

データ同化の数理 (Ver. 0.4)

竹田航太

2025 年 11 月 19 日

■更新情報

- 2025/10/09: Ver. 0.1
- 2025/10/20: Ver. 0.2
- 2025/10/23: Ver. 0.3
- 2025/11/19: Ver. 0.4

目次

第 I 部	はじめに	5
第 1 章	データ同化の概要	7
第 2 章	読者のためのガイド	9
2.1	本書の位置付け	9
2.2	記号リスト	9
第 II 部	データ同化の数学的準備	11
第 3 章	線形逆問題	13
3.1	本章の要点	13
3.2	行列	13
3.3	線形観測	13
3.4	最小二乗法	14
第 4 章	確率・統計とベイズ逆問題	17
4.1	本章の要点	17
4.2	確率論	17
4.3	確率的線形逆問題	18
4.4	ベイズ逆問題	18
4.5	グラフィカルモデル	18
第 5 章	微分方程式と力学系	19
5.1	本章の要点	19
5.2	常微分方程式	19
第 6 章	常微分方程式の数値計算法	21
6.1	本章の要点	21

第 III 部	データ同化問題	23
第 7 章	データ同化問題の定式化	25
7.1	本章の要点	25
7.2	状態空間	25
7.3	ベイズデータ同化問題	27
第 8 章	連続時間フィルタリング問題	31
8.1	本章の要点	31
第 9 章	連続データ同化	33
9.1	本章の要点	33
第 IV 部	データ同化アルゴリズム	35
第 10 章	フィルタリングアルゴリズム	37
10.1	本章の要点	37
10.2	カルマンフィルタとその拡張	37
10.3	粒子フィルタ	39
10.4	改良技術	39
第 11 章	スムージングアルゴリズム	41
11.1	本章の要点	41
11.2	変分法	41
11.3	MCMC	41
第 V 部	数学解析	43
第 VI 部	応用例と実装	45
	参考文献	49

第Ⅰ部

はじめに

第 1 章

データ同化の概要

本書は洋書 [5, 6] などをもとに作成している.

第 2 章

読者のためのガイド

2.1 本書の位置付け

本書に含まれていない内容については洋書を参照せよ．例えば，確率微分方程式のフィルタリング理論については [3] がある．

(WIP)

2.2 記号リスト

\mathbb{R} 実数全体の集合

\mathbb{R}^n n 次元ユークリッド空間

$\mathbb{R}^{m \times n}$ 実 $m \times n$ 行列全体の集合

\mathbb{N} 自然数全体の集合

\mathbb{Z} 整数全体の集合

\mathbb{E} 期待値

Ω 標本空間

\mathcal{F} σ -代数

\mathbb{P} 確率測度

\mathcal{U} 状態空間（一般の場合）

N_u 状態空間の次元（有限次元の場合）

\mathbf{u} 状態ベクトル

\mathcal{Y} 観測空間（一般の場合）

N_y 観測空間の次元（有限次元の場合）

\mathbf{y} 観測ベクトル

h 観測関数

H 観測作用素（行列）

$\mathbf{u}_{k:l}$ ベクトルの集合（時刻 k から l まで）

$\text{diag}(d_1, \dots, d_n)$ 対角成分を d_1, \dots, d_n とする対角行列

$\mathbf{u}(t; \mathbf{u}_0)$ 初期値 \mathbf{u}_0 に対する常微分方程式の解

第 II 部

データ同化の数学的準備

第 3 章

線形逆問題

3.1 本章の要点

データ同化では、観測データから状態ベクトルを推定することが基本的な目的である。数学的には、観測という操作に対する逆問題として定式化される。本章では、その中でも基本的な線形逆問題を扱う。

3.2 行列

行列 $A \in \mathbb{R}^{N_1 \times N_2}$ について、 $\text{rank}(A) = \min(N_1, N_2)$ であるとき、 A はフルランクであるといい、そうでない時はランク落ちもしくは退化しているという。行列の計算については、[2]などを参照せよ。

3.3 線形観測

状態ベクトル $\mathbf{u} \in \mathbb{R}^{N_u}$ から観測ベクトル $\mathbf{y} \in \mathbb{R}^{N_y}$ を得る過程を観測という。

