lambda} (1 - e^{-\lambda t}). \quad (2.25)$$

これは、散逸力学系における本質的な性質である。一般に、(2.24) は次の形式のリアプノフ関数 $\mathcal{E}(\cdot) = |\cdot|^2$ の存在とみなされる：

$$\frac{d}{dt}\mathcal{E}(u_t) \leq -\lambda\mathcal{E}(u_t) + K, \quad \lambda, K > 0. \quad (2.26)$$

これらの概念は、離散時間の力学系にも適用可能である [88]。節 5 では、地球流体の重要な例を伴う散逸力学系における大域的アトラクターの存在やその他の性質について議論する。

3. 状態空間モデルとデータ同化

3.1 有限次元問題

3.1.1 状態空間モデル

有限次元の状態空間 $\mathcal{H} = \mathbb{R}^{N_u}$ と観測空間 $\mathcal{Y} = \mathbb{R}^{N_y}$ ($N_u \geq N_y$) を考える。真の状態の時間発展は、次の形式の離散時間確率過程 $U : \mathbb{N} \times \Omega \rightarrow \mathbb{R}^{N_u}$ によってモデル化されるものとする：

$$U_n = \Psi(U_{n-1}) + \xi_n, \quad (3.1)$$

ここで、 $U_0 \in \mathbb{R}^{N_u}$ は不確定な初期状態であり、 $\Psi : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_u}$ は連続写像とする。ノイズ列 $(\xi_n)_{n \in \mathbb{N}} \subset \mathbb{R}^{N_u}$ は独立同分布 (i.i.d.) の確率誤差であり、モデリングおよび近似誤差を表す。その平均は 0 で、共分散行列は $Q \in \mathbb{R}^{N_u \times N_u}, Q \succeq 0$ で表される。

未知の真の状態に関する情報は、次の形式のノイズを含む観測 $Y : \mathbb{N} \times \Omega \rightarrow \mathbb{R}^{N_y}$ を通じて得られる：

$$Y_n = h(U_n) + \eta_n, \quad n \in \mathbb{N}, \quad (3.2)$$

ここで、 $h : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_y}$ は連続な観測関数であり、 $(\eta_n)_{n \in \mathbb{N}} \subset \mathbb{R}^{N_y}$ は確率密度関数 (PDF) p_Y を持つ独立同分布のノイズ列とする。

データ同化アルゴリズムを節 4 で定式化するために、観測ノイズに対して次の仮定を課す。

仮定 3.1. 任意の $n \in \mathbb{N}$ に対し、 $\eta_n \sim \mathcal{N}(0, R)$ である。ここで、 $R \in \mathbb{R}^{N_y \times N_y}$ かつ $R \succ 0$ とする。

理想的な設定において、データ同化アルゴリズムを解析する際には完全観測 (full observation) が考慮される。

仮定 3.2 (完全観測). 状態は完全に観測される、すなわち、 $h = id_{\mathcal{H}}$ かつ $R = r^2 I_{\mathcal{H}}$ ($r > 0$) である。観測関数が線形作用素 $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$ 、すなわち $h(u) = Hu$ である場合、 $H = I_{\mathcal{H}}$ と仮定する。

数学解析に便利な連続時間での有限次元状態空間モデルを定義する。確率過程 $U : [0, \infty) \times \Omega \rightarrow \mathbb{R}^{N_u}$ に対して、次の確率微分方程式 (SDE) を考える：

$$dU_t = \mathcal{F}(U_t)dt + Q^{\frac{1}{2}}dW_t, \quad (3.3)$$

ここで, $\mathcal{F} : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_u}$ は連続写像, $W = (W_t)_{t \geq 0}$ は N_u 次元のウィーナー過程, $Q \in \mathbb{R}^{N_u \times N_u}$ で $Q \succeq 0$ とする. SDE に関する詳細は, 解の存在と一意性定理を含む基本的な教科書 [73] を参照されたい. 同様に, 連続時間での確率的な観測を次のように採用する:

$$dY_t = h(U_t)dt + R^{\frac{1}{2}}dB_t, \quad (3.4)$$

ここで, $h : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_y}$ は連続写像, $(B_t)_{t \geq 0}$ は N_y 次元のウィーナー過程であり, W と独立, $R \in \mathbb{R}^{N_y \times N_y}$ で $R \succeq 0$ とする. 定義 3.1 に対応する次の仮定を考える.

仮定 3.3. 観測ノイズの共分散は正定値である, $R \succ 0$.

3.1.2 ベイズ的データ同化問題

$\mathcal{T} = \mathbb{N} \cup 0$ または $[0, \infty) \subset \mathbb{R}$ を時間指標集合とする. 確率過程 $U : \mathcal{T} \times \Omega \rightarrow \mathcal{H}$ および $Y : \mathcal{T} \times \Omega \rightarrow \mathcal{Y}$ はそれぞれ真の状態 (true state) と観測 (observation) を表す. 時刻 t までの観測を $\mathbf{Y}_t = \{Y_s \mid 0 \leq s \leq t\}$ と表記する. まず, 観測を用いて真の状態からの L^2 誤差を最小化する状態 (信号) 推定問題を定式化する.

定義 3.4. $t \in \mathcal{T}$ に対して, 確率変数 $V_t : \Omega \rightarrow \mathcal{H}$ が観測 \mathbf{Y}_t に基づく推定量 (estimator) と呼ばれるのは, V_t が \mathcal{F}_t^Y -可測である場合である. さらに, 次を満たすとき, V_t は最適 (optimal) であるという:

$$\mathbb{E}[|U_t - V_t|^2] = \inf\{\mathbb{E}[|U_t - V|^2] \mid V \in \mathcal{K}_t\}.$$

ここで, $\mathcal{K}_t = \{V : \Omega \rightarrow \mathcal{H} \mid V \in L^2(\Omega, \mathbb{P}) \text{ かつ } Y_t \text{ に基づく推定量}\}$. また, $L^2(\Omega, \mathbb{P})$ は Ω 上で \mathbb{P} に関して 2乗可積分な関数の空間である. 状態推定問題とは, 観測 \mathbf{Y}_t に基づいて最適推定量 V_t を構築または近似することである.

次の命題は, 最適推定量が条件付き期待値によって得られることを示す.

命題 3.5 (最適推定 [73]). 状態推定問題の最適推定量 V_t は次によって与えられる:

$$V_t = \mathbb{E}[U_t \mid \mathcal{F}_t^Y].$$

次に, 状態推定問題のベイズ定式化を考える. この場合, 推定は条件付き分布によって表される. ここでは離散時間系を考える. $N \in \mathbb{N}$ に対して, 離散時間区間 $0 \leq n \leq N$ での観測の実現値を $\mathbf{y}_N = \{y_n \mid 0 \leq n \leq N\}$ と表記する.

定義 3.6 (データ同化問題). 未知の真の状態を U , 観測値 $N \in \mathbb{N}$ までの観測の実現値を \mathbf{y}_N とする. $n \in \mathbb{N}$ に対して, 確率変数 V_n を構築する問題を考える. ここで, V_n の確率分布が \mathbf{y}_N に関する U_n の条件付き確率分布に対応するようにする, $\mathbb{P}^{V_n} = \mathbb{P}^{U_n}(\cdot | \mathbf{y}_N)$. この問題を (ベイズ的) データ同化問題 (*data assimilation problem*) と呼ぶ.

データ同化問題は, n と N に応じて次の 3つのタイプに分類される:

- $n > N$ の場合は予測 (*prediction*);
- $n = N$ の場合はフィルタリング (*filtering*);
- $n < N$ の場合はスムージング (*smoothing*).

予測分布 $\mathbb{P}^{U_n}(\cdot | \mathbf{y}_N)$ ($n > N$) は, モデルダイナミクスによってフィルタ分布の押し出しとして得られるため, フィルタリングとスムージングの問題を扱えば十分である. 定義 3.6において, これらの分布は定義 2.27 のベイズ公式を用いて事後分布として得られる.

多くの実世界の応用では, 観測は離散時間ステップごとに得られることが多い. 離散時間のフィルタリング問題では, 分布 \mathbb{P}^{V_n} を逐次更新することが有用である. モデルノイズ ξ_n には確率密度関数 (PDF) p_ξ があると仮定する. モデルが決定論的である場合, すなわち (3.1) のノイズ列 $(\xi_n)_{n \in \mathbb{N}}$ の共分散が O である場合には, p_ξ の代わりにディラックのデルタ関数を用いる.

定義 3.7 (有限次元状態空間における逐次データ同化). U および Y がそれぞれ (3.1) および (3.2) に従うとする. 初期状態 U_0 の不確定性を PDF p_{U_0} で表す. 次の逐次更新により, $p_{V_0} = p_{U_0}$ から始めて, 任意の $n \in \mathbb{N}$ に対して厳密なフィルタ分布 $p_{V_n} = p_{U_n}(\cdot | \mathbf{Y}_n)$ を得る:

(I) (予測: $p_{V_{n-1}} \rightarrow p_{\widehat{V}_n}$) モデルダイナミクスを用いて $p_{V_{n-1}}$ を $p_{\widehat{V}_n}$ に伝播する:

$$p_{\widehat{V}_n}(v) = \int_{\mathbb{R}^{N_u}} p_\xi(v - \Psi(v')) p_{V_{n-1}}(v') dv'. \quad (3.5)$$

(II) (解析: $p_{\widehat{V}_n}, y_n \rightarrow p_{V_n}$) ベイズの公式を用いて $p_{\widehat{V}_n}$ を p_{V_n} に更新する:

$$p_{V_n}(v) = \frac{p_Y(y_n | v) p_{\widehat{V}_n}(v)}{\int_{\mathbb{R}^{N_u}} p_Y(y_n | v') p_{\widehat{V}_n}(v') dv'}, \quad (3.6)$$

ここで, $p_Y(y | u)$ は $u \in \mathbb{R}^{N_u}$ に関する Y の条件付き PDFである.

(I)のステップは予測 (prediction または forecast) ステップと呼ばれ, $\mathbb{P}^{\widehat{V}_n}(dv) = p_{\widehat{V}_n}(v) dv$ は予測分布 (prediction または forecast distribution) と呼ばれる. 同様に, (II)のステッ

プロセスは解析 (analysis または update) ステップと呼ばれ, $\mathbb{P}^{V_n}(dv) = p_{V_n}(v) dv$ は解析分布またはフィルタ分布 (analysis または filtering distribution) と呼ばれる. [63] では, 定義 3.7 の 2 つのステップ (I) および (II) は次のように作用素形式で表される:

$$\mathbb{P}^{\widehat{V}_n} = \mathcal{P} \mathbb{P}^{V_{n-1}}, \quad \mathbb{P}^{V_n} = L_{y_n} \mathbb{P}^{\widehat{V}_n},$$

ここで, $\mathcal{P} : \mathcal{M}_1(\mathbb{R}^{N_u}) \rightarrow \mathcal{M}_1(\mathbb{R}^{N_u})$ は (3.1) に関連するマルコフ遷移作用素であり, $L_{y_n} : \mathcal{M}_1(\mathbb{R}^{N_u}) \rightarrow \mathcal{M}_1(\mathbb{R}^{N_u})$ は観測 y_n を用いたベイズ更新を表している.

厳密なフィルタ分布 $\mathbb{P}^{V_n}(dv)$ を近似するための数値アルゴリズムを構築する必要がある. これについては節 4 で議論する. フィルタリング問題における真の状態を推定するためのいくつかの階層を図 2 に示す. 階層 (a): モデルダイナミクスによって生成される隠れた真の状態を表している. 階層 (b): 真の状態からノイズを含む観測値が得られる. 階層 (c): モデルダイナミクスにより, 真の状態の条件付き分布 $\mathbb{P}^{U_{n-1}}(\cdot | y_{n-1})$ が $\mathbb{P}^{U_n}(\cdot | y_{n-1})$ に伝播され, 観測データによって条件付けられることで $\mathbb{P}^{U_n}(\cdot | y_n)$ となる. 階層 (d): 逐次データ同化プロセスにより定義される厳密なフィルタ分布 $\mathbb{P}^{\widehat{V}_n}$ と \mathbb{P}^{V_n} を表しており, 階層 (c) の条件付き分布を再現している. 階層 (e): 逐次データ同化プロセスを近似するフィルタリングアルゴリズムを説明している. 近似された操作は $\tilde{\mathcal{P}}$ および \tilde{L}_{y_n} で表される.

フィルタ分布とスマージング分布の関係は次のように与えられる. 時刻 $n = N$ までの観測 y_N に関して, 状態 $v = (v_0, \dots, v_N) \in \mathbb{R}^{N_u \times (N+1)}$ のスマージング分布の PDF を $p_V(v | y_N)$ とする.

命題 3.8 ([63]). 離散時間区間 $0 \leq n \leq N$ におけるスマージング分布 $p_V(v | y_N)$ と時刻 $n = N$ でのフィルタ分布 $p_{V_N}(v_N | y_N)$ について, スマージング分布の v_N に関する周辺分布はフィルタ分布に等しい:

$$\int p_V(v | y_N) dv_0 \dots dv_{N-1} = p_{V_N}(v_N | y_N).$$

決定論的モデルダイナミクスの場合, 全観測を用いて初期状態を推定することは, 最終状態の推定を自動的に意味する.

命題 3.9 ([63]). 決定論的モデルダイナミクス (3.1) において, $Q = O$ の場合, 初期状態 v_0 のスマージング分布の押し出しは最終状態 v_N のフィルタ分布に等しい:

$$(\Psi^{(N)})_* \mathbb{P}^{V_0}(\cdot | y_N) = \mathbb{P}^{V_N}(\cdot | y_N).$$

ここで, $\Psi^{(N)}$ は Ψ の N 回の合成を表す.

次に, データ同化のベイズ定式化における頑健性, すなわち事後分布の観測データに対する連続依存性について概観する. 定義 2.32 の結果として, スマージング分布 $\mathbb{P}^{V_0}(\cdot | y_N)$

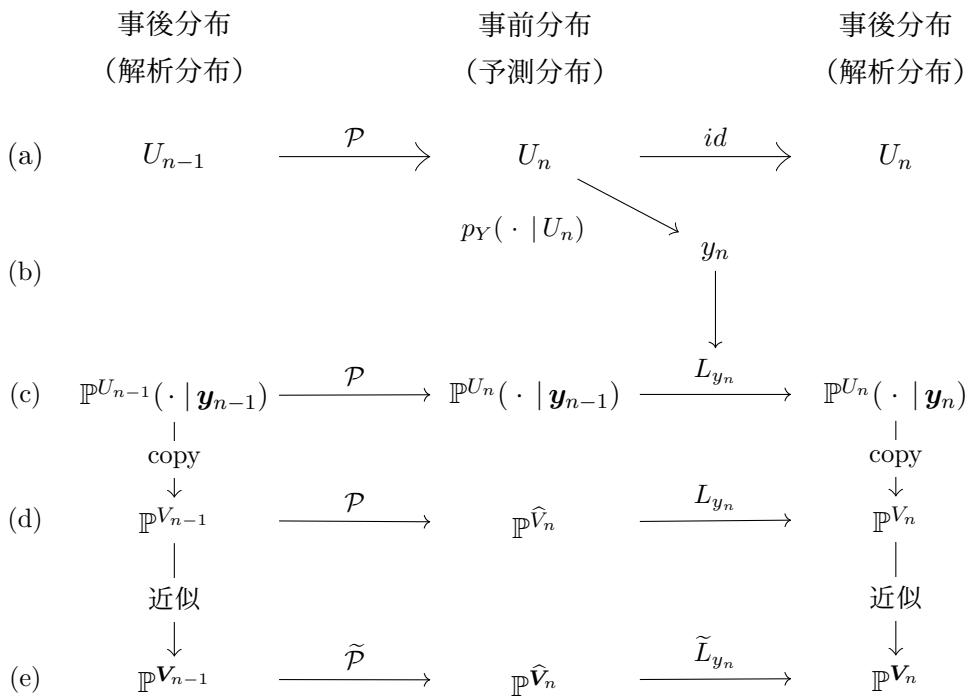


図 2: フィルタリング問題における階層. 階層は (a) モデルダイナミクス, (b) 観測, (c) 条件付き分布, (d) 厳密なフィルタ分布, 及び (e) 近似されたフィルタ分布である.

の適切性 (well-posedness) が、決定論的モデル ($Q = O$) と確率的モデル ($Q \neq O$) の両方について示される [63]. これらの結果および定義 2.33 から、初期状態の推定値の平均

$$\varpi_0 = \mathbb{E}[V_0 | \mathbf{y}_N]$$

は \mathbf{y}_N に関して連続であることが導かれる.

3.2 無限次元問題

3.2.1 離散時間状態空間モデル

無限次元ヒルベルト空間 \mathcal{H} を考える. 離散時間確率過程 $U : \mathbb{N} \times \Omega \rightarrow \mathcal{H}$ が次を満たすとする:

$$U_n = \Psi(U_{n-1}) + \xi_n. \quad (3.7)$$

ここで、 $U_0 \in \mathcal{H}$ は不確定な初期状態、 $\Psi : \mathcal{H} \rightarrow \mathcal{H}$ は連続写像、 $(\xi_n)_{n \in \mathbb{N}} \in \mathcal{H}$ は独立同分布 (i.i.d.) かつ平均 0 のノイズ列で、共分散 $Q \in \mathcal{L}_{sa}(\mathcal{H})$ が $Q \in \mathcal{K}_1(\mathcal{H})$ および $Q \succeq 0$ を満たすとする. 観測は次の形式の確率過程 $Y : \mathbb{N} \times \Omega \rightarrow \mathcal{Y}$ で与えられる:

$$Y_n = h(U_n) + \eta_n, \quad n \in \mathbb{N}. \quad (3.8)$$

ここで、 $h : \mathcal{H} \rightarrow \mathcal{Y}$ は連続写像、 $(\eta_n)_{n \in \mathbb{N}} \subset \mathcal{Y}$ は i.i.d. ノイズ列とする. 観測空間が有限次元の場合 ($\mathcal{Y} = \mathbb{R}^{N_y}$)、定義 2.29 および定義 2.32 における一般化ベイズ公式を用いて、事後分布 \mathbb{P}^{V_n} をそのラドン=ニコディム微分の形で定義できる. この場合、定義 3.1 のようなガウス尤度を考える. したがって、定義 3.6 に従い、ベイズデータ同化問題を考える.

定義 3.10 (無限次元状態空間における逐次データ同化). U および Y がそれぞれ (3.7) および (3.8) に従うとする. $\mathcal{Y} = \mathbb{R}^{N_y}$ かつ観測ノイズが定義 3.1 を満たすと仮定する. 次の逐次更新により、 $\mathbb{P}^{V_0} = \mathbb{P}^{U_0}$ から始めて、任意の $n \in \mathbb{N}$ に対して厳密なフィルタ分布 $\mathbb{P}^{V_n} = \mathbb{P}^{U_n}(\cdot | \mathbf{Y}_n)$ を得る:

(I) (予測: $\mathbb{P}^{V_{n-1}} \rightarrow \mathbb{P}^{\widehat{V}_n}$) モデルダイナミクスを用いて $\mathbb{P}^{V_{n-1}}$ を $\mathbb{P}^{\widehat{V}_n}$ に伝播する:

$$\mathbb{P}^{\widehat{V}_n}(dv) = \int_{\mathcal{H}} K(v_{n-1} dv) \mathbb{P}^{V_{n-1}}(dv_{n-1}), \quad (3.9)$$

ここで、 $K : \mathcal{H} \times \mathcal{B}(\mathcal{H}) \rightarrow [0, 1]$ は (3.7) に関連する遷移核である.

(II) (解析: $\mathbb{P}^{\widehat{V}_n}, y_n \rightarrow \mathbb{P}^{V_n}$) $\Phi(u; y) = \frac{1}{2}|y - h(u)|_R^2$ を用いて, 一般化ベイズ公式により次のように $\mathbb{P}^{V_n} \in \mathcal{M}_1(\mathcal{H})$ を定義する:

$$\frac{d\mathbb{P}^{V_n}}{d\mathbb{P}^{\widehat{V}_n}}(v) \propto \exp(-\Phi(v; y)). \quad (3.10)$$

無限次元状態空間においても, 定義 3.8 and 3.9 と同様の関係が成り立つ [16]. さらに, スムージング分布の well-posedness が [26] で示されている.

無限次元の観測を伴う状態空間モデルを考えることは簡単ではない. \mathcal{Y} が無限次元である場合 [58], ノイズ共分散 $R \succ 0$ の正定値性は $\text{Tr } R = \infty$ を意味する. したがって, 定義 2.26 より, R は \mathcal{Y} 上のガウス分布の共分散にはなり得ない. その結果, 定義 3.1 は適用できない.

一方, $\text{Tr } R < \infty$ の場合, R は可逆でなくなるため, $|\cdot|_R = |R^{-\frac{1}{2}} \cdot|$ は Cameron-Martin 空間 $\text{Ran}(R^{\frac{1}{2}})$ 上でのみ定義される. この場合, 事後分布 \mathbb{P}^{V_n} の正規化定数が 0 となり, 逐次的なベイズデータ同化は考えられない. 無限次元空間 \mathcal{H} と \mathcal{Y} の両方におけるデータ同化問題の詳細な定式化については [51] を参照.

ベイズ定式化の代わりに, 定義 3.4 で定義された状態 (信号) 推定問題を考えることができる. したがって, 節 4 でデータ同化アルゴリズムを定義するために使用される定義 3.1 に代わる次の仮定を導入する:

仮定 3.11. 任意の $n \in \mathbb{N}$ に対して, $\eta_n \sim \mathcal{N}(0, \tilde{R})$ とし, $\tilde{R} \in \mathcal{K}_1(\mathcal{Y})$, $\tilde{R} \succeq 0$, $\tilde{R} \preceq R$ ($R \succ 0$) を満たす.