定義 3.1 (線形観測). 観測行列 $H \in \mathbb{R}^{N_y \times N_u}$ を用いて、観測ベクトル $\mathbf{y} \in \mathbb{R}^{N_y}$ が以下のように得られるとき、線形観測という。

$$\mathbf{y} = H\mathbf{u} + \boldsymbol{\eta}. \quad (3.1)$$

ここで、 $\boldsymbol{\eta} \in \mathbb{R}^{N_y}$ は観測ノイズを表すベクトルである。特に、ノイズがない場合 ($\boldsymbol{\eta} = \mathbf{0}$) の場合はノイズなし線形観測という。

観測行列 H がフルランクであるとき、ノイズがない線形観測値 \mathbf{y} に逆行列 H^{-1} をかけることで、状態ベクトル $H^{-1}\mathbf{y} = H^{-1}H\mathbf{u} = \mathbf{u}$ が復元される。このように、線形観測から状態ベクトルを推定することを線形逆問題という。一般に、逆問題は不良設定問題であることが多い。つまり、解が存在しない、一意に定まらない、もしくは解が \mathbf{y} の変化に敏感に依存するような場合がある。

例 3.1 (ノイズなし線形観測の不良設定の例). $N_u = 2$, $N_y = 1$ のとき, 観測行列 H が以下のような場合を考える.

$$H = \begin{pmatrix} 1 & 0 \end{pmatrix}. \quad (3.2)$$

このとき, 観測ベクトル \mathbf{y} は以下のように表される.

$$\mathbf{y} = H\mathbf{u} = (u_1). \quad (3.3)$$

ここで, u_1 は \mathbf{u} の第 1 成分である. このとき, $\mathbf{u} = (u_1, u_2)$ と $\mathbf{u}' = (u_1, u'_2)$ は同じ \mathbf{y} を与える.

3.4 最小二乗法

線形逆問題をある種の最適化問題として定式化する際に, 最小二乗法がよく用いられる. ここでは, 最小二乗法を幾何学的・代数的な視点から説明する.

補題 3.1 (射影定理). 空でない閉凸部分集合 $K \subset \mathbb{R}^{N_u}$ と $\mathbf{y} \in \mathbb{R}^{N_u}$ を考える. このとき, \mathbf{y} の K への射影 $\mathbf{u}^* = \Pi_K \mathbf{y}$ が一意に定まる. つまり, 以下の最小二乗解が一意に定まる.

$$\mathbf{u}^* = \operatorname{argmin}_{\mathbf{u} \in K} \|\mathbf{y} - \mathbf{u}\|^2. \quad (3.4)$$

Proof. (WIP) □

補題 3.2 (残差の直交性). 閉部分空間 $W \subset \mathbb{R}^{N_u}$ と $\mathbf{b} \in \mathbb{R}^{N_u}$ に対して, 以下は同値.

$$\mathbf{u}^* = \operatorname{argmin}_{\mathbf{u} \in W} \|\mathbf{u} - \mathbf{b}\|^2 \Leftrightarrow (\mathbf{u}^* - \mathbf{b}) \perp W. \quad (3.5)$$

Proof. (WIP) □

補題 3.3 (正規方程式). 観測行列 $H : \mathbb{R}^{N_y \times N_u}$ と観測ベクトル $\mathbf{y} \in \mathbb{R}^{N_y}$ に対し, 以下は同値.

$$\mathbf{u}^* \in \operatorname{argmin}_{\mathbf{u} \in \mathbb{R}^{N_u}} \|\mathbf{y} - H\mathbf{u}\|^2 \Leftrightarrow H^\top H\mathbf{u}^* = H^\top \mathbf{y}. \quad (3.6)$$

Proof. (WIP) □

観測行列 H がフルランクでない場合は正規方程式 (3.6) の解は一意でないことに注意が必要である.

補題 3.4 (Tikhonov 正則化).

Proof. (WIP) □

より一般に状態空間, 観測空間での内積を取り換えることで, 重み付きの最小二乗法を考えることができる.

補題 3.5 (重みつき Tikhonov 正則化).

Proof. (WIP)

□

第 4 章

確率・統計とベイズ逆問題

4.1 本章の要点

3 章で扱った観測値から状態を逆推定する問題を確率的に定式化するために必要な確率論の基礎を導入する。推定対象が決定論的であっても、「わからなさ」を表すために確率分布を使って状態推定を表す。

4.2 確率論

学部生までの確率論の知識は前提とする。必要に応じて確率論の教科書を参照せよ。標本空間 Ω , σ -代数 \mathcal{F} , および確率測度 \mathbb{P} に対し, $(\Omega, \mathcal{F}, \mathbb{P})$ を完備確率空間とする。確率変数 $U : \Omega \rightarrow \mathbb{R}^l$ は \mathbb{R}^l 上の確率測度 μ_U を以下のように誘導する。

$$\mu_U(A) = \mathbb{P}(U^{-1}(A)). \quad (4.1)$$

この μ_U を U の分布と呼ぶ。ある関数 $p_U : \mathbb{R}^l \rightarrow [0, \infty)$ が存在して、任意のボレル集合 $A \subset \mathbb{R}^l$ に対し

$$\mu_U(A) = \int_A p_U(u) du, \quad (4.2)$$

が成り立つ時、 p_U を U の確率密度関数 (Probability Density Function: PDF) と呼ぶ。確率変数 U と確率測度 μ について、 $\mu_U = \mu$ が成り立つ時、 U は μ に従うといい、 $U \sim \mu$ と表記する。確率密度関数 p についても、同様の表記 $U \sim p$ を用いる。

定義 4.1 (\mathbb{R}^l 上のガウス分布)。ベクトル $\varpi \in \mathbb{R}^l$ および行列 $C \in \mathcal{L}_{sa}(\mathbb{R}^l)$ で $C \succ 0$ を満たすものに対し、平均 ϖ および共分散 C を持つガウス分布は、以下の確率密度関数を持つ確率分布である。

$$p(u) = \frac{1}{\sqrt{\det C (2\pi)^l}} \exp \left(-\frac{1}{2} |u - \varpi|_C^2 \right). \quad (4.3)$$

これを $\mathcal{N}(\varpi, C) \in \mathcal{M}_1(\mathbb{R}^l)$ と表記する。さらに、 \mathbb{R}^l -値確率変数 U がガウス確率変数であるとは、

ある $\varpi \in \mathbb{R}^l$ および $C \in \mathcal{L}_{sa}(\mathbb{R}^l)$ で $C \succ 0$ を満たすものが存在して、 $U \sim \mathcal{N}(\varpi, C)$ を満たすことをいう。

命題 4.1. ガウス確率変数のアフィン変換はガウス確率変数である。

命題 4.2. 互いに独立なガウス確率変数の線形結合はガウス確率変数である。

4.3 確率的線形逆問題

線形逆問題に対する最小二乗法を確率・統計的視点から正当化する。(WIP)

4.4 ベイズ逆問題

(WIP)

定理 4.1 (ベイズの公式)。

命題 4.3. ノイズあり線形観測定義 3.1 に対して、事前分布にがうす分布を仮定すると、事後分布もガウス分布となる。

4.5 グラフィカルモデル

確率変数間の依存関係を視覚的に示すために、同時確率分布の分解を表現する有向グラフィカルモデルを導入する。詳しくは [1] の 8 章などを参照せよ。(WIP)

第 5 章

微分方程式と力学系

5.1 本章の要点

データ同化における典型的な力学モデルとしての微分方程式や力学系の基礎を導入する.

5.2 常微分方程式

Gronwall の不等式や半群の理論について, [5] や https://kotatakeda.github.io/math/pdf/dynamical_model.pdf も参照せよ. (WIP)

第 6 章

常微分方程式の数値計算法

6.1 本章の要点

常微分方程式に基づいて状態の時間発展を計算するための数値計算法について紹介する.
(WIP)

第 III 部

データ同化問題

第 7 章

データ同化問題の定式化

7.1 本章の要点

データ同化は時系列観測からの状態推定を目的とし、フィルタリングの理論に基づいて定式化される [3]. ある一定の時間間隔ごとに観測値が得られる状況を考える. 対象の系が連続時間のモデル化をしていても, 観測時刻間の状態の時間発展を一つの写像として扱うことで, 離散時間状態空間モデルの定式化を考えることができる.