3.2.2 連続時間状態空間モデル

連続時間定式化について, 無限次元空間での確率過程 (たとえば確率偏微分方程式の解) を扱うのを避けるため, 確率論的ケースは考えない. その代わりに, 無限次元力学系を考える. これを扱うために, まずヒルベルト空間 \mathcal{H} における発展方程式を導入する:

$$\frac{du}{dt} = \mathcal{F}(u). \quad (3.11)$$

その後, 次元が低い空間 \mathcal{Y} ($\dim(\mathcal{Y}) \leq \dim(\mathcal{H})$) でノイズのない観測を考える:

$$y(t) = h(u(t)), \quad (3.12)$$

ここで, $h : \mathcal{H} \rightarrow \mathcal{Y}$ は観測作用素である. ノイズのない状況におけるデータ同化問題は, 偏微分方程式のフィードバック制御に起因する [11]. 制御理論の観点から, 真の状態を再構築するために, シミュレーションされた状態にどのような有限次元の制御入力が必要か, またはその数を決定することが重要である. このような問題は, 特に非圧縮性 2 次

元 Navier-Stokes 方程式 [10] や非圧縮性 3 次元 Navier-Stokes-alpha 方程式 [2] について研究されている。この問題については節 5 でさらに議論する。

離散時間および連続時間、有限次元および無限次元、決定論的および確率論的な設定間の関係について以下に述べる。

注意 3.12. $Q \neq O$ の場合、(3.1) は連続時間ダイナミクスの離散化として用いられることが多い。モデル誤差 ξ_n は、時間間隔 $[t_{n-1}, t_n]$ および空間領域における離散化誤差の累積として解釈される [20]。理論の単純化のため、 ξ_n はしばしばガウスノイズであると仮定される。

注意 3.13. 多くの応用では、未知の真の状態は連続時間過程としてモデル化される。しかし、ノイズを含む観測は通常、時間間隔 $\tau > 0$ で離散時間ステップで得られる。図 3 を用いて、(3.3) と (3.1)，および (3.11) と (3.7) の関係を説明する。

決定論的ケース (3.11)（または (3.3) で $Q = O$ の場合）では、任意の $u_0 \in \mathcal{H}$ について一意な解が存在し、それが $t \geq 0$ についての半群 $\Psi_t : \mathcal{H} \rightarrow \mathcal{H}$ を生成すると仮定する。このとき、 $\Psi = \Psi_\tau$ および $U_n = u_{n\tau}$ とすれば、(3.7)（または (3.1) で $Q = O$ ）が得られる。詳細については [54, 82] を参照。

確率ダイナミクス (3.3) の場合、一意な解を \tilde{U}_t とするとき、 $U_0 = u$ から始めて、 $\Psi_t(u) = \mathbb{E}^u[\tilde{U}_t]$ 、 $\Psi = \Psi_\tau$ 、および $\xi_n = \tilde{U}_{n\tau} - \Psi(\tilde{U}_{n(\tau-1)})$ とすれば十分である。詳細は [88] も参照せよ。

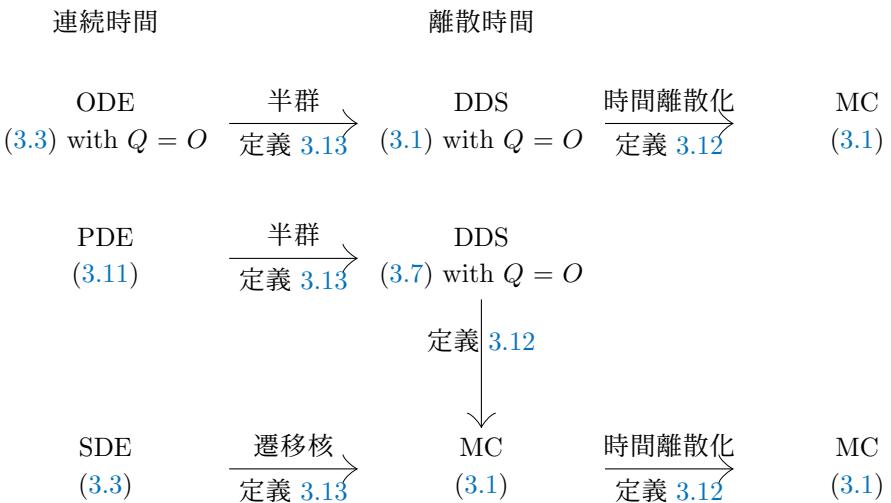


図 3: モデルダイナミクスのさまざまな定式化間の関係. 略語は次のように定義する:ODE: Ordinary Differential Equation (常微分方程式), DDS: Discrete Dynamical System (離散力学系), MC: Markov Chain (マルコフ連鎖), PDE: Partial Differential Equation (偏微分方程式), SDE: Stochastic Differential Equation (確率微分方程式).

4. フィルタリングアルゴリズム

4.1 カルマンフィルタ

状態空間と観測空間の次元を $N_u, N_y \in \mathbb{N}$ とし, $F \in \mathbb{R}^{N_u \times N_u}$ および $H \in \mathbb{R}^{N_y \times N_u}$ とする. 離散時間, 有限次元の線形ガウス系を次のように考える:

$$U_n = FU_{n-1} + \xi_n, \quad Y_n = HU_n + \eta_n, \quad (4.1)$$

ここで, $\xi_n \sim \mathcal{N}(0, Q)$ は誤差共分散行列 $Q \succeq 0$ を持つガウスノイズ, 観測ノイズ η_n は定義 3.1 を満たすとする. 線形ガウス系 (4.1) は (3.1) の特殊な場合に対応する. 初期状態の不確実性は $U_0 \sim \mathcal{N}(\varpi_0, P_0)$ に従うと仮定し, $\varpi_0 \in \mathbb{R}^{N_u}, P_0 \in \mathbb{R}^{N_u \times N_u}$ であり, $P_0 \succ 0$ とする. カルマンフィルタ (Kalman Filter, KF) は, Kalman によって初めて提案され [49], システム (4.1) におけるフィルタ (定義 3.7) を厳密かつ明示的に表現する.

定義 4.1 (KF). ガウス分布 $V_{n-1} \sim \mathcal{N}(\varpi_{n-1}, P_{n-1})$ が与えられているとする. カルマンフィルタ (KF) のアルゴリズムは以下の通り.

(I) (予測: $\varpi_{n-1}, P_{n-1} \rightarrow \hat{\varpi}_n, \hat{P}_n$) 平均と誤差共分散の時間発展を計算する:

$$\hat{\varpi}_n = F\varpi_{n-1}, \quad (4.2)$$

$$\hat{P}_n = FP_{n-1}F^* + Q. \quad (4.3)$$

(II) (更新: $\hat{\varpi}_n, \hat{P}_n, y_n \rightarrow \varpi_n, P_n$) 定義 2.28 を用いて事後分布の平均と誤差共分散を計算する:

$$\varpi_n = \hat{\varpi}_n + K_n(y_n - H\hat{\varpi}_n), \quad (4.4)$$

$$P_n = (I - K_nH)\hat{P}_n. \quad (4.5)$$

ここで, K_n はカルマンゲインを表し,

$$K_n = \hat{P}_nH^*(H\hat{P}_nH^* + R)^{-1}. \quad (4.6)$$

フィルタ分布は $V_n \sim \mathcal{N}(\varpi_n, P_n)$ として表される.

KF の性質を説明するために線形推定の概念を導入する.

定義 4.2 (線形推定). $t \in \mathcal{T}$ に対して, 確率変数 $V_n : \Omega \rightarrow \mathbb{R}^{N_u}$ が以下を満たすとき, 線形推定量と呼ぶ:

$$\mathbb{E}[|U_n - V_n|^2] = \inf\{|\tilde{V}_n - U_n|^2 \mid \tilde{V}_n \in \text{span}(\mathbf{Y}_n)\}.$$

命題 4.3 (KF [4, 24, 73]). 線形ガウス系 (4.1) に対して, 以下が成り立つ.

- (1) 厳密なフィルタ分布 $\mathbb{P}^{U_n}(\cdot | \mathbf{Y}_n)$ はガウス分布となる.
- (2) 線形推定量が最適推定量を達成する.
- (3) ガウス分布の平均と誤差共分散の逐次更新は定義 4.1 で与えられる.

式 (4.4) はカルマン更新として知られている. 以下の重要な補題は, Woodbury の公式 (定義 2.16) から得られる.

補題 4.4. 次の恒等式が成り立つ.

$$I_{N_u} - K_n H = (I_{N_u} + \hat{P}_n H^* R^{-1} H)^{-1}. \quad (4.7)$$

Proof. 定義 2.16 における (2.7) を, $A = I_{N_u}$, $B = \hat{P}_n H^*$, $C = H$, $D = R^{-1}$ として適用すると,

$$(I_{N_u} + \hat{P}_n H^* R^{-1} H)^{-1} = I_{N_u} - \hat{P}_n H^* (R + H \hat{P}_n H^*)^{-1} H = I_{N_u} - K_n H.$$

これが示すべき式である. \square

注意 4.5. $Q, P_0 \succ 0$ の場合, 全ての $n \in \mathbb{N}$ において $P_n, \hat{P}_n \succ 0$ が帰納法により従う. これは以下の誤差共分散行列の逆行列の計算により確認される. 式 (4.3) について, (2.7) における $A = Q$, $B = F$, $C = P_{n-1}$, $D = F^*$ を用いると次が得られる.

$$\hat{P}_n^{-1} = (Q + FP_{n-1}F^*)^{-1} = Q^{-1} - Q^{-1}F(P_{n-1}^{-1} + F^*Q^{-1}F)^{-1}F^*Q^{-1}.$$

式 (4.5) については, 定義 4.4 により次の式が得られる.

$$P_n^{-1} = \hat{P}_n^{-1}(I_{N_u} - K_n H)^{-1} = \hat{P}_n^{-1}(I_{N_u} + \hat{P}_n H^* R^{-1} H) = \hat{P}_n^{-1} + H^* R^{-1} H.$$

これらの等式により, \hat{P}_n^{-1} と P_n^{-1} を逐次計算することが可能であり, \hat{P}_n と P_n を直接評価する必要がない. 誤差共分散行列の逆行列は精度行列 (*precision matrices*) として知られている. 詳細については [63] を参照.

注意 4.6. Q が縮退していても, $R \succ 0$ の場合, KF は縮退した誤差共分散行列 P_n でも定義できる. さらに, 定義 4.1 は無限次元の線形ガウス系にも適用可能であり, 以下の仮定をおく.

- 線形モデル: $F \in \mathcal{L}(\mathcal{H})$,
- ガウスモデルノイズ: $\xi_n \sim \mathcal{N}(0, Q)$ であり, $Q \in \mathcal{L}_{sa}(\mathcal{H}), Q \succeq 0, \text{Tr } Q < \infty$,

- 線形観測作用素: $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$ であり, $\mathcal{Y} \subset \mathcal{H}$,
- ガウス観測ノイズ: η_n は定義 3.11 を満たす,
- ガウス初期不確実性: $U_0 \sim \mathcal{N}(\varpi_0, P_0)$ であり, $\varpi_0 \in \mathcal{H}, P_0 \in \mathcal{K}_1(\mathcal{H}), P_0 \succeq 0$.

注意 4.7 (KF の制約). KF は理論的には明快であるが, 高次元かつ複雑なシステム (例えば大気モデル) に適用する際には以下の 2つの重要な制約がある.

- KF は Ψ が線形であることを仮定するが, 大気力学系モデルは通常非線形である.
- 状態空間の次元が非常に大きくなる (例えば 10^9 次元) 場合がある. これにより, 誤差共分散行列は $10^9 \times 10^9$ サイズとなり, コンピュータメモリに格納するのは不可能である.

4.2 非線形力学系への拡張

この節では, Hilbert 空間 \mathcal{H} と \mathcal{Y} , および線形観測 $h(u) = Hu$ ($H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$) を持つ離散時間非線形力学系 (3.7) と (3.8) を考える. 観測ノイズは 定義 3.11 を満たすと仮定する. 初期の不確実性は $U_0 \sim \mathcal{N}(\varpi_0, P_0)$ であり, ここで $\varpi_0 \in \mathcal{H}, P_0 \in \mathcal{K}_1(\mathcal{H}), P_0 \succeq 0$ であるとする. 拡張カルマンフィルタ (Extended Kalman Filter, ExKF) は, 定義 4.7 に述べた制限事項 (1) を解消するために, Ψ の線形化を用いるカルマンフィルタ (KF) の拡張である.

定義 4.8 (ExKF). Ψ が Fréchet 微分可能であると仮定し, その点 $u \in \mathcal{H}$ での導関数を $D\Psi_u \in \mathcal{L}(\mathcal{H})$ とする. 事前分布 V_{n-1} が平均 ϖ_{n-1} , 共分散行列 P_{n-1} を持つ場合, 拡張カルマンフィルタ (ExKF) のアルゴリズムは以下の通りである.

(I) (予測: $\varpi_{n-1}, P_{n-1} \rightarrow \widehat{\varpi}_n, \widehat{P}_n$)

$$\widehat{\varpi}_n = \Psi(\varpi_{n-1}), \quad (4.8)$$

$$\widehat{P}_n = D\Psi_{\varpi_{n-1}} P_{n-1} D\Psi_{\varpi_{n-1}}^* + Q. \quad (4.9)$$

(II) (解析: $\widehat{\varpi}_n, \widehat{P}_n, y_n \rightarrow \varpi_n, P_n$)

$$\varpi_n = \widehat{\varpi}_n + K_n(y_n - H\widehat{\varpi}_n), \quad (4.10)$$

$$P_n = (I - K_n H) \widehat{P}_n \quad (4.11)$$

ここでカルマンゲインは次式で与えられる.

$$K_n = \widehat{P}_n H^* (H \widehat{P}_n H^* + R)^{-1}.$$

定義 4.8において、線形化された $D\Psi$ および $D\Psi^*$ はそれぞれ接線線形モデルおよび随伴モデルと呼ばれる。通常、高次元または複雑なモデルダイナミクスにおいて導関数 $D\Psi$ を数値的に計算することは困難である。

三次元変分法 (3DVar) は、非線形予測と固定されたモデル共分散を用いたカルマン更新を繰り返す、より簡単なアルゴリズムである。

定義 4.9 (3DVar). 定数のモデル共分散 $\hat{P}_n = P_0$ に対して、3DVar のアルゴリズムは以下の通りである。

(I) (予測: $\varpi_{n-1} \rightarrow \hat{\varpi}_n$)

$$\hat{\varpi}_n = \Psi(\varpi_{n-1}). \quad (4.12)$$

(II) (解析: $\hat{\varpi}_n, y_n \rightarrow \varpi_n$)

$$\varpi_n = \hat{\varpi}_n + K(y_n - H\hat{\varpi}_n) \quad (4.13)$$

ここでカルマンゲイン K は以下で与えられる。

$$K = P_0 H^* (H P_0 H^* + R)^{-1}.$$

定義 4.9 のステップ (II) において、カルマンゲイン K は時間 $n \in \mathbb{N}$ に依存せず、3DVar は最適内挿とも呼ばれる。3DVar は Ψ の導関数を計算する必要がないが、定義 4.7 に指摘された完全共分散行列の大規模ストレージ要求という問題は依然として抱えている。

4.3 アンサンブルカルマンフィルタ

前節と同じ状態空間モデルを考える。アンサンブルカルマンフィルタ (Ensemble Kalman Filter, EnKF) は、アンサンブル $V_n = [v_n^{(1)}, \dots, v_n^{(m)}] \in \mathcal{H}^m$ (サイズ $m \in \mathbb{N}$) の経験分布によってフィルタ分布を近似する。

$$\mathbb{P}^{V_n}(\cdot) \approx \frac{1}{m} \sum_{k=1}^m \delta_{v_n^{(k)}}(\cdot).$$

ここで $\delta_u(\cdot)$ は u におけるディラック測度を表す。同様に、予測ステップの分布 $\mathbb{P}^{\hat{V}_n}$ は次のように近似される。

$$\mathbb{P}^{\hat{V}_n}(\cdot) \approx \frac{1}{m} \sum_{k=1}^m \delta_{\hat{v}_n^{(k)}}(\cdot).$$

ここで $\widehat{\mathbf{V}}_n = [\widehat{v}_n^{(1)}, \dots, \widehat{v}_n^{(m)}] \in \mathcal{H}^m$ であり, 各アンサンブルメンバーは非線形ダイナミカルモデル (3.1) に従って進化する. モデル共分散はアンサンブル共分散 $\widehat{P}_n = \text{Cov}_m \widehat{\mathbf{V}}_n := \frac{1}{m-1} \sum_{k=1}^m (\widehat{v}_n^{(k)} - \bar{\widehat{v}}_n) \otimes (\widehat{v}_n^{(k)} - \bar{\widehat{v}}_n)$ によって近似され, $\bar{\widehat{v}}_n = \frac{1}{m} \sum_{k=1}^m \widehat{v}_n^{(k)}$ である. EnKF の変種は, アンサンブル \mathbf{V}_{n-1} から予測アンサンブル $\widehat{\mathbf{V}}_n$ を得るために同じ予測ステップを使用する. しかし, 各変種は, 予測アンサンブル $\widehat{\mathbf{V}}_n$ と観測データ y_n を用いて解析アンサンブル \mathbf{V}_n を生成する解析ステップにおいて異なるアプローチを採用する. EnKF の単純かつ確率的な実装として, 摂動観測 (Perturbed Observation: PO) 法が知られている [19].

定義 4.10 (PO). 初期アンサンブル $\mathbf{V}_0 = [v_0^{(k)}]_{k=1}^m \in \mathcal{H}^m$ を与える. 摂動観測 (PO) 法のアルゴリズムは以下の 2つのステップから成る. 任意の $n \in \mathbb{N}$ に対して,

(I) (予測: $\mathbf{V}_{n-1} \rightarrow \widehat{\mathbf{V}}_n$) 次を計算する.

$$\widehat{v}_n^{(k)} = \Psi(v_{n-1}^{(k)}) + \xi_n^{(k)}, \quad \xi_n^{(k)} \sim \mathcal{N}(0, Q), \quad k = 1, \dots, m, \quad (4.14)$$

そして $\widehat{\mathbf{V}}_n = [\widehat{v}_n^{(k)}]_{k=1}^m \in \mathcal{H}^m$ を設定する.

(II) (解析: $\widehat{\mathbf{V}}_n, y_n \rightarrow \mathbf{V}_n$) $\widehat{P}_n = \text{Cov}_m(\widehat{\mathbf{V}}_n)$ を設定し, ランダムな摂動を加えることで観測を複製する.

$$y_n^{(k)} = y_n + \eta_n^{(k)}, \quad \eta_n^{(k)} \sim \mathcal{N}(0, R), \quad k = 1, \dots, m, \quad (4.15)$$

そしてアンサンブルを更新する.

$$v_n^{(k)} = \widehat{v}_n^{(k)} + K_n(y_n^{(k)} - H\widehat{v}_n^{(k)}), \quad k = 1, \dots, m. \quad (4.16)$$

ここでカルマンゲイン K_n は

$$K_n = \widehat{P}_n H^* (H\widehat{P}_n H^* + R)^{-1} \quad (4.17)$$

によって与えられる. 最後に $\mathbf{V}_n = [v_n^{(k)}]_{k=1}^m \in \mathcal{H}^m$ を設定する. この $\widehat{\mathbf{V}}_n$ から \mathbf{V}_n への写像を $\mathbf{V}_n = \mathbf{V}_{PO}(\widehat{\mathbf{V}}_n; y_n, \widehat{P}_n)$ と表記する.

命題 4.11 (PO 法の well-definedness [51]). $R \succ 0$ と仮定すると, PO 法は well-defined である, すなわち $H\widehat{P}_n H^* + R$ は可逆である.

PO 法を直接 \widehat{P}_n を評価せずに実装することが可能である.

補題 4.12. PO 法の解析ステップ (II) は, \widehat{P}_n を評価せずに以下のステップで置き換えることができる.

(II') (解析: $\hat{\mathbf{V}}_n, y_n \rightarrow \mathbf{V}_n$) $\hat{\mathbf{V}}_n = \bar{v}_n \mathbf{1} + d\hat{\mathbf{V}}_n$ と分解する. $d\hat{\mathbf{Y}}_n = Hd\hat{\mathbf{V}}_n$ を設定する. $y_n^{(k)}$ および $v_n^{(k)}$ をステップ (II) と同様に定義し,

$$K_n = \frac{1}{m-1} d\hat{\mathbf{V}}_n d\hat{\mathbf{Y}}_n^* \left(\frac{1}{m-1} d\hat{\mathbf{Y}}_n d\hat{\mathbf{Y}}_n^* + R \right)^{-1} \quad (4.18)$$

を設定する. この写像を $\mathbf{V}_n = \mathbf{V}_{PO'}(\hat{\mathbf{V}}_n; y_n)$ と表記する.

Proof. 式 (4.18) は次の理由によって成り立つ.

$$\frac{1}{m-1} d\hat{\mathbf{V}}_n d\hat{\mathbf{Y}}_n^* = \frac{1}{m-1} d\hat{\mathbf{V}}_n d\hat{\mathbf{V}}_n^* H^* = \hat{P}_n H^*$$

かつ

$$\frac{1}{m-1} d\hat{\mathbf{Y}}_n d\hat{\mathbf{Y}}_n^* = \frac{1}{m-1} Hd\hat{\mathbf{V}}_n d\hat{\mathbf{V}}_n^* H^* = H\hat{P}_n H^*.$$

□

[91] で述べられているように、PO 法の解析ステップで人工的なノイズを加えると、解析分布の近似に追加の誤差が生じる. この問題を回避するため、アンサンブル平方根フィルタ (Ensemble Square Root Filter, ESRF) と呼ばれる EnKF の決定論的なバージョンが提案されている [5, 14, 91]. これらのアルゴリズムは、解析アンサンブルを決定論的に生成するために行列平方根を計算する. ESRF アルゴリズムの一つの実装として、アンサンブル変換カルマンフィルタ (Ensemble Transform Kalman Filter, ETKF) がある [14].

定義 4.13 (ETKF). 初期アンサンブル $\mathbf{V}_0 = [v_0^{(k)}]_{k=1}^m \in \mathcal{H}^m$ を与える. アンサンブル変換カルマンフィルタ (ETKF) のアルゴリズムは以下の 2 つのステップからなる.

(I) (予測: $\mathbf{V}_{n-1} \rightarrow \hat{\mathbf{V}}_n$) このステップは定義 4.10 の PO 法のステップ (I) と同じである.

(II) (解析: $\hat{\mathbf{V}}_n, y_n \rightarrow \mathbf{V}_n$) $\hat{\mathbf{V}}_n = \bar{v}_n \mathbf{1} + d\hat{\mathbf{V}}_n$ と分解し, $\hat{P}_n = \text{Cov}_m(d\hat{\mathbf{V}}_n)$ を設定する. 平均を更新する.

$$\bar{v}_n = \bar{v}_n + K_n(y_n - H\bar{v}_n) \quad (4.19)$$

ここで、カルマンゲイン $K_n = \hat{P}_n H^* (H\hat{P}_n H^* + R)^{-1}$ を用いる. 対称行列 $T_n \in \mathbb{R}^{m \times m}$ を次を満たすようにとる.

$$\frac{1}{m-1} d\hat{\mathbf{V}}_n T_n (d\hat{\mathbf{V}}_n T_n)^* = (I_{\mathcal{H}} - K_n H) \hat{P}_n, \quad (4.20)$$

そしてアンサンブル摂動を $d\mathbf{V}_n = d\hat{\mathbf{V}}_n T_n$ に変換する. 行列 T_n を変換行列と呼ぶ. 最後に解析アンサンブルを $\mathbf{V}_n = \bar{v}_n \mathbf{1} + d\mathbf{V}_n$ に設定する. この写像を $\mathbf{V}_n = \mathbf{V}_{ETKF}(\hat{\mathbf{V}}_n; y_n, \hat{P}_n)$ と表記する.

定理 4.1 (ETKF の well-defined 性 [82]). $R \succ 0$ と仮定する. 任意の $\hat{V}_n \in \mathcal{H}^m$ に対し、(4.20) を満たす一意な対称変換行列 $T_n \in \mathbb{R}^{m \times m}$ が存在する. この行列は次で与えられる.

$$T_n = \left(I_m + \frac{1}{m-1} d\hat{V}_n^* H^* R^{-1} H d\hat{V}_n \right)^{-\frac{1}{2}}. \quad (4.21)$$

定理 4.1 の証明のため、カルマンゲイン K_n の性質を調べる.

補題 4.14. $\hat{P}_n \succeq 0$, $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$, および $R \in \mathcal{L}(\mathcal{Y})$ で $R \succ 0$ とする. カルマンゲインは次を満たす：

$$K_n = (I_{\mathcal{H}} - K_n H) \hat{P}_n H^* R^{-1}. \quad (4.22)$$

また、(4.19) は次と同値である：

$$(I_{\mathcal{H}} + \hat{P}_n H^* R^{-1} H) \bar{v}_n = \bar{\bar{v}}_n + \hat{P}_n H^* R^{-1} y_n. \quad (4.23)$$

Proof. 簡単のため、以下の証明では時間インデックス n を省略する. $\hat{P} \succeq 0$ および $R \succ 0$ であることから、 $R + H\hat{P}H^* \succ 0$ が成り立ち、これが可逆であることがわかる. さらに、 $I_{\mathcal{Y}} + R^{-1}H\hat{P}H^* = R^{-1}(R + H\hat{P}H)$ も可逆である. これは、正定値作用素の積が正のスペクトルを持つことにより保証される（定義 2.5 参照）. (2.8) および可逆な A, B に対する $(AB)^{-1} = B^{-1}A^{-1}$ を用いると、以下を得る：

$$\begin{aligned} (R + H\hat{P}H^*)^{-1} &= (I_{\mathcal{Y}} + R^{-1}H\hat{P}H^*)^{-1}R^{-1} \\ &= [I_{\mathcal{Y}} - (I_{\mathcal{Y}} + R^{-1}H\hat{P}H^*)^{-1}R^{-1}H\hat{P}H^*]R^{-1} \\ &= [I_{\mathcal{Y}} - (R + H\hat{P}H^*)^{-1}H\hat{P}H^*]R^{-1}. \end{aligned}$$

これより次が成り立つ：

$$\begin{aligned} K &= \hat{P}H^*(R + H\hat{P}H^*)^{-1} = \hat{P}H^*[I_{\mathcal{Y}} - (R + H\hat{P}H^*)^{-1}H\hat{P}H^*]R^{-1} \\ &= (I_{\mathcal{H}} - KH)\hat{P}H^*R^{-1}, \end{aligned}$$

これは (4.22) を示している.

一方、(4.22) から次が従う：

$$\bar{v} = \bar{\bar{v}} + K(y - H\bar{v}) = (I_{\mathcal{H}} - KH)\bar{\bar{v}} + Ky = (I_{\mathcal{H}} - KH)\bar{\bar{v}} + (I_{\mathcal{H}} - KH)\hat{P}H^*R^{-1}y.$$

(4.23) を示すには、(4.6) を用いて $(I_{\mathcal{H}} + \hat{P}H^*R^{-1}H)(I_{\mathcal{H}} - KH) = I_{\mathcal{H}}$ を確認するだけで十分である. これは 定義 4.4 から従う. \square

証明 (定理 4.1). まず, (4.20) を満たす T_n の存在を示す. $d\mathbf{Y} = Hd\hat{\mathbf{V}}$ とする. このとき, 演算子 $R + \frac{1}{m-1}d\mathbf{Y}d\mathbf{Y}^* = R + H\hat{P}H^*$ は可逆である. 次に, 対称行列

$$S = I_m - \frac{1}{m-1}d\mathbf{Y}^* \left(R + H\hat{P}H^* \right)^{-1} d\mathbf{Y} \in \mathbb{R}^{m \times m}.$$

を考える. すると次が成り立つ:

$$\begin{aligned} \frac{1}{m-1}d\hat{\mathbf{V}}Sd\hat{\mathbf{V}}^* &= \frac{1}{m-1}d\hat{\mathbf{V}}d\hat{\mathbf{V}}^* - \frac{1}{m-1}d\hat{\mathbf{V}}d\mathbf{Y}^* \left(R + H\hat{P}H^* \right)^{-1} \frac{1}{m-1}d\mathbf{Y}d\hat{\mathbf{V}}^* \\ &= \hat{P} - \hat{P}H^*(R + H\hat{P}H^*)^{-1}H\hat{P} = (I - KH)\hat{P}, \end{aligned}$$

ここで $\frac{1}{m-1}d\mathbf{Y}d\hat{\mathbf{V}}^* = H\hat{P}$ を用いた. (2.11) より,

$$S = \left(I_m + \frac{1}{m-1}d\mathbf{Y}^*R^{-1}d\mathbf{Y} \right)^{-1} \quad (4.24)$$

が成り立ち, $S \succ 0$ が得られる. 最後に, 変換行列 $T = S^{\frac{1}{2}}$ を定義する. これはまさに (4.21) に他ならない. また, 定義より T は対称である. \square

定理 4.1 は, ETKF アルゴリズムの適切性 (well-definedness) の Hilbert 空間への拡張であり, これには無限次元空間も含まれる [60]. また, 定義 4.12 に示した ETKF の実用的な実装についても検討する.

補題 4.15. 以下の解析ステップは, ETKF のステップ (II) と等価である:

(II') (解析: $\hat{\mathbf{V}}_n, y_n \rightarrow \mathbf{V}_n$) $\hat{\mathbf{V}}_n = \bar{v}_n \mathbf{1} + d\hat{\mathbf{V}}_n$ と分解する. 修正された変換行列を次で定義する:

$$\tilde{T}_n = \frac{1}{m-1}T_n^2 d\hat{\mathbf{V}}_n^* H^* R^{-1} (y_n - H\bar{v}_n) \mathbf{1} + T_n, \quad (4.25)$$

そして, 次のように変換する:

$$\mathbf{V}_n = \bar{v}_n \mathbf{1} + d\hat{\mathbf{V}}_n \tilde{T}_n.$$

この写像を $\mathbf{V}_n = \mathbf{V}_{ETKF'}(\hat{\mathbf{V}}_n; y_n)$ と表記する.

この代替ステップでは, \hat{P}_n を明示的に評価する必要がないため, 実用的な数値計算において冗長なメモリ割り当てを回避できることに注意する.

Proof. (4.25) の第 1 項に $d\hat{\mathbf{V}}$ を掛けることで, T_n の定義より以下が成り立つ:

$$\begin{aligned} d\hat{\mathbf{V}}_n \frac{1}{m-1}T_n^2 d\hat{\mathbf{V}}_n^* H^* R^{-1} (y_n - H\bar{v}_n) &= (I_H - K_n H) \hat{P}_n H^* R^{-1} (y_n - H\bar{v}_n) \\ &= K_n (y_n - H\bar{v}_n). \end{aligned}$$

ここで, 最後の等号はカルマンゲインの関係式 (4.22) から従う. \square

もう一つの ESRF の実装として、アンサンブル調整カルマンフィルタ (Ensemble Adjustment Kalman Filter, EAKF) [5] が存在する。

定義 4.16 (EAKF). 初期アンサンブルを $V_0 = [v_0^{(k)}]_{k=1}^m \in \mathcal{H}^m$ とする。アンサンブル調整カルマンフィルタ (EAKF) のアルゴリズムは次のように与えられる。

(I) (予測: $V_{n-1} \rightarrow \hat{V}_n$) このステップは、PO 法 (定義 4.10) のステップ (I) と同じである。

(II) (解析: $\hat{V}_n, y_n \rightarrow V_n$) $\hat{V}_n = \bar{v}_n \mathbf{1} + d\hat{V}_n$ と分解し、 $\hat{P}_n = \text{Cov}_m(\hat{V}_n)$ を設定する。平均を次で更新する：

$$\bar{v}_n = \bar{\hat{v}}_n + K_n(y_n - H\bar{\hat{v}}_n), \quad (4.26)$$

ここでカルマンゲイン K_n は次で与えられる：

$$K_n = \hat{P}_n H^*(H\hat{P}_n H^* + R)^{-1}. \quad (4.27)$$

適切な調整作用素 (*adjustment operator*) $A_n \in \mathcal{L}(\mathcal{H})$ を選び、次を満たすようにする：

$$\frac{1}{m-1} A_n d\hat{V}_n (A_n d\hat{V}_n)^* = (I_{\mathcal{H}} - K_n H) \hat{P}_n, \quad (4.28)$$

そして、アンサンブル摂動を $dV_n = A_n d\hat{V}_n$ として変換する。最後に解析アンサンブルを $V_n = \bar{v}_n \mathbf{1} + dV_n$ と設定する。この写像を $V_n = V_{EAKF}(\hat{V}_n; y_n, \hat{P}_n)$ と表記する。

調整作用素 A_n の存在は次の定理によって保証される。

定理 4.2 (EAKF の well-defined 性). $R \succ 0$ と仮定する。任意の $\hat{V}_n \in \mathcal{H}^m$ に対して、(4.28) を満たす調整作用素 $A_n \in \mathcal{L}(\mathcal{H})$ が存在する。それは次のように与えられる：

$$A_n = \Phi \Sigma E(I_{\kappa} + \Lambda)^{-\frac{1}{2}} \Sigma^{-1} \Phi^*. \quad (4.29)$$

ここで、 $\kappa = \text{rank } \hat{P}_n$, $\Sigma = \text{diag}(s_1, \dots, s_{\kappa})$ は \hat{P}_n のゼロでない固有値の平方根を降順に並べた対角行列, $\Phi = [\phi_1, \dots, \phi_{\kappa}] \in \mathcal{H}^{\kappa}$ は対応する固有ベクトルであり、次を満たす：

$$\hat{P}_n = \Phi \Sigma^2 \Phi^*. \quad (4.30)$$

また、対角行列 $\Lambda \in \mathbb{R}^{\kappa \times \kappa}$ と行列 $E \in \mathbb{R}^{\kappa \times \kappa}$ は対称行列 $\Sigma \Phi^* H^* R^{-1} H \Phi \Sigma \in \mathbb{R}^{\kappa \times \kappa}$ の固有値および固有ベクトルから成り、次を満たす：

$$\Sigma \Phi^* H^* R^{-1} H \Phi \Sigma = E \Lambda E^*. \quad (4.31)$$

EAKF の解析ステップ (II) の簡略版は以下の通りである.

補題 4.17. 以下の解析ステップは定義 4.16 における EAKF のステップ (II) と同値である.

(II') (解析: $\hat{\mathbf{V}}_n, y_n \rightarrow \mathbf{V}_n$) $\hat{\mathbf{V}}_n = \bar{\hat{v}}_n \mathbf{1} + d\hat{\mathbf{V}}_n$ と分解する. 平均を次のように更新する:

$$\bar{v}_n = \bar{\hat{v}}_n + K_n(y_n - H\bar{\hat{v}}_n)$$

ここで, Kalman ゲインは (4.18) で与えられる. 次に, 特異値分解を適用する:

$$\frac{1}{\sqrt{m-1}}d\hat{\mathbf{V}} = \Phi\Sigma\tilde{E}^*, \quad (4.32)$$

ここで, $\tilde{E} \in \mathbb{R}^{\kappa \times \kappa}$ はユニタリ行列である. また, 固有値分解を適用する:

$$\frac{1}{m-1}d\hat{\mathbf{V}}^*H^*R^{-1}Hd\hat{\mathbf{V}} = E'\Lambda'(E')^*, \quad (4.33)$$

ここで, $E' \in \mathbb{R}^{\kappa \times \kappa}$ はユニタリ行列である. 調整作用素を以下のように定義する:

$$\tilde{A}_n = \Phi\Sigma\tilde{E}^+E'(I_\kappa + \Lambda')^{-\frac{1}{2}}\Sigma^{-1}\Phi^*. \quad (4.34)$$

このアルゴリズムの残りの部分は定義 4.16 のステップ (II) と同じであり, ただし $A_n = \tilde{A}_n$ とする. この写像を $\mathbf{V}_n = \mathbf{V}_{EAKF'}(\hat{\mathbf{V}}_n; y_n)$ と表記する.

EnKF は, Ψ の導関数を計算することなく, 共分散の低ランク近似を与える. しかし, 有限サイズのアンサンブルを使用することに起因する別の問題が依然として残る.

注意 4.18. 予測共分散 \hat{P}_n が m 個のベクトルからなるアンサンブルで近似されるため, そのランクは次のように制限される:

$$\text{rank } \hat{P}_n \leq m-1.$$

さらに, 解析アンサンブルは予測アンサンブルによって張られる部分空間内に存在する. これは, EnKF およびそれに関連するアルゴリズム [78, 47, 90] に共通して現れるいわゆる部分空間特性 (Subspace property) と呼ばれる. この制約に対する数値的アプローチについては, 節 4.4 で議論されている.

4.4 EnKF に対する数値テクニック

分析ステップにおける予測共分散 \hat{P}_n の選択は, フィルタリングアルゴリズムにおいて重要である. EnKF では, \hat{P}_n は前の時刻からの分析共分散 P_{n-1} を伝播させることで

近似される。実際の数値計算においては、高次元状態空間の不確実な方向において、限定的なアンサンブルサイズ $m \ll N_u$ に起因して \hat{P}_n が過小評価されることがあり、これが状態推定の精度低下を招く。この問題を解決するために、共分散インフレーション技術が導入され、追加のパラメータ α を用いてデータ同化アルゴリズムを拡張する。共分散インフレーションのアイデアは、分析ステップの前に \hat{P}_n を膨張させることである [8, 91, 69]。インフレーションの導入方法は、フィルタリングアルゴリズムによって異なり、以下に示される。

定義 4.19 (PO 法における加法的インフレーション). $\alpha \geq 0$ とする。定義 4.10 の分析ステップ (II) において、インフレーションされた共分散 $\hat{P}_n^\alpha = \hat{P}_n + \alpha^2 I_{\mathcal{H}}$ を定義し、 $\mathbf{V}_n = \mathbf{V}_{PO}(\hat{\mathbf{V}}_n; y_n, \hat{P}_n^\alpha)$ を計算する。

このアプローチは共分散の加法的インフレーションと呼ばれ、 α はインフレーションパラメータと呼ばれる。

定義 4.20 (EnKF における乗法的インフレーション). $\alpha \geq 1$ とする。EnKF アルゴリズムの分析ステップにおいて、以下のように乗法的インフレーションを導入する。

- (1) 定義 4.10 の (II) において、インフレーションされた共分散 $\hat{P}_n^\alpha = \alpha^2 \hat{P}_n$ を定義し、 $\mathbf{V}_n = \mathbf{V}_{PO}(\hat{\mathbf{V}}_n; y_n, \hat{P}_n^\alpha)$ を計算する。
- (1') 定義 4.10 の (II) において、インフレーションされた摂動 $d\hat{\mathbf{V}}_n^\alpha = \alpha d\hat{\mathbf{V}}_n$ とアンサンブル $\hat{\mathbf{V}}^\alpha = \bar{v}_n \mathbf{1} + d\hat{\mathbf{V}}^\alpha$ 、共分散 $\hat{P}_n^\alpha = \text{Cov}_m(d\hat{\mathbf{V}}_n^\alpha)$ を定義し、 $\mathbf{V}_n = \mathbf{V}_{PO}(\hat{\mathbf{V}}_n^\alpha; y_n, \hat{P}_n^\alpha)$ を計算する。同時に、 $\mathbf{V}_n = \mathbf{V}_{PO'}(\hat{\mathbf{V}}_n^\alpha; y_n)$ を定義 4.12 に基づいて計算することもできる。
- (2) 定義 4.13 の (II) において、 $d\hat{\mathbf{V}}_n^\alpha, \hat{P}_n^\alpha$ を (1') に従って用い、 $\mathbf{V}_n = \mathbf{V}_{ETKF}(\hat{\mathbf{V}}_n^\alpha; y_n, \hat{P}_n^\alpha)$ を計算する。同時に、 $\mathbf{V}_n = \mathbf{V}_{ETKF'}(\hat{\mathbf{V}}_n^\alpha; y_n)$ を定義 4.15 に基づいて計算することもできる。
- (3) 定義 4.16 の (II) において、 $d\hat{\mathbf{V}}_n^\alpha, \hat{P}_n^\alpha$ を (1') に従って用い、 $\mathbf{V}_n = \mathbf{V}_{EAKF}(\hat{\mathbf{V}}_n^\alpha; y_n, \hat{P}_n^\alpha)$ を計算する。同時に、 $\mathbf{V}_n = \mathbf{V}_{EAKF'}(\hat{\mathbf{V}}_n^\alpha; y_n)$ を定義 4.17 に基づいて計算することもできる。

注意 4.21. 加法的インフレーションは、直接的に \hat{P} のランクを改善し、 \hat{P}^α がフルランクであることを保証する。しかし、乗法的インフレーションはこれを保証しない。代わりに、定義 4.20 の (1'), (2), (3) におけるアンサンブルの乗法的インフレーションは、分析ステップにおける収縮の前にアンサンブル $\hat{\mathbf{V}}_n^\alpha$ を膨張させるため、連続的なデータ同化プロセスにおいてアンサンブル共分散のランクを維持する可能性がある。この観察は、ESRF の誤差評価の際に重要な役割を果たし、節 7 において議論される。

注意 4.22. ETKF および EAKF における予測アンサンブルと分析アンサンブルの関係を以下に要約する。平均の更新 (4.23) は以下で与えられる：

$$(I_{\mathcal{H}} + \alpha^2 \hat{P}_n H^* R^{-1} H) \bar{v}_n = \bar{v}_n + \alpha^2 \hat{P}_n H^* R^{-1} y_n, \quad (4.35)$$

また、分析共分散は以下を満たす：

$$P_n = \frac{\alpha^2}{m-1} d\hat{V}_n (I_m + \alpha^2 d\hat{V}_n^* H^* R^{-1} H d\hat{V}_n)^{-1} d\hat{V}_n^*. \quad (4.36)$$

PO 法の場合、これらの等式は条件付き期待値の意味においてのみ成り立つ。ここで、 \hat{V}_n と y_n が与えられているとする。

共分散インフレーション技術は、乗法的インフレーションを伴う ESRF における状態推定誤差を実用的に改善する [69, 91]。また、加法的インフレーションを伴う PO 法においても改善が報告されている [54]。大規模な大気モデルでは、大きな α が必要になることが多く、 α の手動調整は計算コストが高い [45]。これを回避するため、適応的な調整アルゴリズムが開発されている [6, 7, 72]。別のアプローチでは、観測値とシミュレーションされた（または予測された）観測値との差である残差に注目し、残差が所定の範囲内に収まることを保証するための乗法的インフレーション α の条件を導出している [68]。

関連する数値技術として、緩和法 (relaxation-to-prior methods) がある。これらは分析ステップにおけるアンサンブルの収縮を緩和する。例えば、予測摂動の緩和 (relaxation to prior perturbation, RTPP) 法 [94] は、予測アンサンブル摂動と分析アンサンブル摂動を $\alpha \in [0, 1]$ の比率で補間する：

$$d\mathbf{V}_{RTPP}^\alpha = \alpha d\hat{\mathbf{V}} + (1 - \alpha) d\mathbf{V}.$$

また、予測スプレッドの緩和 (relaxation to prior spread, RTPS) 法 [92] は、状態空間の各成分について決定される係数を掛けることで分析アンサンブルスプレッドの収縮を緩和する：

$$d\mathbf{V}_{RTPS}^\alpha[i] = \alpha_i d\mathbf{V}[i], \quad \alpha_i = \frac{\alpha |d\hat{\mathbf{V}}[i]|_2 + (1 - \alpha) |d\mathbf{V}[i]|_2}{|d\mathbf{V}[i]|_2},$$

ここで、 $\mathbf{V}[i] = [(v^{(k)})^i]_{k=1}^m$ は $\mathbf{V} \in \mathbb{R}^{N_u \times m}$ における各ベクトルの i 番目の成分のアンサンブルを表す。RTPP および RTPS の適応的な調整手法も提案されている [57, 93]。

共分散の過小評価を回避するもう一つの技術として、局所化 (localization) と呼ばれる方法がある [40, 46, 91]。その中心的なアイデアは、分析ステップにおいて、各状態変数から遠く離れた領域の観測値を無視することである。このアプローチを使用したよく知られたアルゴリズムが、局所アンサンブル変換カルマンフィルタ (Local Ensemble Transform Kalman Filter, LETKF) である [46]。局所化技術は EnKF において重要な手法であるが、本論文では局所化を伴う数学解析は扱っていない。詳細は [31, 87, 90] を参照されたい。

4.5 連続時間アルゴリズム

(3.3) と (3.4) に対する KF の連続版は、カルマン・ブーシーフィルタ (Kalman Bucy Filter, KBF) と呼ばれる [73, 80]. (3.3) および (3.4) に対して、連続時間かつ有限次元の線形系を次の形で考える.

$$\mathcal{F}(u) = Fu, \quad h(u) = Hu, \quad (4.37)$$

ここで、 $F \in \mathbb{R}^{N_u \times N_u}, H \in \mathbb{R}^{N_y \times N_u}$, $N_u, N_y \in \mathbb{N}$ であり、初期の不確実性は $\varpi_0 \in \mathcal{H}, P_0 \in \mathbb{R}^{N_u \times N_u}, P_0 \succeq 0$ を持つ正規分布 $U_0 \sim \mathcal{N}(\varpi_0, P_0)$ とする.

命題 4.23 (カルマン・ブーシーフィルタ [73, 80]). (3.4) および (4.37) における観測ノイズが定義 3.3 を満たすとする. このとき、 $V_t = \mathbb{E}[U_t | \mathcal{F}_t^Y]$ は定義 4.24 の式を満たす.

定義 4.24 (カルマン・ブーシーフィルタ). カルマン・ブーシーフィルタ (KBF) は、以下の式によって定義される.

$$\begin{aligned} dV_t &= (F - P_t H^* R^{-1} H) V_t dt + P_t H^* R^{-1} dY_t \\ &= FV_t dt + K_t (dY_t - HV_t dt), \end{aligned}$$

$$V_0 = \varpi_0,$$

ただし、 $K_t = P_t H^* (HP_t H^* + R)^{-1}$ であり、 V_t の共分散 P_t は以下を満たす.

$$\frac{dP_t}{dt} = FP_t + P_t F^* + Q - P_t H^* R^{-1} H P_t = FP_t + P_t F^* + Q - K_t R K_t^*.$$

EnKF の連続時間拡張版も提案されており、その詳細は [13] にまとめられている. ここでは、非線形系 (3.3) と (3.4) に利用可能な決定論的バージョンであるアンサンブル・カルマン・ブーシーフィルター (EnKBF) を紹介する.

定義 4.25 (アンサンブル・カルマン・ブーシーフィルター [12]). モデルダイナミクスが (3.3) に従い、 $V_0 \in \mathcal{H}^m$ とする. (3.4) の観測ノイズが定義 3.3 を満たしていると仮定する. このとき、EnKBF は次式で与えられる:

$$dV_t^{(k)} = \mathcal{F}(V_t^{(k)}) dt + Q P_t^\dagger (V_t^{(k)} - \bar{v}_t) dt - \frac{1}{2} \tilde{P}_t R^{-1} (h(V_t^{(k)}) dt + \bar{h}_t dt - 2 dY_t)$$

ここで、 $k = 1, \dots, m$, P_t^\dagger は P_t の疑似逆行列を表し、

$$\begin{aligned} \bar{v}_t &= \frac{1}{m} \sum_{k=1}^m V_t^{(k)}, \quad P_t = \frac{1}{m-1} \sum_{k=1}^m (V_t^{(k)} - \bar{v}_t) \otimes (V_t^{(k)} - \bar{v}_t), \\ \bar{h}_t &= \frac{1}{m} \sum_{k=1}^m h(V_t^{(k)}), \quad \tilde{P}_t = \frac{1}{m-1} \sum_{k=1}^m (V_t^{(k)} - \bar{v}_t) \otimes (h(V_t^{(k)}) - \bar{h}_t). \end{aligned}$$

5. 散逸力学系

発展方程式 (3.11) としてヒルベルト空間 \mathcal{H} 上の散逸的な力学系に注目し、特に大気および海洋モデルに現れるカオス的な系を対象とする。流体力学における力学系のさらなる例については、[38] を参照されたい。

5.1 散逸力学系とその例

データ同化アルゴリズムの数学解析において、軌道が有界であることを保証し、小さな初期摂動を持つ二つの軌道間の誤差の時間発展を評価することは重要である。以下では、力学系が示す解吸性を特徴付けるために、決定論的モデルダイナミクス (3.11) に関する二つの仮定を考える。

仮定 5.1. 発展方程式 (3.11) は任意の $u_0 \in \mathcal{H}$ に対して一意な解を持ち、これにより半群 $\Psi_t : \mathcal{H} \rightarrow \mathcal{H}$ が生成される。さらに、ある $\rho > 0$ が存在し、 Ψ_t は吸収的な球 $B(\rho) = \{v \in \mathcal{H} \mid |v| \leq \rho\}$ を持つ。すなわち、任意の $v \in B(\rho)$ および $t \geq 0$ について $\Psi_t(v) \in B(\rho)$ が成り立つ。

仮定 5.2. ある $\beta \in \mathbb{R}$ が存在し、任意の $u \in B(\rho)$ および $v \in \mathcal{H}$ について

$$\langle \mathcal{F}(u) - \mathcal{F}(v), u - v \rangle \leq \beta |u - v|^2 \quad (5.1)$$

が成り立つ。

Kelly ら [54] は、周期境界条件付きの二次元 Navier-Stokes 方程式に適用するため、(5.1) に加えて \mathcal{F} に関する特定の仮定を課している。定義 5.2 は、力学系に沿った誤差の成長を評価するための以下の補題を導く。

補題 5.3 (誤差成長の上限 [54, 82]). 定義 5.1 および定義 5.2 が成り立つと仮定する。このとき、任意の $u \in B(\rho), v \in \mathcal{H}$ および $t > 0$ に対して、

$$|\Psi_t(u) - \Psi_t(v)| \leq e^{\beta t} |u - v| \quad (5.2)$$

が成り立つ。

この補題より、 β は吸収的な球内での誤差成長率（の上限）として解釈できる。 $\beta < 0$ の場合、摂動は指数関数的に速く収束し、力学系がカオス的でないことを示している。なお、定義 5.3 は初期誤差 $|u - v|$ の規模についての仮定を含まない。

以下の誤差解析では、 u と v はそれぞれ真の状態と解析値を表す。解析値としてアンサンブル平均を考える場合、誤差成長を評価するために別の不等式が必要である。そのため、以下のより強い仮定を考える。

仮定 5.4. 定義 5.1 における $\rho > 0$ および (3.11) の \mathcal{F} に対して, ある $\beta > 0$ が存在し, 任意の $u, v \in B(\rho)$ について

$$|\mathcal{F}(u) - \mathcal{F}(v)| \leq \beta |u - v|$$

が成り立つ.