7.2 状態空間

7.2.1 状態空間モデル

有限次元の状態空間 $\mathcal{U} = \mathbb{R}^{N_u}$ 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 $\mathbf{u} : \mathbb{N} \times \Omega \rightarrow \mathbb{R}^{N_u}$ によってモデル化されるものとする:

$$\mathbf{u}_n = \Psi_n(\mathbf{u}_{n-1}) + \boldsymbol{\xi}_n, \quad n \in \mathbb{N}. \quad (7.1)$$

ここで, $\mathbf{u}_0 : \Omega \rightarrow \mathbb{R}^{N_u}$ は不確定な初期状態であり, $\Psi_n : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_u}$ は連続写像とする. モデルノイズ $\boldsymbol{\xi}_n$ はモデリングおよび近似誤差を表している. このモデルノイズのモデル化については議論の余地がある. 以下に典型的な例を示す.

確率的なノイズ 分布は確率密度関数 (PDF) $p_{\boldsymbol{\xi}_n}$ を持ち, 平均は $\mathbf{0}$ で, 共分散行列は $\mathbf{Q}_n \in \mathbb{R}^{N_u \times N_u}, \mathbf{Q}_n \succeq \mathbf{0}$ とする.

決定論的かつ完全モデル 対象が決定論的かつモデル化の誤差がないと仮定する場合, $\boldsymbol{\xi}_n = \mathbf{0}$ とする. 確率的なモデル化と整合的になるように, $p_{\boldsymbol{\xi}_n}$ の代わりにディラックのデルタ関数 $\delta_0(\boldsymbol{\xi}_n)$ を用いる.

有界なノイズ $\epsilon > 0$ に対し, $|\boldsymbol{\xi}_n| \leq \epsilon$ などとする.

推定したい時系列の時間発展をモデル化した (7.1) のことを, 離散時間ダイナミカルモデル (動的モデル) やシステムモデルと呼ぶ.

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

$$\mathbf{y}_n = h_n(\mathbf{u}_n) + \boldsymbol{\eta}_n, \quad n \in \mathbb{N}. \quad (7.2)$$

ここで、 $h_n : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{N_y}$ は連続な観測関数であり、 $(\boldsymbol{\eta}_n)_{n \in \mathbb{N}} \subset \mathbb{R}^{N_y}$ は PDF $p_{\boldsymbol{\eta}_n} : \mathbb{R}^{N_y} \rightarrow [0, \infty)$ を持つ独立同分布のノイズ列とし、モデルノイズ $\boldsymbol{\xi}_n$ とも独立であるとする。推定したい状態と得られている観測値の関係を表した (7.2) のことを、観測モデルと呼ぶ。

7.2.2 例

次の例では、常微分方程式の時間離散化及び観測時間間隔との対応づけを観る。以下は、[8] や [7] で取り上げられていた線形常微分方程式の例である。

例 7.1 (単振動とその観測). 角振動数 $\omega > 0$ に対して、線形常微分方程式

$$\frac{d^2 x}{dt^2} = -\omega^2 x \quad (7.3)$$

を考える。 $x_0 \in \mathbb{R}$ とする。速度を $v = \frac{dx}{dt}$ とし、状態ベクトル $\mathbf{u} = (x, v)^\top \in \mathbb{R}^2$ に対する常微分方程式に書き換える。

$$\frac{d}{dt} \begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} v \\ -\omega^2 x \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix}. \quad (7.4)$$

次に、(7.4) に従う時間 Δt の時間発展を行列 $M_{\Delta t} \in \mathbb{R}^{2 \times 2}$ を用いて

$$\mathbf{u}(\Delta t; \mathbf{u}_0) \approx M_{\Delta t} \mathbf{u}_0 \quad (7.5)$$