補題 5.5 (アンサンブル平均の誤差成長の上限 [82]). 定義 5.1 および定義 5.4 が成り立つ, $u_0 \in B(\rho)$ および $v_0^{(n)} \in B(\rho)$ ($n = 1, \dots, N$) とする. $u_t = \Psi_t(u_0)$ および $\bar{v}_t = \frac{1}{N} \sum_{n=1}^N \Psi_t(v_0^{(n)})$ を定義する. このとき, 任意の $\epsilon > 0$ および $t > 0$ に対して,

$$|\bar{v}_t - u_t|^2 \leq e^{2(\beta+\epsilon)t} (|\bar{v}_0 - u_0|^2 + D) - D \quad (5.3)$$

が成り立つ. ここで, $D = \frac{\beta^2 \rho^2}{(\beta+\epsilon)\epsilon}$ である.

5.1.1 有限次元系の例

有限次元の散逸力学系の例を 2 つ挙げる. まず, ローレンツ 63 方程式 (L63) は, 次のような 3 次元非線形常微分方程式で与えられる:

$$\frac{dx}{dt} = -\sigma x + \sigma y, \quad (5.4a)$$

$$\frac{dy}{dt} = \varrho x - y - xz, \quad (5.4b)$$

$$\frac{dz}{dt} = -bz + xy. \quad (5.4c)$$

ここで, $\sigma > 0$, $b > 1$, $\varrho > 0$.

この方程式は, もともと Lorenz によって提案された [65]. 散逸性を評価するために, [85] では, L63 方程式の並行移動版が考えられている. 変数変換 $(x, y, z) \mapsto (x, y, z - \varrho - \sigma)$ を用いると, (5.4) は以下の形に変形される:

$$\frac{dx}{dt} = -\sigma x + \sigma y, \quad (5.5a)$$

$$\frac{dy}{dt} = -\sigma x - y - xz, \quad (5.5b)$$

$$\frac{dz}{dt} = -bz + xy - b(\varrho + \sigma). \quad (5.5c)$$

L63 方程式 (5.5) は, 次のように力学系の仮定を満たす.

命題 5.6 ([41, 85]). $\mathcal{H} = \mathbb{R}^3$ とし, $u = (x, y, z)^* \in \mathbb{R}^3$ とする. このとき, 任意の $u_0 \in \mathcal{H}$ に対して, (5.5) の初期条件 $u(0) = u_0$ を持つ一意な解 $u(t) \in \mathcal{H}$ が存在する. L63 方程式は, $\rho = \frac{b(\varrho+\sigma)}{\sqrt{4(b-1)}}$ を満たす定義 5.1, $\beta = 2\rho - 1$ を満たす定義 5.2, およびある $\beta > 0$ を満たす定義 5.4 を満たす. さらに, グローバルアトラクター \mathcal{A} が存在する.

図 4 は、典型的なパラメータ $\sigma = 10, b = \frac{8}{3}, \varrho = 28$ を持つ L63 方程式のグローバルアトラクターの 2 成分への射影を示している。L63 方程式の軌道のインタラクティブなアニメーションを以下のウェブサイトで確認できる：<https://kotatakedo.github.io/lorenz-webgl> [81]。

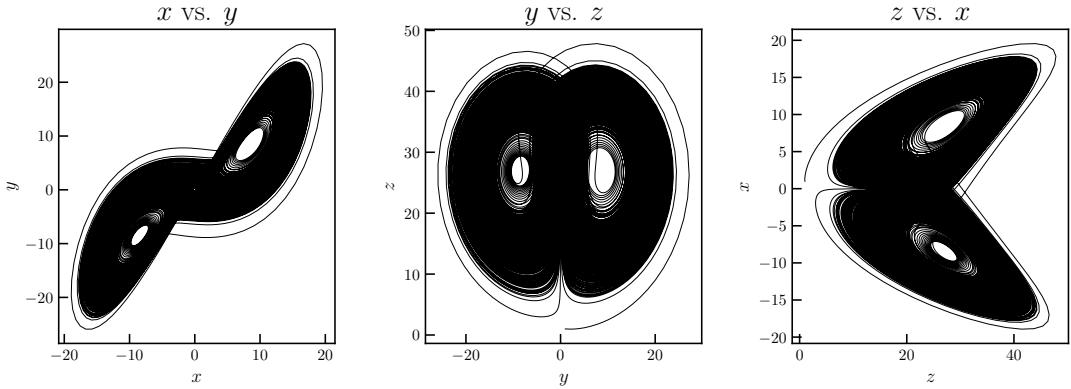


図 4: $\sigma = 10, b = \frac{8}{3}, \varrho = 28$ を持つ L63 方程式のグローバルアトラクター。

もう一つの重要な例は、気象学における空間的に拡張されたカオス的現象モデルであるローレンツ 96 (L96) 方程式である [66, 67]。L96 方程式もまた、仮定を満たす。

例 5.7. 与えられた成分数 $J \in \mathbb{N}$ に対して、状態ベクトル $u = (u^1, \dots, u^J)^* \in \mathbb{R}^J$ に対する L96 方程式は次で与えられる：

$$\frac{du^i}{dt} = (u^{i+1} - u^{i-2})u^{i-1} - u^i + f, \quad i = 1, \dots, J. \quad (5.6)$$

ここで、 $u^{-1} = u^{J-1}, u^0 = u^J, u^{J+1} = u^1, f \in \mathbb{R}$ は外力項である。

命題 5.8 ([62]). $\mathcal{H} = \mathbb{R}^J$ とする。このとき、任意の $u_0 \in \mathcal{H}$ に対して、(5.6) の初期条件 $u(0) = u_0$ を持つ一意な解 $u(t) \in \mathcal{H}$ が存在する。L96 方程式は、 $\rho = \sqrt{2J}|f|$ を満たす定義 5.1, $\beta = 2\rho - 1$ を満たす定義 5.2, および $\beta > 0$ を満たす定義 5.4 を満たす。さらに、グローバルアトラクター \mathcal{A} が存在する。

5.1.2 無限次元系の例

Ω を \mathbb{R}^d の開部分集合 ($d \in \mathbb{N}$) とする。 $1 \leq p \leq \infty$ の場合、 $L^p(\Omega)$ は Ω 上の L^p 関数の空間を表す。同様に、 $n \in \mathbb{N}$ かつ $1 \leq p \leq \infty$ の場合、 $W^{n,p}(\Omega)$ は Ω 上の L^p 関数で、その k 階弱微分が $L^p(\Omega)$ に属するものの空間 ($k = 1, \dots, n$) を表す。特に、 $H^n(\Omega) = W^{n,2}(\Omega)$

とする. 標準的な L^2 および H^1 ノルムをそれぞれ $|\cdot|$ および $\|\cdot\|$ で表す. さらに, バナッハ空間 \mathcal{X} に対して, $1 \leq p \leq \infty$ (または $n \in \mathbb{N}$), $-\infty \leq a < b \leq \infty$ の場合, $L^p(a, b; \mathcal{X})$ (または $H^n(a, b; \mathcal{X})$) は (a, b) から \mathcal{X} への L^p (または H^n) 関数の空間を表す. 同様に, $-\infty < a < b < \infty$ の場合, $[a, b]$ から \mathcal{X} への連続関数の空間を $C([a, b]; \mathcal{X})$ で表す.

2次元 Navier-Stokes 方程式は以下のように与えられる.

例 5.9 (2次元 Navier-Stokes 方程式). $L > 0$ とし, $\Omega = [0, L]^2$ とする. \mathcal{V} を, $\nabla \cdot \phi = 0$ かつ $\int_{\Omega} \phi = 0$ を満たす $\phi : \Omega \rightarrow \mathbb{R}^2$ のベクトル値 L 周期三角多項式の集合とする. \mathcal{V} を L^2 ノルムについて閉包したものを $\mathcal{H} = \overline{\mathcal{V}}^{|\cdot|}$ とし, H^1 ノルムについて閉包したものを $\mathcal{V} = \overline{\mathcal{V}}^{\|\cdot\|}$ と定義する. Leray-Helmholtz 射影 $P_{\mathcal{H}}$, すなわち L^2 直交射影 $P_{\mathcal{H}} : L^2(\Omega) \rightarrow \mathcal{H}$ を考える. [41, 54, 59, 85] の記法を用いると, 周期境界条件付きの非圧縮性の 2次元 Navier-Stokes 方程式は以下のように与えられる.

$$\frac{du}{dt} + \mathcal{A}u + \mathcal{B}(u, u) = f. \quad (5.7)$$

ここで, 非有界線形作用素 $\mathcal{A} : \mathcal{H} \rightarrow \mathcal{V}^*$ は $\mathcal{A} = -\nu \Delta$ と定義され, 対称双線形作用素 $\mathcal{B} : \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}^*$ は $\mathcal{B}(u, v) = \frac{1}{2}[P_{\mathcal{H}}(u \cdot \nabla)v + P_{\mathcal{H}}(v \cdot \nabla)u]$ と定義される. $\nu > 0$ は運動粘性係数, $f \in \mathcal{H}$ は時間に依存しない外力を表す. \mathcal{H} 上の \mathcal{A} の定義域を $D(\mathcal{A})$ とする.

周期境界条件付きの非圧縮性 2次元 Navier-Stokes 方程式を, トーラス上の 2D-NSE と呼ぶ.

命題 5.10 (トーラス上の 2次元 Navier-Stokes 方程式 [41, 84, 85]). $u_0, f \in \mathcal{H}$ とすると, $u(0) = u_0$ を満たす (5.7) の一意的な弱解が存在し, 次を満たす.

$$u \in C([0, T]; \mathcal{H}) \cap L^2(0, T; \mathcal{V}), \quad \frac{du}{dt} \in L^2(0, T; \mathcal{V}^*).$$

ここで, $T > 0$ は任意とする. 半群 $\Psi_t : \mathcal{H} \ni u_0 \mapsto u_t \in \mathcal{H}$ は $t \geq 0$ に対して well-defined であり, \mathcal{H} 上で連続である. \mathcal{H} 内のボール $B(\rho) = B_{\mathcal{H}}(\rho)$ で $\rho = \frac{|f|}{\nu \lambda_1}$ は吸収集合であり, ここで $\lambda_1 > 0$ は \mathcal{A} の最小固有値を表す. 定義 5.2 に対応する $\beta \in \mathbb{R}$ が存在する. さらに, $u_0 \in \mathcal{V}$ である場合, 一意的な強解が存在し, 次を満たす.

$$u \in C([0, T]; \mathcal{V}) \cap L^2(0, T; D(\mathcal{A})), \quad \frac{du}{dt} \in L^2(0, T; \mathcal{H}).$$

ここで, $T > 0$ は任意とする. 半群 $\Psi_t : \mathcal{V} \rightarrow \mathcal{V}$ は $t \geq 0$ に対して定義される. 定義 5.1 と定義 5.2 は, \mathcal{H} の代わりに \mathcal{V} に対して, ある $\rho > 0$ と $\beta \in \mathbb{R}$ で成り立つ. さらに, $\mathcal{A} \subset \mathcal{V}$ としてグローバルアトラクターが存在する.

定義 5.10 と同様の結果は, 境界付き領域における 2D-NSE を滑りなしディリクレ境界条件 [84, 85] で考える場合にも成り立つ. 一方で, 3次元 Navier-Stokes 方程式にお

いて、時間大域的に定義された滑らかな解の存在を証明することは困難である。しかし、Camassa-Holm 方程式や Navier-Stokes-alpha 方程式などの正則化バージョンに対しては、時間大域的な解の存在と一意性が証明されている [35, 71]。同様に、Leray-alpha 方程式 [22] や Navier-Stokes-omega 方程式 [64] においても適切性の結果が得られている。

5.2 部分観測からの状態の再構成

データ同化の重要な役割の一つは、有限次元空間における部分観測から真の状態を再構築することである（節 1 参照）。ここでは、理想的な設定で部分観測の問題を議論するために、ヒルベルト空間上でのノイズのない観測 (3.12) を考察する。この問題は、制御理論に基づいて力学系の自由度を推定することと関連している。

5.2.1 Navier-Stokes 方程式に対する連続データ同化

連続データ同化 (Continuous data assimilation, CDA) として知られる問題は、不完全な観測データの時系列を用いて数値天気予報における適切な初期状態を得るために定式化された [21, 28]。類似の数値的研究 [18, 42] では、大気運動において小規模なダイナミクスが大規模なダイナミクスに従属するかどうかが調査されている。平坦なトーラス \mathbb{T}^2 上での 2D-NSE (5.7) を考える。任意の L^2 関数 $a : \mathbb{T}^2 \rightarrow \mathbb{R}^2$ は以下のように表される：

$$a = \sum_{k \in \mathcal{I}} \hat{a}_k \phi_k,$$

ここで、 $\mathcal{I} = \{2\pi m \mid m \in \mathbb{Z}^2 \setminus \{0\}\}$ はインデックス集合、 $\phi_k(x) = e^{ik \cdot x}$ は \mathcal{H} の直交基底、 $\hat{a}_k = \overline{\hat{a}_{-k}}$ である。直交射影 \mathcal{P}_λ を $\lambda > 0$ に対して以下のように定義する：

$$\mathcal{P}_\lambda a = \sum_{|k|^2 \leq \lambda} \hat{a}_k \phi_k.$$

この射影は、最小スケール長 $\lambda^{-\frac{1}{2}}$ を持つスペースな観測作用素とみなされ、 $\mathcal{Q}_\lambda = I - \mathcal{P}_\lambda$ と表記する。これは、 $h(u) = \mathcal{P}_\lambda u$ と設定することで、(3.11) と (3.12) の特別な場合とみなすことができる。このとき、(5.7) の二つの解 u_1 と u_2 を大規模成分と小規模成分に分解して考える：

$$u_i(t) = p_i(t) + q_i(t), \quad p_i(t) = \mathcal{P}_\lambda u_i(t), \quad q_i = \mathcal{Q}_\lambda u_i(t), \quad i = 1, 2. \quad (5.8)$$

数値天気予報の文脈では、 u_1 は真の状態、 u_2 はその近似を表す。観測データとして得られるのは p_1 のみである。

以下のアルゴリズムは、観測データを直接挿入する手法であり、連続データ同化 (CDA) または同期フィルター (Synchronization filter) と呼ばれる。 $u_1(0)$ の初期状態に対する初期推定値として $\varpi \in \mathcal{V}$ を与える。

定義 5.11 (トーラス上の 2D-NSE に対する連続データ同化). u_1 を推定するために、 u_1 の大規模成分 p_1 を u_2 の大規模成分としてコピーする：

$$p_2(t) = p_1(t). \quad (5.9)$$

小規模成分 q_1 を次の方程式を積分することで近似する：

$$\frac{dq_2}{dt} + \mathcal{A}q_2 + \mathcal{Q}_\lambda \mathcal{B}((p_1 + q_2), (p_1 + q_2)) = \mathcal{Q}_\lambda f, \quad q_2(0) = \mathcal{Q}_\lambda \varpi. \quad (5.10)$$

同期のために必要な長さスケールを特定するための重要な概念として、決定モード (Determining modes) がある [36, 74].

定義 5.12. 決定モードの数とは、次が成り立つ射影 P_λ の最小ランクである：方程式 (5.7) の解 u_1, u_2 が

$$\lim_{t \rightarrow \infty} |P_\lambda u_1(t) - P_\lambda u_2(t)| = 0$$

を満たすとき、次の関係

$$\lim_{t \rightarrow \infty} |u_1(t) - u_2(t)| = 0$$

も成り立つ。

次の結果は λ を下限を与える。

命題 5.13 ([48, 74]). (5.7) の 2つの解 u_1, u_2 を対応する時間非依存の外力 $f_1, f_2 \in \mathcal{H}$ と初期条件 $u_1(0), u_2(0) \in \mathcal{V}$ を持つものとする。このとき、 ν, f_i 、および任意の初期条件に依存しない定数 $c > 0$ が存在し、 $\lambda > c \text{Gr}(f_1)(2\pi/L)^2$ に対して以下が成り立つ：

$$\lim_{t \rightarrow \infty} |P_\lambda u_1(t) - P_\lambda u_2(t)| = 0$$

が成り立つとき、次の関係

$$\lim_{t \rightarrow \infty} \|u_1(t) - u_2(t)\| = 0$$

も成り立つ。ここで、 $\text{Gr}(f) = (L/2\pi\nu)^2 \limsup_{t \rightarrow \infty} |f(t)|$ を Grashof 数と呼び、 $\|\cdot\|$ は H^1 -ノルムを表す。

より詳細には、収束レートも [74] で評価されている。この解析は、観測誤差が存在するデータ同化問題においても、真の解を推定するために必要な観測の次元に関する指針を与える。また、Korn [55, 56] は、正則化された Navier-Stokes 方程式に対する連続データ同化を研究した。

注意 5.14 (各点観測). 実際の応用では、観測データを直接モデルの状態に挿入するのは困難なことが多い。例えば、測定データが離散的な空間点での真の解の値である場合、正確な空間微分を取得することはできない。Azouani ら [10, 11] は、制御理論に基づく一般的な観測作用素のための新しいアルゴリズムを提案した。この方法では、観測された真の解の情報を補間作用素を介して発展方程式のベクトル場に挿入する。

5.2.2 有限次元系に対する連続データ同化

L63 方程式や L96 方程式のような有限次元の力学系に対しても、連続データ同化 (CDA) は適切に定義できる。 $\mathcal{H} = \mathbb{R}^{N_u}$ を状態空間とし、 $\mathcal{P} \in \mathbb{R}^{N_u \times N_u}$ を射影行列とする。ここで、各行は \mathbb{R}^{N_u} の標準基底であり、 $\mathcal{Q} = I - \mathcal{P}$ と定義する。真の解を u 、その近似解を v とし、観測関数を $h(u) = \mathcal{P}u$ 、初期推定を $\varpi \in \mathbb{R}^{N_u}$ とする。この設定のもとで、定義 5.11 と同様に連続データ同化を定義できる。

定義 5.15 (常微分方程式 (ODE) の連続データ同化). 真の状態 u を推定するために、 v の大規模成分を次のように置き換える：

$$v = \mathcal{P}u + q. \quad (5.11)$$

小規模成分 q は次の微分方程式を解くことで求める：

$$\frac{dq}{dt} = \mathcal{Q}\mathcal{F}(\mathcal{P}u + q), \quad q(0) = \mathcal{Q}\varpi. \quad (5.12)$$

L96 方程式に対して、 $J = 3J'$ ($J' \in \mathbb{N}$) のとき、以下の射影行列を考える：

$$\mathcal{P} = [\phi_1, \phi_2, 0, \phi_4, \phi_5, 0, \dots] \in \mathbb{R}^{J \times J}, \quad (5.13)$$

ここで、 $(\phi_j)_{j=1}^J$ は \mathbb{R}^J の標準基底である。このとき、 \mathcal{P} のランクは $\text{rank}(\mathcal{P}) = 2J' = \frac{2}{3}J$ となる。

命題 5.16 (L96 方程式に対する CDA [62]). L96 方程式 (5.6) の解 u を考える。初期状態 $u_0 \in \mathcal{B}(\rho)$ は、定義 5.8 によって与えられる $\rho > 0$ を満たすものとする。 v は、(5.11) および (5.12) に従い、(5.13) で定義された \mathcal{P} を用いて得られるものとする。このとき、次の収束が成立する：

$$\lim_{t \rightarrow \infty} |u(t) - v(t)| = 0.$$

L63 方程式に対しても, [41] のように, 次の射影行列を用いることで, 同様の結果が得られる :

$$\mathcal{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (5.14)$$

5.2.3 離散データ同化

\mathcal{H} が無限次元の場合を再び考える. これまでの結果は, 観測データが時間的に連続であるという仮定のもとで得られた. しかし, 実際の応用では, 観測データは通常離散的な時間で得られる. \mathcal{H} 上の有限ランクの直交射影を \mathcal{P} とし, $\mathcal{Q} = I - \mathcal{P}$ と定義する. (3.11) の解 u に対して, Ψ_t を対応するセミグループとする. 増大列 $(t_n)_{n \in \mathbb{N}} \subset [0, \infty)$ は, 観測データが得られる離散時間を表す. 初期推定を $q_0 \in \mathcal{Q}(\mathcal{H})$ とする. Hayden ら [41] は, 連続データ同化の離散版として離散データ同化 (Discrete Data Assimilation, DDA) を提案した.

定義 5.17 (離散データ同化). $n = 0$ のとき, $v_0 = \mathcal{P}u_0 + q_0$ とする. 各 $n \in \mathbb{N}$ に対して, v_n を

$$v_n = \mathcal{P}u(t_n) + \mathcal{Q}\Psi_{t_n-t_{n-1}}(v_{n-1}) \quad (5.15)$$

により定義し, 区分的に連続な関数 $v(t)$ を

$$v(t) = \Psi_{t-t_{n-1}}(v_{n-1}), \quad t \in [t_{n-1}, t_n) \quad (5.16)$$

と定義する.

次の収束結果が得られる.

命題 5.18 (L63 方程式に対する離散データ同化 [41]). $u(t)$ を (5.4) の解とし, その初期値がグローバルアトラクター \mathcal{A} に属するとする. このとき, ある $t^* = t^*(\sigma, b, \varrho) > 0$ が存在し, \mathcal{P} を (5.14) のように設定し, $t_n = \tau n$ としたとき, 任意の $\tau \in (0, t^*]$ に対して近似解 $v(t)$ は $u(t)$ に収束する. すなわち,

$$\lim_{t \rightarrow \infty} |u(t) - v(t)| = 0.$$

注意 5.19 ([41]). $\sigma = 10, b = 8/3, \varrho = 28$ のとき, t^* の推定値は 0.000129 である.

L96 方程式や 2D-NSE に対しても同様の結果が得られる.

命題 5.20 (L96 方程式に対する離散データ同化 [62]). $u(t)$ を (5.6) の解とし, 初期値が定義 5.8 によって与えられる ρ を満たすとする. このとき, ある $t^* = t^*(F, J) > 0$ が存在し, \mathcal{P} を (5.13) のように設定し, $t_n = \tau n$ としたとき, 任意の $\tau \in (0, t^*]$ に対して近似解 $v(t)$ は $u(t)$ に収束する. すなわち,

$$\lim_{t \rightarrow \infty} |u(t) - v(t)| = 0.$$

命題 5.21 (2D-NSE に対する離散データ同化 [41]). $u(t)$ を (5.7) の解とし, 初期値 $u_0 \in \mathcal{A}$, $q_0 \in \mathcal{V}$ とする. このとき, ある $\lambda^* = \lambda^*(\|q_0\|, |f|, \nu, \Omega) > 0$ が存在し, 任意の $\lambda > \lambda^*$ に対して, ある $t^* = t^*(\lambda, \|q_0\|, \rho, \nu, \Omega) > 0$ が存在し, $\mathcal{P} = \mathcal{P}_\lambda$ として $t_n = \tau n$ としたとき, 任意の $\tau \in (0, t^*]$ に対して近似解 $v(t)$ は $u(t)$ に収束する. すなわち,

$$\lim_{t \rightarrow \infty} \|u(t) - v(t)\| = 0.$$

この結果の系として, $\lambda > 0$ を十分大きくとることで任意の時間間隔に対して収束が保証される.

系 5.22 ([41]). 任意の $t^* > 0$ に対して, ある $\lambda = \lambda(\rho, \|q_0\|, \nu, \Omega, t^*) > 0$ が存在し, 近似解 $v(t)$ は $u(t)$ に収束する. すなわち,

$$\lim_{t \rightarrow \infty} \|u(t) - v(t)\| = 0.$$

5.3 確率的ダイナミクス

我々の理論は現在のところ確率的なモデルダイナミクスを扱っていないが, 確率的なモデルダイナミクスを伴うデータ同化問題への理論拡張の指針として, 既存の解析結果を概観する. 確率的なモデルダイナミクス (3.1) および (3.3) に対して, エネルギー原理 (2.25) は [83, 88] に示されるように拡張でき, 解の期待値の有界性を導く.

仮定 5.23 (エネルギー原理). 離散時間確率モデルダイナミクス (3.1) に対して, ある $\lambda \in (0, 1), K > 0$ が存在して

$$|\Psi(u)|^2 + \text{Tr } Q \leq (1 - \lambda)|u|^2 + K \quad (5.17)$$

が成り立つ.

連続時間確率モデルダイナミクス (3.3) に対しては, (3.3) に対応する無限小生成作用素 \mathcal{L} を用いた (2.26) の類似形を考える:

$$\mathcal{L}\mathcal{E}(u) \leq -\lambda'\mathcal{E}(u) + K', \quad (5.18)$$

ここで, $\lambda', K' > 0$ かつ $\mathcal{E}(\cdot) = |\cdot|^2$ である. 定義 3.13 における連続時間モデルと離散時間モデルの関係を用いると, 不等式 (5.18) から 定義 5.23 が導かれる. 定義 5.23 は, シフトされた L63 方程式, L96 方程式, および 2D-NSE に対して 節 5.1 の議論と同様に確認できる. この条件により, 期待値 $\mathbb{E}[|u_t|^2]$ の有界性が保証される. しかし, 決定論的な場合とは異なり, 個々のサンプルパスの有界性は保証できない.

一方, 定義 5.3 のように最大誤差成長率を評価するためには, 確率微分方程式の解析において標準的に仮定される大域リップシツ条件が必要となる.

仮定 5.24 (大域リップシツ条件). 離散時間確率モデルダイナミクス (3.1) に対して, ある $\beta' > 0$ が存在して

$$|\Psi(u) - \Psi(v)| \leq \beta'|u - v| \quad (5.19)$$

が任意の $u, v \in \mathbb{R}^{N_u}$ について成り立つ. 連続時間確率モデルダイナミクス (3.3) に対して, ある $\beta > 0$ が存在して

$$|\mathcal{F}(u) - \mathcal{F}(v)| \leq \beta|u - v| \quad (5.20)$$

が任意の $u, v \in \mathbb{R}^{N_u}$ について成り立つ.

散逸的な確率ダイナミクスの例として, 確率摂動を受けた L63 方程式と L96 方程式が [30, 31] で考察されているが, 定義 5.24 ではなく局所リップシツ条件 (定義 5.4) を満たす. 決定論的な場合には, 軌道が有界であるため, 条件が局所的か大域的かは重要ではない. しかし確率的な場合, 確率的摂動を受けた L96 方程式のサンプルパスの有界性は [31] で数値的にのみ検証されている. 最後に, モデルノイズの非退化性は, 確率モデルダイナミクスを伴う EnKF の解析において本質的な役割を果たす [30, 88]. この点については 節 6 で議論する.

仮定 5.25 (モデルノイズの非退化性). 状態空間モデル (3.1) および (3.3) に対して, ノイズが非退化である, すなわち,

$$Q \succ 0 \quad (5.21)$$

が成り立つ.

6. EnKF の数学解析

6.1 EnKF の基本的な性質

まず、無限アンサンブル極限 $m \rightarrow \infty$ における EnKF を議論する。これは、EnKF がモンテカルロ法を用いて KF を近似するように設計されているためである。以下の結果は、線形ガウス系 (4.1) に対して、 $m \rightarrow \infty$ の極限で EnKF が KF に収束することを保証する。この性質は、KF との整合性 (consistency with the KF) と呼ばれる。

命題 6.1 (KF との整合性 [58, 70]). 線形ガウス系 (4.1) に対して、EnKF は KF に収束する。すなわち、任意の $p \in [1, \infty)$ および $n \in \mathbb{N}$ に対して、

$$\lim_{m \rightarrow \infty} \mathbb{E}[|\bar{v}_n^m - \varpi_n|^p] = 0,$$

$$\lim_{m \rightarrow \infty} \mathbb{E}[|P_n^m - P_n|_{HS}^p] = 0$$

が成り立つ。ここで、 \bar{v}_n^m および P_n^m はそれぞれ EnKF のアンサンブル平均と共分散、 ϖ_n および P_n は KF の平均と共分散を表す。

この整合性は、 \mathcal{H} および \mathcal{Y} が無限次元ヒルベルト空間である場合にも成り立つ [51, 58]。また、連続時間の枠組みにおいても、EnKBF は KBF と整合性を持つことが示されている [13, 30]。EnKF や EnKBF は、システムが線形ガウスでない限り、厳密なフィルタリング分布には収束しないことに注意が必要である。KF との整合性とは対照的に、有限のアンサンブルサイズ m における解析は、PO 法と ESRF の違いを明らかにする。特に、ESRF は、有限のアンサンブルサイズ m のもとで解析分布の平均と共分散を近似する際に、PO 法よりも誤差が小さくなることが示されている [1]。

次に、 $n \rightarrow \infty$ における EnKF の振る舞いを考察する。フィルタリングアルゴリズムが安定であるとは、同じ観測データを用いた場合に、アルゴリズムによって生成された 2 つの推定値 V_n と V'_n が同じ解に収束することを意味し、すなわち

$$\lim_{n \rightarrow \infty} |V_n - V'_n| = 0$$

が成り立つ。また、フィルタリング分布の観点から安定性を定義することもできる、すなわち

$$\lim_{n \rightarrow \infty} d(\mathbb{P}^{V_n}, \mathbb{P}^{V'_n}) = 0.$$

ここで $d(\cdot, \cdot)$ は 2 つの確率分布間の距離を表す。さらに、収束レートが指数的である場合、すなわち、ある定数 $C > 0$ および $\gamma \in (0, 1)$ が存在して

$$d(\mathbb{P}^{V_n}, \mathbb{P}^{V'_n}) \leq C\gamma^n, \quad n \in \mathbb{N}$$

が成り立つとき、この性質をフィルタリングアルゴリズムの指数安定性（または幾何エルゴード性）と呼ぶ。古典的なフィルタリング理論では、線形ガウス系 (4.1) の可制御性と可観測性が Kalman フィルタの安定性を保証する [4, 24].

非線形問題に対しては、さまざまな文脈で安定性が議論されている [9, 83]. 例えば、3DVar に対しては、トーラス上の二次元 Navier-Stokes 方程式に対する指数安定性が証明されている [17, 15]. EnKF の安定性も得られている [88]. ここでは、確率モデルダイナミクス (3.1) と線形観測 (3.2) に対する EnKF の指数安定性について紹介する.

命題 6.2 (EnKF の指数安定性 [88]). (3.1) の解 U_n と、(3.2) に基づいて EnKF によって生成されたアンサンブル $V_n^{(1)}, \dots, V_n^{(m)}$ を考える. ここで、 $\mathbf{X}_n = (U_n, V_n^{(1)}, \dots, V_n^{(m)})$ を状態空間 $\mathcal{X} = \mathbb{R}^{N_u} \times \mathbb{R}^{N_u \times m}$ 上のマルコフ連鎖とし、 P をそのマルコフ遷移カーネルとする. (3.1) が仮定 5.23 および 定義 5.25 を満たし、観測ノイズが仮定 3.1 を満たし、さらに、ある定数 $\lambda > 0, K > 0$ と、正值関数 $\mathcal{E} : \mathcal{X} \rightarrow \mathbb{R}$ が存在して、任意の $c \in \mathbb{R}$ に對してコンパクトなレベル集合 $\{\mathcal{E}(u) \leq c\}$ が存在し,

$$\mathbb{E}[\mathcal{E}(\mathbf{X}_n) | \mathcal{F}_{n-1}^{\mathbf{X}}] \leq (1 - \lambda)\mathcal{E}(\mathbf{X}_{n-1}) + K, \quad n \in \mathbb{N}$$

が成り立つとする. このとき、 $\gamma \in (0, 1)$ が存在して、任意の $\mu, \nu \in \mathcal{M}_1(\mathcal{X})$ に対し、ある定数 $C = C(\mu, \nu) > 0$ が存在して

$$d_{TV}(P^n \mu, P^n \nu) \leq C_{\mu, \nu} \gamma^n, \quad n \in \mathbb{N} \quad (6.1)$$

が成り立つ. ここで、 $d_{TV}(\cdot, \cdot)$ は全変動距離である.

定義 6.2 は、EnKF の初期誤差が時間とともに指数的に減衰することを意味する. フィルタ安定性はフィルタリングアルゴリズムにとって適切な性質であるが、EnKF による状態推定の精度を保証するものではないことに注意する必要がある.

6.2 フィルタリングアルゴリズムの誤差解析

Hilbert 空間 \mathcal{H} 上の発展方程式に関する状態空間モデルについて、EnKF の誤差解析を概観する.

$$\frac{du}{dt} = \mathcal{F}(u), \quad (6.2)$$

これは (3.11) と同じものである. 任意の $u_0 \in \mathcal{H}$ に対して一意な解が存在し、 $t \geq 0$ に対して一パラメータ半群 $\Psi_t : \mathcal{H} \rightarrow \mathcal{H}$ を生成すると仮定する. 定義 3.13 に述べたように、この連続系に対応する離散力学系を

$$u_n = \Psi(u_{n-1}), \quad n \in \mathbb{N}, \quad (6.3)$$

とする. ここで, $\Psi = \Psi_\tau$ は時間間隔 $\tau > 0$ に対応する半群作用素である. 観測空間を Hilbert 空間 $\mathcal{Y} \subset \mathcal{H}$ とする. ノイズを含む観測値 $y_n \in \mathcal{Y}$ は以下のように得られる.

$$y_n = Hu_n + \eta_n, \quad (6.4)$$

ここで, $H \in \mathcal{L}(\mathcal{H}, \mathcal{Y})$ は線形観測作用素, $(\eta_n)_{n \in \mathbb{N}} \subset \mathcal{Y}$ は独立同分布 (i.i.d.) なノイズ列である.

6.2.1 3DVar

3DVar アルゴリズムの解析から始める.

命題 6.3 ([63]). $\dim(\mathcal{H}) < \infty$ とし, u_n を (6.3) の一意な解で $u_0 \in \mathbb{R}^m$ を初期状態とする. 観測ノイズ $(\eta_n)_{n \in \mathbb{N}}$ が以下を満たすと仮定する.

$$\sup_{n \in \mathbb{N}} |\eta_n| = \epsilon. \quad (6.5)$$

さらに, 3DVar のモデル共分散 \hat{P}_0 が適切に選ばれ, 写像 $(I_{\mathcal{H}} - KH)\Psi : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_u}$ が一様リップシツ連続で, そのリップシツ定数が $\theta \in (0, 1)$ であるとする. このとき, 3DVar アルゴリズムによって生成される推定状態列 $(\varpi_n)_{n \in \mathbb{N}}$ に対して, ある定数 $c > 0$ が存在し,

$$\limsup_{n \rightarrow \infty} |\varpi_n - u_n| \leq \frac{c}{1-\theta} \epsilon \quad (6.6)$$

が成り立つ.

Proof. [63] の証明を概観し, 非線形力学系に対するデータ同化アルゴリズムの誤差解析の本質を説明する. (4.12) と (4.13) に (6.4) を適用すると,

$$\varpi_n = \widehat{\varpi}_n + K(y_n - H\widehat{\varpi}_n) = (I_{\mathcal{H}} - KH)\Psi(\varpi_{n-1}) + KHu_n + K\eta_n.$$

また, (6.3) から

$$u_n = (I_{\mathcal{H}} - KH)\Psi(u_{n-1}) + KH\Psi(u_{n-1}).$$

これらの式の差を取り, 三角不等式を適用すると, $e_n = \varpi_n - u_n$ について

$$|e_n| \leq |(I_{\mathcal{H}} - KH)\Psi(\varpi_{n-1}) - (I_{\mathcal{H}} - KH)\Psi(u_{n-1})| + |K\eta_n| \leq \theta|e_{n-1}| + c\epsilon$$

となる. ここで, θ は一様リップシツ定数, ϵ は観測ノイズの上限, $c = |K|_{\mathcal{L}}$ である. この不等式を繰り返し適用すると,

$$|e_n| = \theta^n |e_{n-1}| + c\epsilon \frac{1-\theta^n}{1-\theta} \rightarrow \frac{c\epsilon}{1-\theta} \quad (n \rightarrow \infty).$$

したがって, (6.6) が示された. □

注意 6.4. 同様に、観測ノイズ $(\eta_n)_{n \in \mathbb{N}}$ が定義 3.1 を満たす場合、 $\mathbb{E}[|e_n|^2]$ の誤差評価を得ることができる。

定義 6.3 では、非線形写像 $(I_{\mathcal{H}} - KH)\Psi$ のグローバルリップシツ定数が 1 より小さいことを仮定している。これは、以下の 2 つの条件を組み合わせることで十分条件となる。

- (1) モデルダイナミクス Ψ が定義 5.24 のようにグローバルリップシツ連続であり、定数 $\beta' > 0$ を持つ。
- (2) 解析ステップにおける誤差の収縮が $|I_{\mathcal{H}} - KH|_{\mathcal{L}} < \beta'^{-1}$ で見積もられる。

1 つ目の条件は多くの解析で一般的に仮定される。2 つ目の条件について、(4.7) より

$$|I_{\mathcal{H}} - KH|_{\mathcal{L}} = |(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1}|_{\mathcal{L}}$$

が成り立つ。全観測 (定義 3.2) を仮定し、3DVar で $\widehat{P}_n = P_0 = \alpha^2 I_{\mathcal{H}}$ を選ぶと、

$$|(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1}|_{\mathcal{L}} = |(I_{\mathcal{H}} + \alpha^2 r^{-2} I_{\mathcal{H}})^{-1}|_{\mathcal{L}} = \frac{r^2}{r^2 + \alpha^2} \quad (6.7)$$

となる。したがって、任意の $\beta' > 1$ に対し、十分に大きな $\alpha > 0$ を取ることで $|I_{\mathcal{H}} - KH|_{\mathcal{L}} < \beta'^{-1}$ を満たすことができる。一方で、部分観測を考える場合、観測行列 H はフルランクではなくなる。この場合、 $\text{Ker } H \neq \{0\}$ ならば $|I_{\mathcal{H}} - KH|_{\mathcal{L}} \geq 1$ となるため、上記の手法は利用できない。一般に、次の補題が成り立つ。

補題 6.5. \mathcal{H} と \mathcal{Y} を Hilbert 空間とする。もし $\widehat{P}_n H^* R^{-1} H \in \mathcal{L}(\mathcal{H})$ がフルランクでないならば、

$$|I_{\mathcal{H}} - KH|_{\mathcal{L}} = |(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1}|_{\mathcal{L}} \geq 1.$$

Proof. 仮定より、ある $u \in \mathcal{H}$ に対して $u \neq 0$ かつ $\widehat{P}_n H^* R^{-1} H u = 0$ となるものが存在する。したがって、

$$(I_{\mathcal{H}} + \widehat{P}_n H^* R^{-1} H)^{-1} u = u$$

が成り立つ。よって結論が得られる。 \square

3DVar の誤差評価は、節 5 で導入された部分観測を伴う散逸的力学系について得られる。証明は、節 5.2 における離散および連続データ同化の証明と類似している。さらに、3DVar の連続時間版についても同様の結果が成り立つ。結果の概要を表 1 に示す。

アルゴリズム \ モデル	L63	L96	2D-NSE
3DVar	[61]	[62]	[17]
連続時間 3DVar	[61]	[62]	[15]

表 1: 3DVar の誤差評価に関する参考文献。

6.2.2 PO 法

EnKF の誤差評価には、以下の 2 つの困難がある。

(i-1) EnKF では、リップシット定数 $L > 0$ を推定する必要があるが、

$$|u_n - \bar{\hat{v}}_n| \leq L|u_{n-1} - \bar{v}_{n-1}|$$

となる L を求めるのが困難である。特に、 $\Psi(\bar{v}_{n-1}) \neq \bar{\hat{v}}_n$ となるため、定義 5.3 を直接適用できない。

(i-2) 3DVar と異なり、予測共分散 \hat{P}_n が明示的に決まらず、たとえ完全観測の設定であっても $|I_{\mathcal{H}} + \hat{P}_n H^* R^{-1} H|_{\mathcal{L}}$ の評価が難しい。

以下の補題は、EnKF における誤差収縮の評価を与える。

補題 6.6. 定義 3.2 が成立すると仮定する。このとき、EnKF における予測共分散 \hat{P}_n は次の不等式を満たす。

$$|I_{\mathcal{H}} - K_n H|_{\mathcal{L}} = |(I_{\mathcal{H}} + r^{-2} \hat{P}_n)^{-1}|_{\mathcal{L}} \leq 1.$$

Proof. $\hat{P}_n \succeq 0$ であることから、定義 2.3 および定義 2.17 を適用することで結論が得られる。 \square

Kelly ら [54] は、各アンサンブルメンバーの状態の誤差を評価するため、 $e_n^{(k)} = v_n^{(k)} - u_n$ を定義し、各メンバーの誤差成長を定義 5.3 を用いて評価することで、上記の課題 (1) を回避する。まず、定義 6.6 を利用し、PO 法の適切性と呼ばれる性質を証明する。すなわち、有限時間での誤差の爆発が起こらないことを示す。

命題 6.7 (PO 法の適切性、[54] の修正)。モデルダイナミクス (6.2) に対して定義 5.1 および定義 5.2 が成立し、観測 (6.4) に対して定義 3.1 および定義 3.2 が成立するとする。さらに、 u_n を (6.3) の解で、 $u_0 \in \mathcal{B}(\rho)$ とし、 V_n を定義 4.10 によって生成されるアンサンブルとする。このとき、任意の $k = 1, \dots, m$ と $n \in \mathbb{N}$ に対して、

$$\mathbb{E} [|e_n^{(k)}|^2] \leq e^{2\beta\tau n} \mathbb{E} [|e_0^{(k)}|^2] + 2mr^2 \frac{e^{2\beta\tau n} - 1}{e^{2\beta\tau} - 1} \quad (6.8)$$

が成り立つ。

定義 6.7 は、 $\beta > 0$ の場合、すなわちカオス的なモデルダイナミクスに対して、時間に対して一様な誤差評価を保証しない。そのため、定義 4.19 で述べた加法的インフレー

ションを適用する。完全観測（定義 3.2）のもとで、加法的インフレーションを適用した予測共分散 \widehat{P}_n^α に対して、(6.7) と同様に

$$\begin{aligned} |(I_{\mathcal{H}} + \widehat{P}_n^\alpha H^* R^{-1} H)^{-1}|_{\mathcal{L}} &= |(I_{\mathcal{H}} + r^{-2}(\widehat{P}_n + \alpha^2 I_{\mathcal{H}}))^{-1}|_{\mathcal{L}} \\ &\leq |(I_{\mathcal{H}} + \alpha^2 r^{-2} I_{\mathcal{H}})^{-1}|_{\mathcal{L}} = \frac{r^2}{r^2 + \alpha^2} \end{aligned}$$

が成り立つ。この評価は、加法的インフレーションを適用した PO 法の誤差評価において重要な役割を果たす。

命題 6.8 (加法的インフレーションを適用した PO 法の誤差評価, [54] の修正). モデルダイナミクス (6.2) に対して定義 5.1 および定義 5.2 が成立し、観測 (6.4) に対して定義 3.1 および定義 3.2 が成立するとする。さらに、 u_n を (6.3) の解で、 $u_0 \in \mathcal{B}(\rho)$ とし、 \mathbf{V}_n を定義 4.10 に基づいて生成され、加法的インフレーション定義 4.19 を適用するとする。ここで、 $\theta = (1 + \frac{\alpha^2}{r^2})^{-2} e^{2\beta h} < 1$ と定義する。このとき、任意の $k = 1, \dots, m$ と $n \in \mathbb{N}$ に対して、

$$\mathbb{E} [|e_n^{(k)}|^2] \leq \theta^n \mathbb{E} [|e_0^{(k)}|^2] + 2mr^2 \frac{1 - \theta^n}{1 - \theta} \quad (6.9)$$

が成り立つ。特に、 $\theta < 1$ の場合、

$$\lim_{n \rightarrow \infty} \mathbb{E} [|e_n^{(k)}|^2] \leq \frac{2mr^2}{1 - \theta}$$

が成立する。

注意 6.9. 定義 6.7 および定義 6.8 は、定義 3.1 の代わりに定義 3.11 を仮定しても成立する。これは、次の補題による。

補題 6.10. ヒルベルト空間 \mathcal{H} に対して、 $\eta \sim \mathcal{N}(0, \tilde{R})$ かつ $\tilde{R} \succeq 0$ で $\text{Tr } \tilde{R} < \infty$, $\tilde{R} \preceq r^2 I_{\mathcal{H}}$ とする。さらに、 $\mathcal{P} : \mathcal{H} \rightarrow \mathcal{H}$ を直交射影とし、 $\dim(\mathcal{P}(\mathcal{H})) < \infty$ とする。このとき、

$$\mathbb{E}[|\mathcal{P}\eta|^2] \leq \dim(\mathcal{P}(\mathcal{H}))r^2.$$

次に、連続時間版の EnKF である EnKBF（定義 4.25 で定義）について考える。ここでは、有限次元の状態空間 $\mathcal{H} = \mathbb{R}^{N_u}$ および観測空間 $\mathcal{Y} = \mathbb{R}^{N_y}$ を仮定する。完全観測および小さいノイズの下で、EnKBF の一様時間誤差評価が得られる。

命題 6.11 (EnKBF の精度評価 [30]). 確率的モデルダイナミクス (3.3) に対して、定義 5.24 および定義 5.25 が成立し、観測 (3.4) に対して定義 3.2 が成立すると仮定する。また、観測ノイズの分散 r^2 は十分小さく選ばれるとする。初期アンサンブル $\mathbf{V}_0 \in \mathbb{R}^{N_u \times m}$ は、初期アンサンブル共分散 $P_0 \in \mathbb{R}^{N_u \times N_u}$ が可逆であり、

$$\lambda_{\max}(P_0) \leq C_1 r, \quad \lambda_{\min}(P_0) \geq C_2 r$$

を満たすように選ばれる（ここで， $C_1, C_2 > 0$ は定数）．確率変数 U_t を (3.3) の解， $\mathbf{V}_t \in \mathbb{R}^{N_u \times m}$ を定義 4.25 によって生成されるアンサンブルとする．このとき，

$$\lim_{t \rightarrow \infty} \mathbb{E}[|U_t - \bar{v}_t|^2] = O(r) \quad (6.10)$$