のように近似することを考える。例えば、*Symplectic Euler* 法により離散化すると

$$M_{\Delta t} = \begin{pmatrix} 1 & \Delta t \\ -\omega^2 \Delta t & 1 - \omega^2 (\Delta t)^2 \end{pmatrix} \quad (7.6)$$

を得る [7]。Heun 法により時間離散化すると、

$$M_{\Delta t} = \begin{pmatrix} 1 - \frac{1}{2} \omega^2 (\Delta t)^2 & \Delta t \\ -\omega^2 \Delta t & 1 - \frac{1}{2} \omega^2 (\Delta t)^2 \end{pmatrix} \quad (7.7)$$

を得る [8]。観測は $n_{\text{interval}} \in \mathbb{N}$ を用いて、 $\tau = n_{\text{interval}} \Delta t$ の時間が経つごとに得られると仮定する。離散的な観測時刻 $t_n = \tau n$ での状態を $\mathbf{u}_n = (x_n, v_n)^\top \in \mathbb{R}^2$ 、離散化に伴う誤差をモデルノイズとして $\boldsymbol{\xi}_n \in \mathbb{R}^2$ と書くと、

$$\mathbf{u}_n = M_{\Delta t}^{n_{\text{interval}}} \mathbf{u}_{n-1} + \boldsymbol{\xi}_n, \quad n \in \mathbb{N} \quad (7.8)$$

という離散時刻の間の時間発展を近似計算する線形ダイナミカルモデルを得る。これで、2次元の状態空間上で離散時間モデルが設定された。ここで、モデルノイズ $\boldsymbol{\xi}_n$ は離散化誤差の蓄積を表すと

考えられるが、具体的な与え方は自明でなく、体系的な理論が存在しない。例えば、 Δt と n_{interval} に依存する共分散行列 Q をもつガウスノイズとして確率的に表すことができる。

$$\xi_n \sim \mathcal{N}(\mathbf{0}, Q). \quad (7.9)$$

もしくは、それらに依存して上限 ϵ を定めることもできる。

次に、観測関数を行列 $H = [1, 0] \in \mathbb{R}^{1 \times 2}$ を用いて、

$$h_n(\mathbf{u}) = H\mathbf{u} = x \quad (7.10)$$

と取ると、 x 成分への射影となる。1次元観測空間上で観測ノイズは $r > 0$ に対して r^2 を分散にもつガウス分布に従う、つまり、

$$\eta_n \sim \mathcal{N}(0, r^2) \quad (7.11)$$

とすると、

$$y_n = H\mathbf{u}_n + \eta_n, \quad n \in \mathbb{N} \quad (7.12)$$

という観測モデルを得る。

7.2.3 有向グラフィカルモデルと条件付き独立性

状態空間モデル (7.1) および (7.2) に基づいて推定を行う上で条件付き独立性が有効に働く。この顕著な性質を視覚的に理解するために、節 4.5 で紹介したグラフィカルモデルを利用する。この状態空間モデルは、有向グラフィカルモデルを用いて図 7.1 のように表せる。二つの重要な条件付き独立性

$$p(\mathbf{u}_{n+1} \mid \mathbf{u}_{1:n}, \mathbf{y}_{1:n}) = p(\mathbf{u}_{n+1} \mid \mathbf{u}_n), \quad (7.13)$$

$$p(\mathbf{y}_n \mid \mathbf{u}_{1:n}, \mathbf{y}_{1:n-1}) = p(\mathbf{y}_n \mid \mathbf{u}_n) \quad (7.14)$$

が直ちにわかる。(7.13) はダイナミカルモデルのマルコフ性を表しており、 \mathbf{u}_{n+1} への過去時刻からの影響は1時刻前の状態 \mathbf{u}_n からの影響に集約される。また、(7.14) は観測の直接の依存範囲を示しており、現在の状態 \mathbf{u}_n を固定すると、 \mathbf{y}_n は他の時刻の状態や観測とは独立に定まる。