が成り立つ．

インフレーションを用いずに，この結果はモデルノイズの非退化性（定義 5.25）によって保証される．

7. ESRF の誤差解析

7.1 ESRF の適切性

ここでも、(6.2)-(6.4) で与えられる状態空間モデルを考える。節 6.2 においてこのモデルを用いる。状態推定誤差のアンサンブルを $\mathbf{E}_n = [e_n^{(k)}]_{k=1}^m \in \mathcal{H}^m$ と定義し、 \mathcal{F}_n を初期不確実性 V_0 と観測ノイズ系列 $(\eta_k)_{k=1}^n$ によって生成される σ -代数とする。ここでは ETKF の適切性を示す。

定理 7.1 (ETKF の適切性 [82]). ダイナミクス (6.2) が 定義 5.1 および 定義 5.4 を満たし、観測 (6.4) が 定義 3.2 および 定義 3.11 を満たすとする。 (6.3) の解 u_n は初期条件 $u_0 \in B(\rho)$ を満たし、ETKF (定義 4.13) によってアンサンブル V_n が生成されるとする。このとき、以下の評価が成り立つ：

$$\mathbb{E} [|\mathbf{E}_n|_2^2] \leq e^{2\beta hn} \mathbb{E} [|\mathbf{E}_0|_2^2] + (m-1)r^2 \frac{e^{2\beta hn} - 1}{e^{2\beta h} - 1}, \quad n \in \mathbb{N}. \quad (7.1)$$

ETKF の解析ステップ (定義 4.13) に必要な補題を示す。これは、PO 法に関する [54] の Proposition 3.2 に対応している。

補題 7.1. 変換行列について、次の関係が成り立つ。

$$dV_n \mathbf{1}^* = d\hat{V}_n \mathbf{1}^* = 0 \in \mathcal{H}, \quad (7.2)$$

$$T_n \mathbf{1}^* = \mathbf{1}^*. \quad (7.3)$$

さらに、アンサンブルについて、次の関係式が成り立つ。

$$(I_{\mathcal{H}} + \hat{P}_n H^* R^{-1} H) V_n = \hat{V}_n T_n^{-1} + \hat{P}_n H^* R^{-1} y_n \mathbf{1}. \quad (7.4)$$

Proof. 以下の証明では、簡単のために時間インデックス n を省略する。定義 2.14 の式 (2.5) より、

$$S^{-1} \mathbf{1}^* = \left(I_m + \frac{1}{m-1} dV^* H^* R^{-1} H dV \right) \mathbf{1}^* = \mathbf{1}^*$$

となる。ここで、 $S = T^2$ は (4.24) で定義される。したがって、

$$S \mathbf{1}^* = \mathbf{1}^* \quad (7.5)$$

が成り立つ。

次に, $\mathbf{1}^*$ が $T = S^{\frac{1}{2}}$ の固有ベクトルであり, 固有値が 1 であることを示す. S は対称行列であるため, ユニタリ行列 $U \in \mathbb{R}^{m \times m}$ と対角行列 $D \in \mathbb{R}^{m \times m}$ を用いて $S = UDU^*$ と対角化できる. このとき, (7.5) は次のように書き換えられる.

$$S\mathbf{1}^* = \mathbf{1}^* \Leftrightarrow UDU^*\mathbf{1}^* = \mathbf{1}^* \Leftrightarrow DU^*\mathbf{1}^* = U^*\mathbf{1}^*.$$

ここで, $u = U^*\mathbf{1}^* = (u^1, \dots, u^m)^* \in \mathbb{R}^m$ とし, D の対角要素を $d^j > 0$ ($j = 1, \dots, m$) とおくと, 各成分について次の関係が成り立つ.

$$d^j u^j = u^j, \quad j = 1, \dots, m.$$

これは, $d^j = 1$ または $u^j = 0$ を意味する. したがって,

$$(d^j)^{\frac{1}{2}} u^j = u^j, \quad j = 1, \dots, m.$$

となり, $D^{\frac{1}{2}}U^*\mathbf{1}^* = U^*\mathbf{1}^*$ が成り立つ. 定義より $T = UD^{\frac{1}{2}}U^*$ であるから, $T\mathbf{1}^* = \mathbf{1}^*$ が示された. これは (7.3) である.

最後に, (7.4) を示す. (4.23) より,

$$(I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)\bar{v} = \bar{v} + \widehat{P}H^*R^{-1}y \in \mathcal{H}.$$

が成り立つ. また, $\widehat{P} = \text{Cov}_m(\widehat{\mathbf{V}})$, (4.24), および $S = T^2$ を用いると,

$$\begin{aligned} (I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)d\mathbf{V} &= (I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)d\widehat{\mathbf{V}}T \\ &= d\widehat{\mathbf{V}} \left[I_m + \frac{1}{m-1}d\widehat{\mathbf{V}}^*H^*R^{-1}Hd\widehat{\mathbf{V}} \right] T \\ &= d\widehat{\mathbf{V}}S^{-1}T = d\widehat{\mathbf{V}}T^{-1} \in \mathcal{H}^m. \end{aligned}$$

最後に, (4.23) と $\bar{v}\mathbf{1}T^{-1} = \bar{v}\mathbf{1}$ を用いると,

$$\begin{aligned} (I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)V &= (I_{\mathcal{H}} + \widehat{P}H^*R^{-1}H)(\bar{v}\mathbf{1} + d\mathbf{V}) \\ &= \bar{v}\mathbf{1} + \widehat{P}H^*R^{-1}y\mathbf{1} + d\widehat{\mathbf{V}}T^{-1} \\ &= \bar{v}\mathbf{1}T^{-1} + d\widehat{\mathbf{V}}T^{-1} + \widehat{P}H^*R^{-1}y\mathbf{1} \\ &= \widehat{\mathbf{V}}T^{-1} + \widehat{P}H^*R^{-1}y\mathbf{1}. \end{aligned}$$

以上で証明が完了する. \square

(Proof of 定理 7.1). 定義 3.2 より, 関係式 (7.4) は以下と同値である.

$$(I_{\mathcal{H}} + r^{-2}\widehat{P}_n)\mathbf{V}_n = \widehat{\mathbf{V}}_nT_n^{-1} + r^{-2}\widehat{P}_ny_n\mathbf{1}. \quad (7.6)$$

ここで, $\mathbf{U}_n = u_n \mathbf{1} \in \mathcal{H}^m$ と定める. (7.3) より, $\mathbf{U}_n = \mathbf{U}_n T_n^{-1}$ が成り立つ. したがって,

$$(I_{\mathcal{H}} + r^{-2} \widehat{P}_n) \mathbf{U}_n = \mathbf{U}_n + r^{-2} \widehat{P}_n \mathbf{U}_n = \mathbf{U}_n T_n^{-1} + r^{-2} \widehat{P}_n \mathbf{U}_n. \quad (7.7)$$

ここで, $\widehat{\mathbf{E}}_n = \widehat{\mathbf{V}}_n - \mathbf{U}_n$ とおくと, (7.7) を (7.6) から引くことで,

$$(I_{\mathcal{H}} + r^{-2} \widehat{P}_n) \mathbf{E}_n = \widehat{\mathbf{E}}_n T_n^{-1} + r^{-2} \widehat{P}_n (y_n - u_n) \mathbf{1} = \widehat{\mathbf{E}}_n T_n^{-1} + r^{-2} \widehat{P}_n \eta_n \mathbf{1}$$

を得る.

ここで, $r^{-2} \widehat{P}_n \succeq 0$ であるため, $I_{\mathcal{H}} + r^{-2} \widehat{P}_n$ は可逆である. したがって, 両辺に $(I_{\mathcal{H}} + r^{-2} \widehat{P}_n)^{-1}$ を掛けると,

$$\mathbf{E}_n = (I_{\mathcal{H}} + r^{-2} \widehat{P}_n)^{-1} \widehat{\mathbf{E}}_n T_n^{-1} + (I_{\mathcal{H}} + r^{-2} \widehat{P}_n)^{-1} r^{-2} \widehat{P}_n \eta_n \mathbf{1}$$

を得る. ここで, \mathbf{E}_n を以下の 2 つの項に分けて評価する.

$$R_1 = (I_{\mathcal{H}} + r^{-2} \widehat{P}_n)^{-1} \widehat{\mathbf{E}}_n T_n^{-1}, \quad (7.8)$$

$$R_2 = (I_{\mathcal{H}} + r^{-2} \widehat{P}_n)^{-1} r^{-2} \widehat{P}_n \eta_n \mathbf{1}. \quad (7.9)$$

ここで, P_n の像空間の次元は高々 $m - 1$ である. これは \widehat{P}_n が m 個のベクトルから構成されるが, 1 つの制約があるためである. よって, \mathcal{H} から $\text{Ran}(P_n)$ への射影を Π_n とすると, $R_2 = (I_{\mathcal{H}} + r^{-2} \widehat{P}_n)^{-1} r^{-2} \widehat{P}_n \Pi_n \eta_n \mathbf{1}$ となる. (2.9) より, $r^{-2} \widehat{P}_n \succeq 0$ であるため, $(I_{\mathcal{H}} + r^{-2} \widehat{P}_n)^{-1} r^{-2} \widehat{P}_n \preceq I$ が成り立つ. したがって,

$$|R_2|_2^2 \leq |\Pi_n \eta_n \mathbf{1}|_2^2 = |\Pi_n \eta_n|^2 \quad (7.10)$$

を得る.

ここで, $J = I_{\mathcal{H}} + r^{-2} \widehat{P}_n$ とおく. このとき, $J, J^{-1} \in \mathcal{L}_{sa}(\mathcal{H})$ であり, $|J^{-1}|_{\mathcal{L}} \leq 1$ である. よって,

$$R_1 R_1^* = J^{-1} \widehat{\mathbf{E}}_n T_n^{-2} \widehat{\mathbf{E}}_n^* J^{-1}$$

となる.

ここで, (7.3) および $\widehat{\mathbf{E}}_n d\widehat{\mathbf{V}}_n^* = d\widehat{\mathbf{V}}_n \widehat{\mathbf{E}}_n^* = d\widehat{\mathbf{V}}_n d\widehat{\mathbf{V}}_n^*$ を用いると,

$$\begin{aligned} \widehat{\mathbf{E}}_n T_n^{-2} \widehat{\mathbf{E}}_n^* &= \widehat{\mathbf{E}}_n \left[I_m + \frac{r^{-2}}{m-1} d\widehat{\mathbf{V}}_n^* d\widehat{\mathbf{V}}_n \right] \widehat{\mathbf{E}}_n^* \\ &= \widehat{e}_n \mathbf{1} (\widehat{e}_n \mathbf{1})^* + d\widehat{\mathbf{V}}_n d\widehat{\mathbf{V}}_n^* + d\widehat{\mathbf{V}}_n d\widehat{\mathbf{V}}_n^* r^{-2} \widehat{P}_n \\ &= \widehat{e}_n \mathbf{1} (\widehat{e}_n \mathbf{1})^* + d\widehat{\mathbf{V}}_n d\widehat{\mathbf{V}}_n^* J \end{aligned}$$

となる. この関係を利用して R_1 の評価を進め, 定義 5.3 を用いることで, 最終的に

$$\mathbb{E}[|\mathbf{E}_n|_2^2] \leq e^{2\beta h} \mathbb{E}[|\mathbf{E}_{n-1}|_2^2] + (m-1)r^2$$

を得る. この不等式を繰り返し適用することで, (7.1) を得る. \square

EAKF に対しても同様の結果が得られる.

定理 7.2 (EAKF の適切性). 定理 7.1 と同じ条件のもとで, EAKF (定義 4.16) によって生成される \mathbf{V}_n について, (7.1) と同じ誤差評価が成り立つ.

Proof. ETKF の場合と同様に, 時間インデックス n は固定し, 証明の簡潔さのため省略する. ETKF との違いは R_1 の評価にある.

明らかに, 次の関係式が成り立つ:

$$u = (I_{\mathcal{H}} - K)u + Ku. \quad (7.11)$$

一方, EAKF における解析アンサンブルは

$$\mathbf{V} = \bar{v}\mathbf{1} + d\mathbf{V} = \left[(I_{\mathcal{H}} - K)\bar{v} + Ky \right] \mathbf{1} + Ad\hat{\mathbf{V}} \quad (7.12)$$

で与えられる. (7.11) を (7.12) から引くことで,

$$\mathbf{E} = (I_{\mathcal{H}} - K)\hat{e}\mathbf{1} + Ad\hat{\mathbf{V}} + K(y - u)\mathbf{1} \quad (7.13)$$

を得る. 定義 4.4 より, $(I_{\mathcal{H}} - K)^{-1} = I_{\mathcal{H}} + r^{-2}\hat{P} = J$ である. さらに, 定義 4.14 から

$$K = (I_{\mathcal{H}} - K)r^{-2}\hat{P} = (I_{\mathcal{H}} + r^{-2}\hat{P})^{-1}r^{-2}\hat{P}$$

が成り立つ. したがって, (7.13) は

$$\mathbf{E} = J^{-1}\hat{e}\mathbf{1} + Ad\hat{\mathbf{V}} + (I_{\mathcal{H}} + r^{-2}\hat{P})^{-1}r^{-2}\hat{P}\eta\mathbf{1}$$

と書ける. ここで, 誤差のアンサンブル \mathbf{E} を次の 2 つの項に分解する:

$$\begin{aligned} R_1 &= J^{-1}\hat{e}\mathbf{1} + Ad\hat{\mathbf{V}}, \\ R_2 &= (I_{\mathcal{H}} + r^{-2}\hat{P})^{-1}r^{-2}\hat{P}\eta\mathbf{1}. \end{aligned}$$

ここで, R_2 は ETKF の場合と同じである. また, $J = (I_{\mathcal{H}} - K)^{-1}$ であるため, (4.28) における調整作用素の定義から,

$$Ad\hat{\mathbf{V}}(Ad\hat{\mathbf{V}})^* = (m-1)(I_{\mathcal{H}} - K)\hat{P} = J^{-1}d\hat{\mathbf{V}}d\hat{\mathbf{V}}^*$$

が成り立つ. また, $d\hat{\mathbf{V}}\mathbf{1}^* = 0$ であることを考慮すると,

$$R_1 R_1^* = J^{-1}\hat{e}\mathbf{1}(J^{-1}\hat{e}\mathbf{1})^* + Ad\hat{\mathbf{V}}d\hat{\mathbf{V}}^*A^* = J^{-1}\hat{e}\mathbf{1}(J^{-1}\hat{e}\mathbf{1})^* + J^{-1}d\hat{\mathbf{V}}d\hat{\mathbf{V}}^*$$

が得られる. これは ETKF の証明における $R_1 R_1^*$ と同じである. \square

注意 7.2. 定理 7.1 と定理 7.2 は, ESRF の短時間における精度を保証する. さらに, これらの結果は \mathcal{H} が無限次元であっても成り立ち, PO 法に関する定義 6.7 と類似している. 定義 6.7 と比較すると, 定理 7.1 および定理 7.2 の評価 (7.1) の第二項には係数 2 が現れない. これは, PO 法が観測ノイズの分散の影響をより受けやすいことを示している.

7.2 ESRF の時間大域的な誤差評価

解析誤差 $e_n = \bar{v}_n - u_n$ の時間大域的な誤差評価を行う。困難 (i-1) から、アンサンブル平均の誤差増大を評価するために、定義 5.3 の代わりに定義 5.5 を用いる必要がある。

定理 7.3 (乗法的インフレーションを用いた ETKF の誤差評価 [82]). $\dim(\mathcal{H}) = N_u < \infty$ とする。 (6.2) のモデルダイナミクスが定義 5.1 および定義 5.4 を満たし、 (6.4) の観測が定義 3.11 および定義 3.2 を満たしているとする。 (6.3) の解 u_n は初期状態 $u_0 \in B(\rho)$ で与えられ、 V_n は定義 4.13 の ETKF に定義 4.20 の乗法的インフレーションを $\alpha \geq 1$ として適用したものとする。さらに、アンサンブルサイズ $m \in \mathbb{N}$ が十分に大きく、 $\lambda_{\min}(P_0) \geq \lambda_0$ を満たすとし、すべての $k = 1, \dots, m$ および $n \in \mathbb{N}$ に対して $v_n^{(k)} \in B(\rho)$ であると仮定する。このとき、任意の $\epsilon > 0$ に対して、 $\alpha_0 = \alpha_0(\rho, \beta, m, \lambda_0, \tau, r, \epsilon) \geq 1$ が存在し、すべての $\alpha \geq \alpha_0$ について以下が成り立つ。

(i) $\lambda_* = \lambda_*(\rho, \beta, m, \lambda_0, \tau, r, \alpha) > 0$ が存在し、すべての $n \in \mathbb{N}$ に対して $\lambda_{\min}(\hat{P}_n) > \lambda_*$ が成り立つ。

(ii) $n \in \mathbb{N}$ および $\theta = (1 + \frac{\alpha^2}{r^2} \lambda_*)^{-2} e^{2(\beta+\epsilon)\tau}$ に対して、

$$\mathbb{E}[|e_n|^2] \leq \theta^n (\mathbb{E}[|e_0|^2] + D) + N_u r^2 \frac{1 - \theta^n}{1 - \theta} + \left(\frac{(1 - \theta^n)(1 - \Theta)}{1 - \theta} - 1 \right) D \quad (7.14)$$

が成り立つ。ここで、 $D = \frac{\beta^2 \rho^2}{(\beta+\epsilon)\epsilon}$ および $\Theta = (1 + \frac{\alpha_0^2}{r^2} \lambda_*)^{-2}$ である。さらに、 $\theta < 1$ ならば、

$$\lim_{n \rightarrow \infty} \mathbb{E}[|e_n|^2] \leq \frac{N_u r^2}{1 - \theta} + \left(\frac{1 - \Theta}{1 - \theta} - 1 \right) D \quad (7.15)$$

が成り立つ。

Proof. 簡単のため、 $\hat{\lambda}_n^{\min} = \lambda_{\min}(\hat{P}_n)$ および $\lambda_n^{\min} = \lambda_{\min}(P_n)$ と書くことにする。まず、予測ステップにおける固有値の変化、すなわち λ_{n-1}^{\min} から $\hat{\lambda}_n^{\min}$ への変化を見積もる。このために、予測ステップを次のように補間する：

$$\hat{v}_t^{(k)} = \Psi_t(v_{n-1}^{(k)}), \quad \hat{P}_t = \frac{1}{m-1} \sum_{k=1}^m (\hat{v}_t^{(k)} - \bar{\hat{v}}_t) \otimes (\hat{v}_t^{(k)} - \bar{\hat{v}}_t), \quad \lambda_t = \lambda_{\min}(\hat{P}_t), \quad t \in [0, \tau].$$

定義 5.1 より、 $\hat{v}_t^{(k)} \in \mathcal{B}(\rho)$ である。 \hat{P}_t を t で微分すると、

$$\frac{d}{dt} \hat{P}_t = \frac{1}{m-1} \sum_{k=1}^m (F(\hat{v}_t^{(k)}) - \bar{F}_t) \otimes (\hat{v}_t^{(k)} - \bar{\hat{v}}_t) + (\hat{v}_t^{(k)} - \bar{\hat{v}}_t) \otimes (F(\hat{v}_t^{(k)}) - \bar{F}_t), \quad (7.16)$$

ここで,

$$\bar{F}_t = \frac{1}{m} \sum_{k=1}^m F(\hat{v}_t^{(k)}).$$

また, $\frac{1}{m} \sum_{k=1}^m \hat{v}_t^{(k)} - \bar{v}_t = 0$ であるため,

$$\begin{aligned} \frac{d}{dt} \hat{P}_t &= \frac{1}{m-1} \sum_{k=1}^m (F(\hat{v}_t^{(k)}) - F(\bar{v}_t)) \otimes (\hat{v}_t^{(k)} - \bar{v}_t) + (\hat{v}_t^{(k)} - \bar{v}_t) \otimes (F(\hat{v}_t^{(k)}) - F(\bar{v}_t)). \end{aligned} \quad (7.17)$$

定義 2.13 によると, $t \in (0, \tau)$ に対して, ある $w = w_t \in \mathcal{H}$ が存在し, $|w| = 1$ であって,

$$\frac{d}{dt} \lambda_t = \left\langle w, \frac{d}{dt} \hat{P}_t w \right\rangle.$$

を満たす. $\frac{d}{dt} \lambda_t$ の下限を導出するため, (7.17) の右辺の絶対値を考える. $|w| = 1$ であり, 定義 5.4 により,

$$\begin{aligned} \left| \left\langle w, \frac{d}{dt} \hat{P}_t w \right\rangle \right| &\leq \left| \frac{2}{m-1} \sum_{k=1}^m \left\langle F(\hat{v}_t^{(k)}) - F(\bar{v}_t), w \right\rangle \left\langle \hat{v}_t^{(k)} - \bar{v}_t, w \right\rangle \right| \\ &\leq 2 \left(\frac{1}{m-1} \sum_{k=1}^m \left\langle F(\hat{v}_t^{(k)}) - F(\bar{v}_t), w \right\rangle^2 \right)^{\frac{1}{2}} \left(\frac{1}{m-1} \sum_{k=1}^m \left\langle \hat{v}_t^{(k)} - \bar{v}_t, w \right\rangle^2 \right)^{\frac{1}{2}} \\ &\leq \beta \frac{1}{m-1} \sum_{k=1}^m |\hat{v}_t^{(k)} - \bar{v}_t|^2 \leq 8 \frac{m}{m-1} \beta \rho^2. \end{aligned}$$

最後の不等式は $\hat{v}_t^{(k)} \in B(\rho)$ であることと, 定理 7.3 の仮定による. したがって,

$$\frac{d}{dt} \lambda_t \geq -a$$

が成り立ち, ここで $a = 8 \frac{m}{m-1} \beta \rho^2 > 0$ である. $t = 0$ から $t = \tau$ まで積分すると,

$$\hat{\lambda}_n^{\min} = \lambda_\tau \geq e^{-a\tau} \lambda_0 = e^{-a\tau} \lambda_{n-1}^{\min}. \quad (7.18)$$

解析ステップにおける固有値の変化を考える. 定義 3.2 および (4.36) より,

$$P_{n-1} = \frac{\alpha^2}{m-1} d\hat{\mathbf{V}}_{n-1} (I_m + \alpha^2 \gamma^{-2} \tilde{P}_{n-1})^{-1} d\hat{\mathbf{V}}_{n-1}^*$$

が成り立つ. ここで,

$$\tilde{P}_{n-1} = \frac{1}{m-1} d\hat{\mathbf{V}}_{n-1}^* d\hat{\mathbf{V}}_{n-1} \in \mathbb{R}^{m \times m}$$

である. 次に, $n \in \mathbb{N}$ を固定し, \widehat{P}_{n-1} の固有ベクトルが P_{n-1} の固有ベクトルでもあることを示す. 実際, $\phi \in \mathcal{H}$ が固有値 $\lambda \geq 0$ に対して $\widehat{P}_{n-1}\phi = \lambda\phi$ を満たすとすると,

$$\widetilde{P}_{n-1}d\widehat{\mathbf{V}}_{n-1}^*\phi = \frac{1}{m-1}d\widehat{\mathbf{V}}_{n-1}^*d\widehat{\mathbf{V}}_{n-1}d\widehat{\mathbf{V}}_{n-1}^*\phi = d\widehat{\mathbf{V}}_{n-1}^*\widehat{P}_{n-1}\phi = \lambda d\widehat{\mathbf{V}}_{n-1}^*\phi$$

となる. したがって,

$$\begin{aligned} P_{n-1}\phi &= \frac{\alpha^2}{m-1}d\widehat{\mathbf{V}}_{n-1}(I_m + \alpha^2\gamma^{-2}\widetilde{P}_{n-1})^{-1}d\widehat{\mathbf{V}}_{n-1}^*\phi \\ &= \frac{\alpha^2}{m-1}d\widehat{\mathbf{V}}_{n-1}\frac{1}{1+\alpha^2\gamma^{-2}\lambda}d\widehat{\mathbf{V}}_{n-1}^*\phi \\ &= \frac{\alpha^2}{1+\alpha^2\gamma^{-2}\lambda}\widehat{P}_{n-1}\phi = \frac{\alpha^2\lambda}{1+\alpha^2\gamma^{-2}\lambda}\phi \end{aligned}$$

が成り立つ. 関数 $\lambda \mapsto \frac{\alpha^2\lambda}{1+\alpha^2\gamma^{-2}\lambda}$ は単調増加するため, 最小固有値の間には

$$\lambda_{n-1}^{min} = \frac{\alpha^2\widehat{\lambda}_{n-1}^{min}}{1 + \frac{\alpha^2}{\gamma^2}\widehat{\lambda}_{n-1}^{min}} \quad (7.19)$$

という関係が成り立つ. (7.18) と (7.19) を組み合わせることで, $\widehat{\lambda}_n^{min}$ の不等式を得る:

$$\widehat{\lambda}_n^{min} \geq \frac{e^{-a\tau}\alpha^2\widehat{\lambda}_{n-1}^{min}}{1 + \frac{\alpha^2}{\gamma^2}\widehat{\lambda}_{n-1}^{min}}.$$

ここで, 固有値に関する離散力学系

$$\lambda_{n+1} = g(\lambda_n), \quad \lambda_0 > 0$$

を考える. ただし, $g(\lambda) = \frac{e^{-a\tau}\alpha^2\lambda}{1 + \frac{\alpha^2}{\gamma^2}\lambda}$ である. このとき, すべての $n \in \mathbb{N} \cup \{0\}$ に対して $\lambda_n > 0$ が成り立つ. この力学系の不動点を $\lambda_\infty = \frac{\gamma^2}{\alpha^2}(e^{-a\tau}\alpha^2 - 1)$ とする. すなわち, $\lambda_\infty = g(\lambda_\infty)$ を満たす. $e^{-a\tau}\alpha^2 > 1$ の場合, $\frac{g(\lambda)}{\lambda}$ は $\lambda \leq \lambda_\infty$ のとき ≥ 1 を満たし, $\lambda > \lambda_\infty$ のとき < 1 を満たす. したがって, $\lim_{n \rightarrow \infty} \lambda_n = \lambda_\infty$ となる. 一方, $e^{-a\tau}\alpha^2 \leq 1$ の場合, $\lim_{n \rightarrow \infty} \lambda_n = 0$ である. したがって, 次の下限評価を得る:

$$\widehat{\lambda}_n^{min} \geq \min \left\{ \widehat{\lambda}_0^{min}, \frac{\gamma^2}{\alpha^2}(e^{-a\tau}\alpha^2 - 1) \right\} = \min \left\{ e^{-a\tau}\lambda_0, \frac{\gamma^2}{\alpha^2}(e^{-a\tau}\alpha^2 - 1) \right\} = \lambda_* > 0 \quad (7.20)$$

ただし, この不等式が成り立つののは $e^{-a\tau}\alpha^2 > 1$ の場合に限る.

最後に, $\mathbb{E}[|e_n|^2]$ に対する漸化不等式を導く. 定義 3.2 により, 等式 (4.35) は次の形に書き換えられる:

$$(I + \alpha^2\gamma^{-2}\widehat{P}_n)\bar{v}_n = \bar{v}_n + \alpha^2\gamma^{-2}\widehat{P}_ny_n.$$

定理 7.1 の証明と同様に、誤差を次の 2 つの項に分解する：

$$r_1 = (I_{\mathcal{H}} + \alpha^2 \gamma^{-2} \widehat{P}_n)^{-1} \widehat{e}_n, \quad (7.21)$$

$$r_2 = (I_{\mathcal{H}} + \alpha^2 \gamma^{-2} \widehat{P}_n)^{-1} \alpha^2 \gamma^{-2} \widehat{P}_n (y_n - u_n), \quad (7.22)$$

ここで、 $\widehat{e}_n = \bar{\widehat{v}}_n - u_n$ である。[\(2.9\)](#) より、

$$|(I_{\mathcal{H}} + \alpha^2 \gamma^{-2} \widehat{P}_n)^{-1} \alpha^2 \gamma^{-2} \widehat{P}_n|_{\mathcal{L}} \leq 1$$

が成り立つ。したがって、

$$|r_2| \leq |y_n - u_n| = |\eta_n| \quad (7.23)$$

を得る。

また、最小固有値の下限 [\(7.20\)](#) より、

$$|(I_{\mathcal{H}} + \alpha^2 \gamma^{-2} \widehat{P}_n)^{-1}|_{\mathcal{L}} \leq (1 + \frac{\alpha^2}{\gamma^2} \lambda_*)^{-1}$$

が成り立つ。したがって、

$$|r_1|^2 \leq \frac{1}{(1 + \frac{\alpha^2}{\gamma^2} \lambda_*)^2} |\widehat{e}_n|^2 \leq \frac{e^{2(\beta+\epsilon)h}}{(1 + \frac{\alpha^2}{\gamma^2} \lambda_*)^2} (|e_{n-1}|^2 + D) - \frac{D}{(1 + \frac{\alpha^2}{\gamma^2} \lambda_*)^2}, \quad (7.24)$$

ここで、 $\epsilon > 0$ に対して 定義 5.5 を用いた。

定理 7.1 の証明と同様に、 r_1 と r_2 の期待値をそれぞれ計算すると、

$$\begin{aligned} \mathbb{E}[|e_n|^2] &= \mathbb{E}[|r_1|^2] + \mathbb{E}[|r_2|^2] \leq \mathbb{E}[|r_1|^2] + \mathbb{E}[|\eta_n|^2] = \mathbb{E}[|r_1|^2] + N_u \gamma^2 \\ &\leq \frac{e^{2(\beta+\epsilon)h}}{(1 + \frac{\alpha^2}{\gamma^2} \lambda_*)^2} (\mathbb{E}[|e_{n-1}|^2] + D) - \frac{D}{(1 + \frac{\alpha^2}{\gamma^2} \lambda_*)^2} + N_u \gamma^2 \\ &\leq \theta (\mathbb{E}[|e_{n-1}|^2] + D) - \frac{D}{(1 + \frac{\alpha^2}{\gamma^2} \lambda_*)^2} + N_u \gamma^2 = \theta (\mathbb{E}[|e_{n-1}|^2] + D) + \delta, \end{aligned}$$

ここで、 $\delta = N_u \gamma^2 - \Theta D$ である。この不等式を利用すると、

$$\mathbb{E}[|e_n|^2] + D \leq \theta (\mathbb{E}[|e_{n-1}|^2] + D) + \delta + D$$

を得る。この不等式を繰り返し適用すると、

$$\begin{aligned} \mathbb{E}[|e_n|^2] &\leq \theta^n (\mathbb{E}[|e_0|^2] + D) + (\delta + D) \frac{1 - \theta^n}{1 - \theta} - D \\ &= \theta^n (\mathbb{E}[|e_0|^2] + D) + (N_u \gamma^2 + (1 - \Theta)D) \frac{1 - \theta^n}{1 - \theta} - D \\ &= \theta^n (\mathbb{E}[|e_0|^2] + D) + N_u \gamma^2 \frac{1 - \theta^n}{1 - \theta} + \left(\frac{(1 - \theta^n)(1 - \Theta)}{1 - \theta} - 1 \right) D \end{aligned}$$

を得る. さらに, $\theta < 1$ の場合, $n \rightarrow \infty$ の極限において

$$\mathbb{E}[|e_n|^2] \leq \frac{N_u \gamma^2}{1 - \theta} + \left(\frac{1 - \Theta}{1 - \theta} - 1 \right) D$$

が成り立つ.

□

同様に, EAKF の誤差評価も得られる.

定理 7.4 (乗法的インフレーションを用いた EAKF の誤差評価). 乗法的インフレーション 定義 4.20 を用いた EAKF によって生成される \mathbf{V}_n に対して, 定理 7.3 と同じ仮定を置く. ここで, インフレーション係数は $\alpha \geq 1$ とする. このとき, 誤差評価 (7.14) および (7.15) が同様に成り立つ.

Proof. 証明は 定理 7.3 の場合と同じである. なぜなら, 定義 4.22 の関係式

$$(I_{\mathcal{H}} + \alpha^2 \widehat{P}_n H^* R^{-1} H) \bar{v}_n = \bar{\bar{v}}_n + \alpha^2 \widehat{P}_n H^* R^{-1} y_n,$$

$$P_n = \frac{\alpha^2}{m-1} d\widehat{V}_n (I_m + \alpha^2 d\widehat{V}_n^* H^* R^{-1} H d\widehat{V}_n)^{-1} d\widehat{V}_n^*$$

を用いるだけだからである. これらの関係式は ETKF と EAKF の両方で共通している.

□

Luo と Hoteit [68] は, 乗法的インフレーション係数 α の上限と下限を導出し, 残差誤差 $e_n^r = y_n - H\bar{v}_n$ がある範囲内に収まることを保証している. この α の範囲は, 各時刻の解析ステップの前に, 予測残差誤差 $\tilde{e}_n^r = y_n - H\bar{\bar{v}}_n$ を用いて適応的に計算される. 彼らの理論は, モデルのダイナミクスに関する仮定を課さない. しかし, この理論は残差誤差 e_n^r に対する評価を保証するのみであり, 実際の誤差 $e_n = u_n - \bar{v}_n$ については保証していない. 一方で, 本研究の結果は, モデルのダイナミクスに関する定数 ρ および β を推定すれば, 実際の誤差 e_n に対する誤差評価を保証できる.

定義 6.8 における加法的インフレーションと比較すると, 定理 7.3 および 定理 7.4 の乗法的インフレーションはアンサンブルの分散を拡大する. その結果, 解析アンサンブルのすべてのメンバー \mathbf{V}_n に対する誤差評価を得ることはできない. 代わりに, 定義 5.5 を用いることで, 解析平均 \bar{v}_n に対する誤差評価を行う. これにより, (7.15) において追加の定数 D が導入される. しかし, この追加項は無視できる程度になり, ある極限においてフィルタリング誤差は観測ノイズのオーダーに収束する傾向がある.

系 7.3. 定理 7.3 の仮定の下で, 観測の精度が無限大に近づく極限 (すなわち $r \rightarrow 0$) では, フィルタリング誤差 (7.15) は次の関係を満たす:

$$\lim_{n \rightarrow \infty} \mathbb{E}[|e_n|^2] = O(r^2).$$

Proof. Θ は次のように表される：

$$\Theta = \left(\frac{r^2}{r^2 + \alpha^2 \lambda_*} \right)^2 = O(r^4),$$

また, $\theta = O(r^4)$ である. このとき,

$$\begin{aligned} \frac{1 - \Theta}{1 - \theta} - 1 &= (1 - \Theta)(1 + \theta + O(\Theta^2)) - 1 \\ &= 1 - \Theta + \theta + O(\Theta^2) - 1 \\ &= \Theta(e^{2(\beta+\epsilon)h} - 1) + O(\Theta^2) = O(r^4) \end{aligned}$$

が成り立つ. よって,

$$\lim_{n \rightarrow \infty} \mathbb{E}[|e_n|^2] = \frac{mr^2}{1 - \theta} + \left(\frac{1 - \Theta}{1 - \theta} - 1 \right) D = mr^2(1 + O(r^4)) + O(r^4) = O(r^2)$$

となる. \square

7.3 数値計算

観測システムシミュレーション実験 (Observing System Simulation Experiment, OSSE) または双子実験は, データ同化アルゴリズムを数値的に検証するための標準的な手法である. この手法では, 実際の観測データではなく, 数値モデルから直接生成された合成データのみを使用する. これにより, 不完全なモデルや未知の測定ノイズに関連する問題を回避できる.

定義 7.4 (OSSE). OSSE の標準的な手順は以下の通りである.

- (1) (6.2) の真の解 $(u_n)_{n=1}^N$ を数値的に計算する.
- (2) (6.4) に従い, ランダムな観測データ $(y_n)_{n=1}^N$ を生成する.
- (3) 観測データ $(y_n)_{n=1}^N$ を同化し, データ同化アルゴリズムを用いて解析値 $(\varpi_n)_{n=1}^N$ を得る.
- (4) 各時刻 $n = 1, \dots, N$ について, u_n と ϖ_n の誤差を計算する.

EnKFにおいて, 解析値はアンサンブル平均に対応し, すなわち $\varpi_n = \bar{v}_n$ である. 時刻 n における二乗誤差 (SE) は次のように定義される :

$$SE_n = |\varpi_n - u_n|^2. \quad (7.25)$$

これは、定理 7.3 の誤差評価 (7.14) および (7.15) で用いられる。また、時刻 n における平均二乗誤差の平方根 (RMSE) も利用される：

$$\text{RMSE}_n = \frac{|\varpi_n - u_n|}{\sqrt{N_u}} = \sqrt{\frac{\text{SE}_n}{N_u}}. \quad (7.26)$$

これは、モデルのダイナミクスの次元で正規化されたものである。これらの誤差には、定義 7.4 の手順 (1) で真の解を数値的に近似しているため、数値誤差と状態推定誤差の両方が含まれる。

OSSE の例として L96 方程式 (5.6) を考える。ここでは、パラメータとして $J = 40$, $f = 8$ を設定する。これらの値のもとで、定義 5.8 に示されるように、定義 5.2 の不等式 (5.1) において

$$\beta = 2\rho - 1 = 2\sqrt{2J}f - 1 > 0$$

が成り立つため、L96 方程式はカオス的な振る舞いを示し得る。

まず、時間ステップ幅 $\Delta t = 0.01$ で 4 次の Runge-Kutta 法を用いて、 $N = 14400$ ステップ分の解 $u(t)$ を $T = N\Delta t$ まで計算する。初期条件は次のように設定する：

$$\begin{aligned} u_0 &= (f * 1.001, f, \dots, f)^* \in \mathbb{R}^J \\ &= (8.008, 8.0, \dots, 8.0)^* \in \mathbb{R}^{40}. \end{aligned}$$

時間区間 $0 \leq t \leq 144$ における解 $u(t)$ の第 1 成分 $u^1(t)$ の時間発展を 図 5(a) に示す。また、第 1 成分と第 2 成分 $(u^1(t), u^2(t)) \in \mathbb{R}^2$ への射影を 図 5(b) に示す。

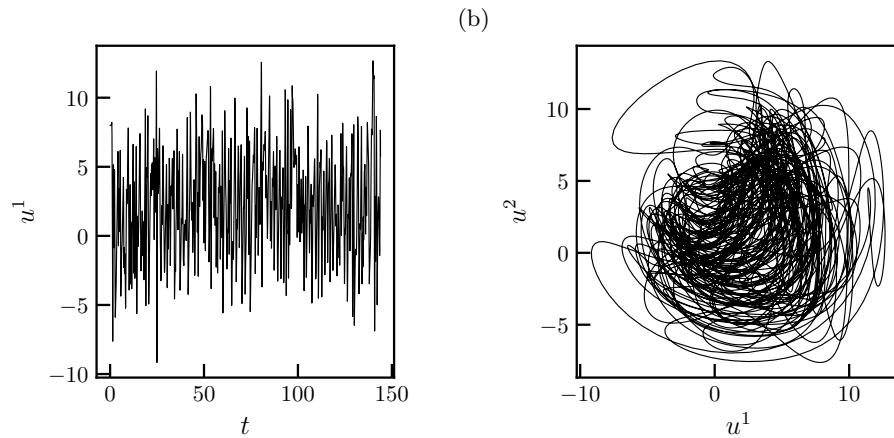


図 5: L96 方程式 (5.6) の解 $((J, f) = (40, 8))$ 。

また、解の有界性を確認するため、 $|u(t)|/\sqrt{J}$ の時間発展を 図 6 に示す。初期段階ではノルムが急激に減少し、これは L96 方程式の散逸性を示している。その後、ノルムは概ね 4.0 の値の周りで非周期的に変動しており、軌道が有界であることが確認できる。

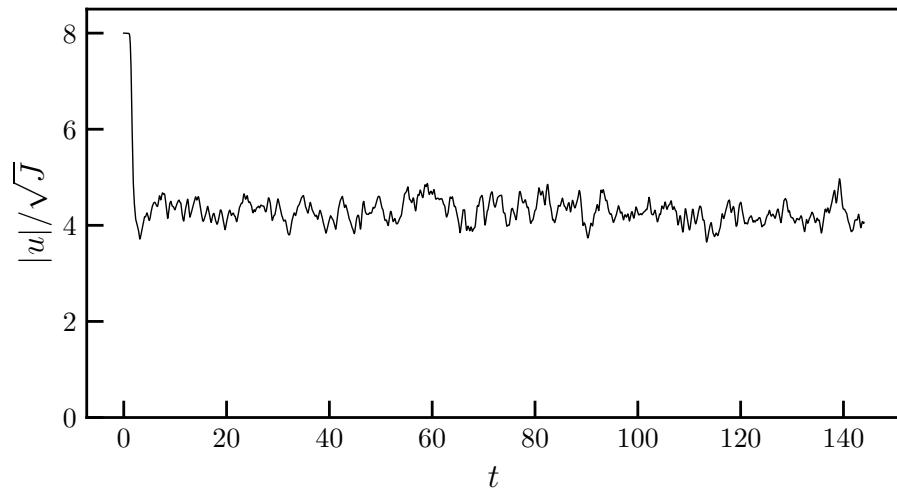


図 6: ノルム $|u(t)|/\sqrt{J}$ の時間発展.

L96 方程式の 2 つの解 $u(t)$ と $u^\epsilon(t)$ を考える。それぞれの初期状態を $u(0)$ および $u^\epsilon(0)$ とし、 $|u^\epsilon(0) - u(0)|/\sqrt{J} \approx \epsilon$ となるように設定する。ここで、 $\epsilon > 0$ は小さなスケールを表す。図 7 には、正規化誤差 $|\delta u(t)|/\sqrt{J}$ の時間発展を示す。ここで、 $\delta u(t) = u(t) - u^\epsilon(t)$ とし、 $\epsilon = 10^{-4}$ を用いる。この結果は、L96 方程式の解が初期擾動に対して鋭敏性を持ち、カオス的な振る舞いを示すことを示唆している。

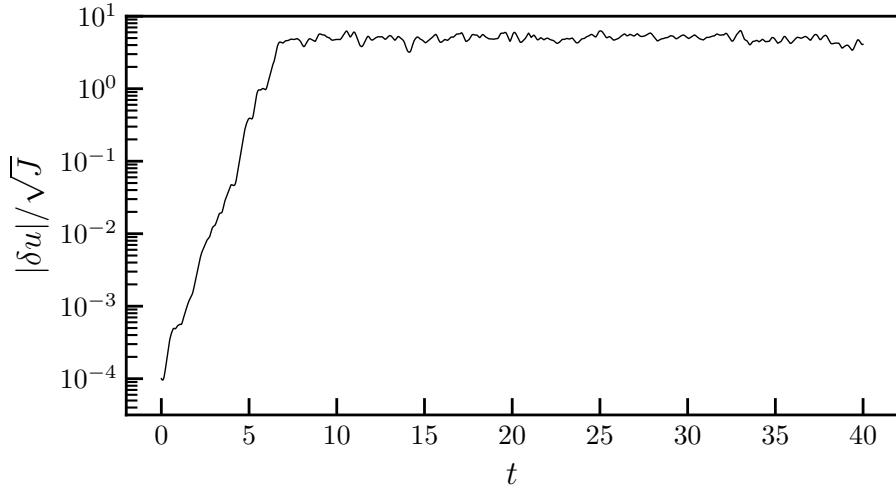


図 7: 正規化誤差 $|\delta u(t)|/\sqrt{J}$ の時間発展 ($\epsilon = 10^{-4}$) の対数プロット.

L96 方程式に対して ETKF を適用し、前節の解析を検証する。時間間隔 $\tau = 5\Delta t = 0.05$ に対し、前述の数値解 $u(t)$ を用いて写像 Ψ_τ を近似する。まず、 $N_0 = 1440$ に対して $t = N_0\tau$ までの解を破棄し、L96 方程式の吸収球内に収まるようにする。次に、隠れた真の状態 $(u_n)_{n=1}^{N_t}$ を次のようにサンプリングする：

$$u_n = u(n\tau + N_0\tau), \quad n = 1, \dots, N_t.$$

ここで、 $N_t = 480$ とする。この処理は 定義 7.4 の OSSE のステップ (1) に対応する。

次に、観測システム (6.4) に関して 定義 3.2 を仮定し、観測系列 $(y_n)_{n=1}^{N_t}$ を以下のように生成する：

$$y_n = u_n + \xi_n, \quad \xi_n \sim N(0, r^2 I), \quad n = 1, \dots, N_t,$$

ここで、観測ノイズの分散は $r^2 = 0.1$ とする。この処理は 定義 7.4 の OSSE のステップ (2) に対応する。OSSE のステップ (3) において、アンサンブルサイズ $m = J + 1$ の ETKF を適用する。初期アンサンブルは、 \mathbb{R}^J の標準基底 e_i ($i = 1, \dots, J$) を用いて

$$\mathbf{V}_0 = [e_1, e_2, \dots, e_J, -\sum_{i=1}^J e_i] \in \mathbb{R}^{J \times m}$$

とする。これにより、定理 7.3 の仮定 $\lambda_{min}(C_0) > 0$ が満たされる。この \mathbf{V}_0 および観測データ $(y_n)_{n=1}^{N_t}$ を用いて、ETKF の予測アンサンブル $(\hat{\mathbf{V}}_n)_{n=1}^{N_t}$ および解析アンサンブル $(\mathbf{V}_n)_{n=1}^{N_t}$ を計算する。ここでは、定義 4.20 に示される乗法的共分散インフレーションを用い、 $\alpha = 1.0, 1.1, 5.0$ の 3 つの値について実験を行う。

期待値 $\mathbb{E}[\text{SE}_n]$ は、観測ノイズを生成するための異なる 20 個の乱数シードに対する平均を用いて近似する。図 8(a) には、 $\alpha = 1.0, 1.1, 5.0$ の場合の $\mathbb{E}[\text{SE}_n]$ を示す。また、誤差成長が十分に抑制されるようなインフレーション ($\theta \ll 1$) のもとでの誤差評価式 (7.15) も示す。この場合、誤差評価は Jr^2 によって近似される。

さらに、図 8(b) には、予測共分散の最小固有値 $\hat{\lambda}_{min} = \lambda_{min}(\hat{P}_n)$ を示す。この図から、インフレーションなし ($\alpha = 1.0$) では $\mathbb{E}[\text{SE}_n]$ が理論的評価値よりも大きい (図 8(a))。また $\hat{\lambda}_{min} \approx 10^{-10}$ であり、観測ノイズの分散 $r^2 = 0.1$ に比べて無視できるほど小さいことが分かる (図 8(b))。

一方、インフレーションパラメータが大きい場合 ($\alpha = 5.0$)、 $\hat{\lambda}_{min}$ は $\approx 10^{-2}$ で下から抑えられ、 $\alpha = 1.0$ の場合に比べて約 8 衡大きい値を取る (図 8(b))。その結果、初期時刻を除けば $\mathbb{E}[\text{SE}_n]$ は理論的評価値以下、もしくは同程度となる (図 8(a))。この結果は、理論評価の近似誤差や数値誤差を考慮しても定理 7.3 を支持する。

さらに、比較的小さいインフレーションパラメータ ($\alpha = 1.1$) の場合、 $\mathbb{E}[\text{SE}_n]$ は $\alpha = 5.0$ の場合よりも小さい値を示すが、 $\hat{\lambda}_{min}$ は依然として小さい。したがって、本理論では最適な α を決定することはできない。しかし、大きなインフレーションパラメータと最小固有値の下限が誤差評価 (7.15) を得るために十分条件であるため、本数値結果は理論と整合的である。