7.3 バイズデータ同化問題

状態空間モデル (7.1) と (7.2) に基づく状態推定問題をバイズの的に定式化する。 $N \in \mathbb{N}$ に対して、離散時間区間 $0 \leq n \leq N$ での観測 (の実現値) を $\mathbf{y}_{1:N} = \{\mathbf{y}_n \mid 0 \leq n \leq N\}$ と表記する。以下、簡単のため確率密度関数 (PDF) を用いて議論する。

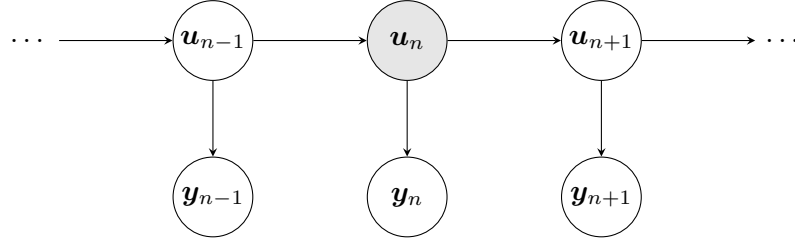


図 7.1 状態空間モデルの有向グラフィカルモデル. 丸は確率変数を表し、矢印は依存関係を表す. 条件付ける現在状態 u_n を灰色で示す.

定義 7.1 (データ同化問題). 推定時刻 $n \in \mathbb{N}$ に対して, 時刻 $N \in \mathbb{N}$ までの観測 $y_{1:N}$ を用いた u_n の推定値 v_n として, 以下のように条件付き確率分布に一致する分布を持つ確率変数を考える:

$$p_{v_n}(\cdot) = p_{u_n}(\cdot \mid y_{1:N}). \quad (7.15)$$

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, 推定時刻 n と (最新の) 観測時刻 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 (*forecast or 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*).

これらの分布は定理 4.1 のベイズ公式を用いて事後分布として得られる. 予測分布 $p_{u_n}(\cdot \mid y_{1:N})$ ($n > N$) 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.

7.3.1 フィルタリング問題

フィルタリング問題では, 観測が得られるごとに分布 $p_{v_n}(\cdot) = p_{u_n}(\cdot \mid y_{1:n})$ を最初から計算するのではなく, 前の時刻の $p_{v_{n-1}}(\cdot) = p_{u_{n-1}}(\cdot \mid y_{1:n-1})$ の情報を利用し逐次更新することが有用である. 予測ステップ: $p_{u_{n-1}}(\cdot \mid y_{1:n-1}) \mapsto p_{u_n}(\cdot \mid y_{1:n-1})$ と解析ステップ: $p_{u_n}(\cdot \mid y_{1:n-1}) \mapsto p_{u_n}(\cdot \mid y_{1:n})$ を交互に行うことで, フィルタ分布を計算する. まず, 予測ステップでは

$$p_{u_n}(\cdot \mid y_{1:n-1}) = \int_{\mathbb{R}^{N_u}} p_{u_n}(\cdot \mid y_{1:n-1}, u_{n-1}) p_{u_{n-1}}(u_{n-1} \mid y_{1:n-1}) du_{n-1} \quad (7.16)$$

$$= \int_{\mathbb{R}^{N_u}} p_{u_n}(\cdot \mid u_{n-1}) p_{u_{n-1}}(u_{n-1} \mid y_{1:n-1}) du_{n-1}. \quad (7.17)$$

と伝播する．ここで， $p_{\mathbf{u}_n}(\cdot \mid \mathbf{y}_{1:n-1}, \mathbf{u}_{n-1}) = p_{\mathbf{u}_n}(\cdot \mid \mathbf{u}_{n-1})$ を用いたが，これは (7.1) のマルコフ性から従う．次に，解析ステップでは，

$$p_{\mathbf{u}_n}(\mathbf{u}_n \mid \mathbf{y}_{1:n}) = p_{\mathbf{u}_n}(\mathbf{u}_n \mid \mathbf{y}_{1:n-1}, \mathbf{y}_n) \quad (7.18)$$

$$= \frac{p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{u}_n, \mathbf{y}_{1:n-1}) p_{\mathbf{u}_n}(\mathbf{u}_n \mid \mathbf{y}_{1:n-1})}{p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{y}_{1:n-1})} \quad (7.19)$$

$$= \frac{p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{u}_n) p_{\mathbf{u}_n}(\mathbf{u}_n \mid \mathbf{y}_{1:n-1})}{p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{y}_{1:n-1})}. \quad (7.20)$$