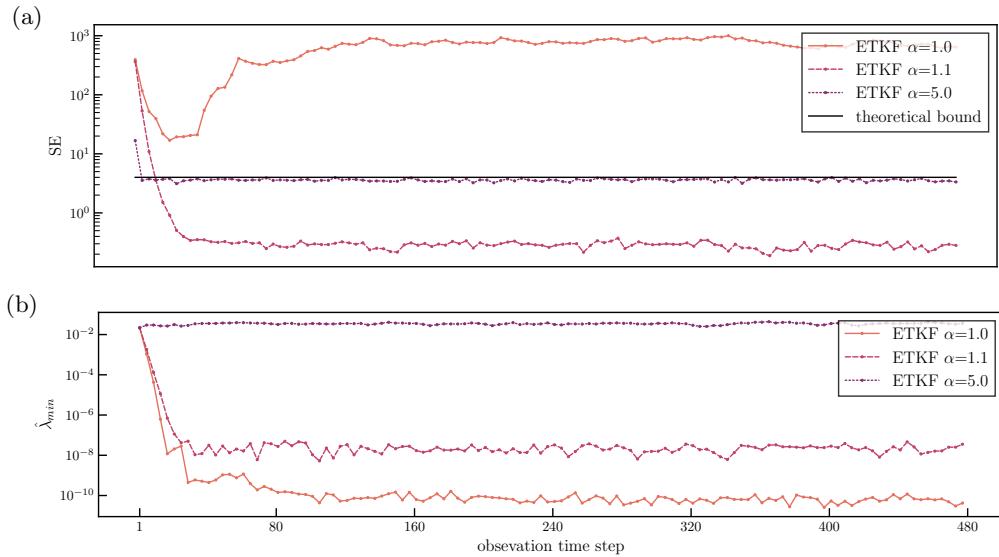


図 8: (a) $\mathbb{E}[\text{SE}_n]$ の時間発展、および (b) 予測共分散の最小固有値 $\hat{\lambda}_{min}$ の観測時間ステップ n に対するプロット (ETKF における乗法的インフレーションパラメータ $\alpha = 1.0, 1.1, 5.0$ の場合)。

次に、時間平均 SE の観測ノイズ分散 r^2 に対する依存性を調べる。図 9 には、 $\alpha = 1.1, 5.0$ における時間平均 SE の r^2 に対する両対数プロットを示す。ここで、 $r = 10^{-5}, \dots, 10^{-1}$

とし、各パラメータの組 (r, α) に対して 1 つの乱数シードを用いて SE を計算する。 $\alpha = 5.0$ の場合、定義 7.3 で示した理論的評価 $SE = O(r^2)$ が検証される。

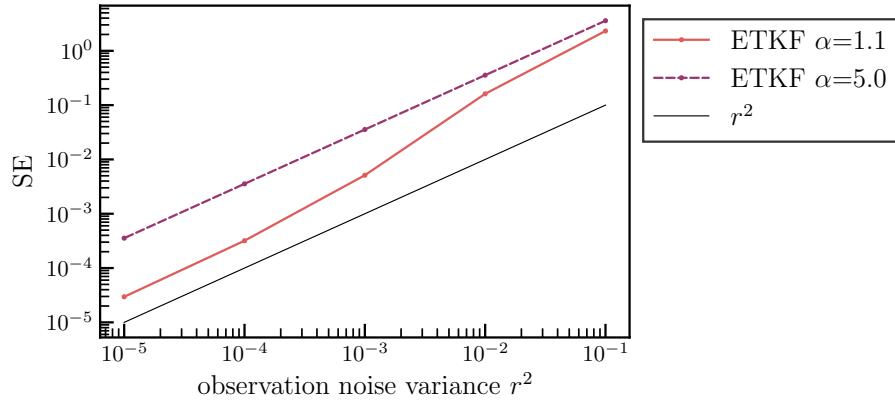


図 9: ETKF における乗法的インフレーション ($\alpha = 1.1, 5.0$) のもとでの時間平均 SE と r^2 の両対数プロット ($r = 10^{-5}, \dots, 10^{-1}$)。

8. まとめと今後の展望

本論文では、ヒルベルト空間上の散逸力学系に対して EnKF の理論的側面を研究した。EnKF に関しては、KF との整合性、指数安定性、および理想条件下での誤差評価が得られている。特に、本論文の主な貢献は以下の 3 点に要約される [82]。

第一に、ESRF のフィルタリング誤差が無限次元の状態空間においても任意の有限時間で有界であることを証明した。第二に、状態空間が有限次元の場合、乗法的インフレーションを用いることで、一様時間誤差評価を得るために必要なインフレーションパラメータ α の最小値を決定した。これらは、PO 法に対する解析を ESRF の場合へと拡張した理論的成果である。最後に、数値例を通じて誤差評価の妥当性を検証し、 α の下限が改良可能であることを示した。

今後の研究の方向性を示す。まず、乗法的インフレーションを用いた ESRF に関する誤差評価（定理 7.3 および定理 7.4）は、PO 法における乗法的アンサンブルインフレーション（定義 4.20 の (1')）のケースへと拡張できると考えられる。このためには、PO 法の解析ステップにおいてランダムに生成されるアンサンブル共分散 P_n の固有値を適切に評価する必要がある。

次に、現状の EnKF の誤差解析は現実的な応用には適さない全観測の場合に限定されている。部分観測のもとでの誤差評価については、表 1 における 3DVar の結果を EnKF に拡張できる可能性がある。例えば、L63 方程式を考えると、部分観測作用素 (5.14) を用いた PO 法において、加法的インフレーションを適用すれば誤差評価が得られるのではないかと予想される。

さらに、アンサンブルサイズが状態空間の次元よりも大きいと仮定しているが、これは理論的な制約である。実際の EnKF の応用では、計算資源の制約によりこの仮定は成り立たないことが多い。小さなアンサンブルサイズで誤差評価を得るために、力学系の次元削減が必要となる。例えば、トーラス上の 2D-NSE における決定モードは有限次元部分空間を定義し、この部分空間からの部分観測によって全空間の状態を再構成できる。したがって、アンサンブルサイズがこの部分空間の次元よりも大きい場合、誤差評価が得られる可能性がある。

最後に、本論文では力学モデルの時間発展が数値誤差なしに近似できることを仮定しているが、これは節 6.2 で議論したように理想的な前提である。今後の研究では、データ同化アルゴリズムの誤差解析に数値誤差を含めるべきであり、数値解析の理論を活用することが求められる。関連するトピックとして、無限次元空間で定式化されたベイズ逆問題の離散化に関する研究 [27] や、不完全モデルを扱う手法 [76] などが挙げられる。

参考文献

- [1] O. AL-GHATTAS AND D. SANZ-ALONSO, *Non-asymptotic analysis of ensemble Kalman updates: Effective dimension and localization*, Information and Inference: A Journal of the IMA, 13 (2024), p. iaad043.
- [2] D. A. F. ALBANEZ, H. J. N. LOPES, AND E. S. TITI, *Continuous data assimilation for the three-dimensional Navier-Stokes-alpha model*, Asymptot. Anal., 97 (2016), pp. 139–164.
- [3] C. D. ALIPRANTIS AND K. C. BORDER, *Infinite Dimensional Analysis: A Hitch-hiker's Guide*, Springer Berlin, Heidelberg, 3 ed., 2006. <https://link.springer.com/book/10.1007/978-3-540-29587-9>.
- [4] B. D. O. ANDERSON AND J. B. MOORE, *Optimal Filtering*, Courier Corporation, May 2012.
- [5] J. L. ANDERSON, *An Ensemble Adjustment Kalman Filter for Data Assimilation*, Mon. Weather Rev., 129 (2001), pp. 2884–2903.
- [6] ——, *An adaptive covariance inflation error correction algorithm for ensemble filters*, Tellus Dyn. Meteorol. Oceanogr., 59 (2007), pp. 210–224.
- [7] ——, *Spatially and temporally varying adaptive covariance inflation for ensemble filters*, Tellus Dyn. Meteorol. Oceanogr., 61 (2009), pp. 72–83.
- [8] J. L. ANDERSON AND S. L. ANDERSON, *A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts*, Mon. Weather Rev., 127 (1999), pp. 2741–2758.
- [9] R. ATAR AND O. ZEITOUNI, *Exponential stability for nonlinear filtering*, Annales de l'Institut Henri Poincaré (B) Probability and Statistics, 33 (1997), pp. 697–725.
- [10] A. AZOUANI, E. OLSON, AND E. S. TITI, *Continuous Data Assimilation Using General Interpolant Observables*, J Nonlinear Sci, 24 (2014), pp. 277–304.
- [11] A. AZOUANI AND E. S. TITI, *Feedback control of nonlinear dissipative systems by finite determining parameters- A reaction-diffusion paradigm*, EECT, 3 (Wed Oct 01 00:00:00 UTC 2014), pp. 579–594.

- [12] K. BERGEMANN AND S. REICH, *An ensemble Kalman-Bucy filter for continuous data assimilation*, Meteorol. Z., 21 (2012), pp. 213–219.
- [13] A. N. BISHOP AND P. DEL MORAL, *On the mathematical theory of ensemble (linear-Gaussian) Kalman–Bucy filtering*, Math. Control Signals Syst., 35 (2023), pp. 835–903.
- [14] C. H. BISHOP, B. J. ETHERTON, AND S. J. MAJUMDAR, *Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects*, Mon. Weather Rev., 129 (2001), pp. 420–436.
- [15] D. BLÖMKER, K. LAW, A. M. STUART, AND K. C. ZYGALAKIS, *Accuracy and stability of the continuous-time 3DVAR filter for the Navier–Stokes equation*, Nonlinearity, 26 (2013), p. 2193.
- [16] C. E. A. BRETT, K. F. LAM, K. J. H. LAW, D. S. MCCORMICK, M. R. SCOTT, AND A. M. STUART, *Stability of Filters for the Navier-Stokes Equation*, Oct. 2011.
- [17] ——, *Accuracy and stability of filters for dissipative PDEs*, Physica D: Nonlinear Phenomena, 245 (2013), pp. 34–45.
- [18] G. L. BROWNING, W. D. HENSHAW, AND H.-O. KREISS, *A numerical investigation of the interaction between the large and small scales of the two-dimensional incompressible Navier–Stokes equations*, Res. Rep. -UR-98-1712 Los Alamos Natl. Lab., (1998).
- [19] G. BURGERS, P. J. VAN LEEUWEN, AND G. EVENSEN, *Analysis Scheme in the Ensemble Kalman Filter*, Mon. Weather Rev., 126 (1998), pp. 1719–1724.
- [20] A. CARRASSI, M. BOCQUET, L. BERTINO, AND G. EVENSEN, *Data assimilation in the geosciences: An overview of methods, issues, and perspectives*, WIREs Clim. Change, 9 (2018), p. e535.
- [21] J. CHARNEY, M. HALEM, AND R. JASTROW, *Use of Incomplete Historical Data to Infer the Present State of the Atmosphere*, J. Atmospheric Sci., 26 (1969). https://journals.ametsoc.org/view/journals/atsc/26/5/1520-0469_1969_026_1160_uoihdt_2_0_co_2.xml.
- [22] A. CHESKIDOV, D. D. HOLM, E. OLSON, AND E. S. TITI, *On a Leray- α model of turbulence*, Proc. R. Soc. Math. Phys. Eng. Sci., 461 (2005), pp. 629–649.

- [23] A. J. CHORIN AND J. E. MARSDEN, *A Mathematical Introduction to Fluid Mechanics*, vol. 4 of Texts in Applied Mathematics, Springer, New York, NY, 1993.
- [24] C. K. CHUI AND G. CHEN, *Kalman Filtering*, Springer International Publishing, Cham, 2017.
- [25] J. B. CONWAY, *A Course in Functional Analysis*, vol. 96 of Graduate Texts in Mathematics, Springer, New York, NY, 2007.
- [26] S. L. COTTER, M. DASHTI, J. C. ROBINSON, AND A. M. STUART, *Bayesian inverse problems for functions and applications to fluid mechanics*, Inverse Problems, 25 (2009), p. 115008.
- [27] S. L. COTTER, M. DASHTI, AND A. M. STUART, *Approximation of Bayesian Inverse Problems for PDEs*, SIAM J. Numer. Anal., 48 (2010), pp. 322–345.
- [28] R. DALEY, *Atmospheric Data Analysis*, Cambridge University Press, Cambridge, UK, 1991.
- [29] M. DASHTI AND A. M. STUART, *The Bayesian Approach to Inverse Problems*, in Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, and H. Owhadi, eds., Springer International Publishing, Cham, 2017, pp. 311–428.
- [30] J. DE WILJES, S. REICH, AND W. STANNAT, *Long-Time Stability and Accuracy of the Ensemble Kalman-Bucy Filter for Fully Observed Processes and Small Measurement Noise*, Siam J. Appl. Dyn. Syst., 17 (2018), pp. 1152–1181.
- [31] J. DE WILJES AND X. T. TONG, *Analysis of a localised nonlinear ensemble Kalman Bucy filter with complete and accurate observations*, Nonlinearity, 33 (2020), pp. 4752–4782.
- [32] L. DIECI AND T. EIROLA, *On smooth decompositions of matrices*, Siam J. Matrix Anal. Appl., 20 (1999), pp. 800–819.
- [33] G. EVENSEN, *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics*, J. Geophys. Res. Oceans, 99 (1994), pp. 10143–10162.
- [34] ———, *Data Assimilation: The Ensemble Kalman Filter*, Springer, Berlin, Heidelberg, 2009.

- [35] 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*, J. Dyn. Differ. Equ., 14 (2001).
- [36] C. FOIAS, O. MANLEY, R. ROSA, AND R. TEMAM, *Navier-Stokes Equations and Turbulence*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, Cambridge, 2001.
- [37] A. L. GIBBS AND F. E. SU, *On Choosing and Bounding Probability Metrics*, Int. Stat. Rev. Rev. Int. Stat., 70 (2002), pp. 419–435.
- [38] Y. GIGA AND A. NOVOTNÝ, *Handbook of Mathematical Analysis in Mechanics of Viscous Fluids*, Springer Cham, 1 ed., 2018.
- [39] G. H. GOLUB AND C. F. V. LOAN, *Matrix Computations*, Johns Hopkins University Press, 2013.
- [40] T. M. HAMILL, J. S. WHITAKER, AND C. SNYDER, *Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter*, Mon. Weather Rev., 129 (2001), pp. 2776–2790.
- [41] K. HAYDEN, E. OLSON, AND E. S. TITI, *Discrete data assimilation in the Lorenz and 2D Navier-Stokes equations*, Phys. -Nonlinear Phenom., 240 (2011), pp. 1416–1425.
- [42] W. D. HENSHAW, H.-O. KREISS, AND J. YSTRÖM, *Numerical Experiments on the Interaction Between the Large- and Small-Scale Motions of the Navier-Stokes Equations*, Multiscale Model. Simul., 1 (2003), pp. 119–149.
- [43] M. HLADNIK AND M. OMLADIC, *Spectrum of the Product of Operators*, Proc. Amer. Math. Soc., 102 (1988), pp. 300–302.
- [44] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [45] P. L. HOUTEKAMER AND F. ZHANG, *Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation*, Mon. Weather Rev., 144 (2016), pp. 4489–4532.
- [46] B. R. HUNT, E. J. KOSTELICH, AND I. SZUNYOGH, *Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter*, Physica D: Nonlinear Phenomena, 230 (2007), pp. 112–126.

- [47] M. A. IGLESIAS, K. J. H. LAW, AND A. M. STUART, *Ensemble Kalman methods for inverse problems*, Inverse Problems, 29 (2013), p. 045001.
- [48] D. A. JONES AND E. S. TITI, *Upper Bounds on the Number of Determining Modes, Nodes, and Volume Elements for the Navier-Stokes Equations*, Indiana Univ. Math. J., 42 (1993), pp. 875–887. <https://www.jstor.org/stable/24897124>.
- [49] R. E. KALMAN, *A New Approach to Linear Filtering and Prediction Problems*, J. Basic Eng, 82 (1960), pp. 35–45.
- [50] E. KALNAY, *Atmospheric Modeling, Data Assimilation and Predictability*, Nov. 2002.
- [51] I. KASANICKY, *Ensemble Kalman Filter on High and Infinite Dimensional Spaces*, PhD thesis, Charles University, Prague, Dec. 2016.
- [52] T. KATO, *Perturbation Theory for Linear Operators*, vol. 132 of Classics in Mathematics, Springer, Berlin, Heidelberg, 1995.
- [53] A. KATOK AND B. HASSELBLATT, *Introduction to the Modern Theory of Dynamical Systems*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, Cambridge, 1995.
- [54] 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 (2014), pp. 2579–2603.
- [55] P. KORN, *Data assimilation for the Navier–Stokes- $\{\alpha\}$ equations*, Physica D: Nonlinear Phenomena, 238 (2009), pp. 1957–1974.
- [56] ——, *On degrees of freedom of certain conservative turbulence models for the Navier–Stokes equations*, Journal of Mathematical Analysis and Applications, 378 (2011), pp. 49–63.
- [57] S. KOTSUKI, Y. OTA, AND T. MIYOSHI, *Adaptive covariance relaxation methods for ensemble data assimilation: Experiments in the real atmosphere*, Q. J. R. Meteorol. Soc., 143 (2017), pp. 2001–2015.
- [58] E. KWIATKOWSKI AND J. MANDEL, *Convergence of the Square Root Ensemble Kalman Filter in the Large Ensemble Limit*, Siam-Asa J. Uncertain. Quantif., 3 (2015), pp. 1–17.

- [59] W. LAMB, *Robinson, J. C.**Infinite-dimensional dynamical systems (Cambridge University Press, 2001) 461pp., 0 521 63564 0 (paperback), £24.95, 0 521 63204 8 (hardback), £70*, Proc. Edinb. Math. Soc., 46 (2003), pp. 252–253.
- [60] T. LANGE AND W. STANNAT, *Mean field limit of Ensemble Square Root filters - discrete and continuous time*, FoDS, 3 (Thu Jan 21 19:00:00 EST 2021), pp. 563–588.
- [61] K. LAW, A. SHUKLA, AND A. STUART, *ANALYSIS OF THE 3DVAR FILTER FOR THE PARTIALLY OBSERVED LORENZ'63 MODEL*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1061–1078.
- [62] 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*, Phys. -Nonlinear Phenom., 325 (2016), pp. 1–13.
- [63] K. J. H. LAW, A. M. STUART, AND K. C. ZYGALAKIS, *Data Assimilation: A Mathematical Introduction*, Springer, 2015.
- [64] W. LAYTON, C. C. MANICA, M. NEDA, AND L. G. REBOLZ, *Numerical analysis and computational comparisons of the NS-alpha and NS-omega regularizations*, Computer Methods in Applied Mechanics and Engineering, 199 (2010), pp. 916–931.
- [65] E. N. LORENZ, *Deterministic Nonperiodic Flow*, J. Atmospheric Sci., 20 (1963), pp. 130–141. https://journals.ametsoc.org/view/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.xml.
- [66] ———, *Predictability: A problem partly solved*, Proc Semin. Predict. Read. UK ECMWF 1996, 1 (1996), pp. 1–18.
- [67] E. N. LORENZ AND K. A. EMANUEL, *Optimal Sites for Supplementary Weather Observations: Simulation with a Small Model*, J. Atmospheric Sci., 55 (1998), pp. 399–414.
- [68] X. LUO AND I. HOTEIT, *Covariance Inflation in the Ensemble Kalman Filter: A Residual Nudging Perspective and Some Implications*, Mon. Weather Rev., 141 (2013), pp. 3360–3368.
- [69] A. J. MAJDA AND J. HARLIM, *Filtering Complex Turbulent Systems*, Filter. Complex Turbul. Syst., (2012), pp. 1–357.

- [70] J. MANDEL, L. COBB, AND J. D. BEEZLEY, *On the convergence of the ensemble Kalman filter*, Appl. Math., 56 (2011), pp. 533–541.
- [71] J. E. MARSDEN AND S. SHKOLLER, *Global well-posedness for the Lagrangian averaged Navier–Stokes (LANS- α) equations on bounded domains*, Philos. Trans. R. Soc. Lond. Ser. Math. Phys. Eng. Sci., 359 (2001), pp. 1449–1468.
- [72] T. MIYOSHI, *The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter*, Mon. Weather Rev., 139 (2011), pp. 1519–1535.
- [73] B. ØKSENDAL, *Stochastic Differential Equations*, Universitext, Springer, Berlin, Heidelberg, 2003.
- [74] E. OLSON AND E. S. TITI, *Determining Modes for Continuous Data Assimilation in 2D Turbulence*, Journal of Statistical Physics, 113 (2003), pp. 799–840.
- [75] K. B. PETERSEN AND M. S. PEDERSEN, *The Matrix Cookbook*. <http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html>, Nov. 2012.
- [76] S. REICH AND C. COTTER, *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press, Cambridge, 2015.
- [77] F. RELLICH, *Perturbation theory of eigenvalue problems*. <https://archive.org/details/perturbationtheo00rell>.
- [78] D. SANZ-ALONSO, A. STUART, AND A. TAEB, *Inverse Problems and Data Assimilation*, London Mathematical Society Student Texts, Cambridge University Press, Cambridge, 2023.
- [79] A. M. STUART, *Inverse problems: A Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559.
- [80] T. SULLIVAN, *Introduction to Uncertainty Quantification*, vol. 63 of Texts in Applied Mathematics, Springer International Publishing, Cham, 2015.
- [81] K. TAKEDA, *Lorenz System*. <https://kotatakedo.github.io/lorenz-webgl/>, 2024.
- [82] K. TAKEDA AND T. SAKAJO, *Uniform error bounds of the ensemble transform Kalman filter for chaotic dynamics with multiplicative covariance inflation*, SIA-MASA J. Uncertain. Quantif., (2024, in press).

- [83] T.-J. TARN AND Y. RASIS, *Observers for nonlinear stochastic systems*, IEEE Trans. Autom. Control, 21 (1976), pp. 441–448.
- [84] R. TEMAM, *Navier-Stokes Equations and Nonlinear Functional Analysis*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Jan. 1995.
- [85] ———, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer New York, NY, 1997.
- [86] M. K. TIPPETT, J. L. ANDERSON, C. H. BISHOP, T. M. HAMILL, AND J. S. WHITAKER, *Ensemble Square Root Filters*, Mon. Weather Rev., 131 (2003). https://journals.ametsoc.org/view/journals/mwre/131/7/1520-0493_2003_131_1485_esrf_2.0.co_2.xml.
- [87] X. T. TONG, *Performance Analysis of Local Ensemble Kalman Filter*, J Nonlinear Sci, 28 (2018), pp. 1397–1442.
- [88] X. T. TONG, A. J. MAJDA, AND D. KELLY, *Nonlinear stability and ergodicity of ensemble based Kalman filters*, Nonlinearity, 29 (2016), pp. 657–691.
- [89] X. T. TONG, A. J. MAJDA, AND D. KELLY, *Nonlinear stability of the ensemble Kalman filter with adaptive covariance inflation*, Comm. Math. Sci., 14 (2016), pp. 1283–1313.
- [90] X. T. TONG AND M. MORZFIELD, *Localized ensemble Kalman inversion*, Inverse Problems, 39 (2023), p. 064002.
- [91] J. S. WHITAKER AND T. M. HAMILL, *Ensemble Data Assimilation without Perturbed Observations*, Mon. Weather Rev., 130 (2002), pp. 1913–1924.
- [92] ———, *Evaluating Methods to Account for System Errors in Ensemble Data Assimilation*, Mon. Weather Rev., 140 (2012), pp. 3078–3089.
- [93] Y. YING AND F. ZHANG, *An adaptive covariance relaxation method for ensemble data assimilation*, Q. J. R. Meteorol. Soc., 141 (2015), pp. 2898–2906.
- [94] F. ZHANG, C. SNYDER, AND J. SUN, *Impacts of Initial Estimate and Observation Availability on Convective-Scale Data Assimilation with an Ensemble Kalman Filter*, Mon. Weather Rev., 132 (2004), pp. 1238–1253.