ここで， $p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{u}_n, \mathbf{y}_{1:n-1}) = p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{u}_n)$ を用いたが，これは観測の条件付き独立性から従う．予測値に対応する予備変数として $\hat{\mathbf{v}}_n$ を導入し，その分布を $p_{\hat{\mathbf{v}}_n}$ と表記すると，以上の2つのステップは次のようにまとめられる．

定義 7.2 (逐次データ同化). \mathbf{u}_n および \mathbf{y}_n がそれぞれ (7.1) および (7.2) に従うとする．逐次データ同化 (*sequential data assimilation*) とは，時刻 $n \in \mathbb{N}$ におけるフィルタ分布 $p_{\mathbf{v}_n}(\cdot) = p_{\mathbf{u}_n}(\cdot \mid \mathbf{y}_{1:n})$ を，次の2つのステップで計算することを言う．

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

$$p_{\hat{\mathbf{v}}_n}(\mathbf{v}) = \int_{\mathbb{R}^{N_u}} p_{\xi}(\mathbf{v} - \Psi(\mathbf{v}')) p_{\mathbf{v}_{n-1}}(\mathbf{v}') d\mathbf{v}'. \quad (7.21)$$

(II) (解析: $p_{\hat{\mathbf{v}}_n}, \mathbf{y}_n \rightarrow p_{\mathbf{v}_n}$) バイズの公式を用いて $p_{\hat{\mathbf{v}}_n}$ を $p_{\mathbf{v}_n}$ に更新する：

$$p_{\mathbf{v}_n}(\mathbf{v}) = \frac{p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{v}) p_{\hat{\mathbf{v}}_n}(\mathbf{v})}{\int_{\mathbb{R}^{N_u}} p_{\mathbf{y}}(\mathbf{y}_n \mid \mathbf{v}') p_{\hat{\mathbf{v}}_n}(\mathbf{v}') d\mathbf{v}'}, \quad (7.22)$$

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

(I) のステップは予測 (forecast または prediction) ステップと呼ばれ， $p_{\hat{\mathbf{v}}_n}(\mathbf{v})$ は予測分布 (forecast または prediction distribution) と呼ばれる．同様に，(II) のステップは解析 (analysis または update) ステップと呼ばれ， $p_{\mathbf{v}_n}(\mathbf{v})$ 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．

実際に，逐次データ同化 (定義 7.2) からフィルタ分布が得られることが次の定理で示される．

定理 7.1. 初期状態 \mathbf{u}_0 の不確かさを PDF $p_{\mathbf{u}_0}$ で表す．定義 7.2 を $p_{\mathbf{v}_0} = p_{\mathbf{u}_0}$ から始めると，任意の $n \in \mathbb{N}$ に対して厳密なフィルタ分布 $p_{\mathbf{v}_n}(\cdot) = p_{\mathbf{u}_n}(\cdot \mid \mathbf{y}_{1:n})$ を得る．

Proof. 上で見たように帰納的に従う． □

注意 7.1. 逐次データ同化は，モデル (7.1) のマルコフ性や観測ノイズの独立性に由来する同時分布の分解を利用している．

第 8 章

連続時間フィルタリング問題

8.1 本章の要点

確率微分方程式と連続時間観測を用いたフィルタリング問題の定式化を紹介する．一般的な有限次元系を想定した定式化が基本である．線形ガウス系に対しては，最適なフィルタとして連続時間のカルマンフィルタが得られる．一般の系に対しては，非線形フィルタリング問題を解くための確率微分方程式が得られるが，解析的には解けないため近似アルゴリズムを用いる．離散時間の定式化と比べて解析的に取り扱いやすい．

第 9 章

連続データ同化

9.1 本章の要点

偏微分方程式などの発展方程式に対して，ノイズのない連続時間観測を用いたデータ同化問題の定式化を紹介する．Navier-Stokes 方程式に代表されるように，「散逸構造」を持ち，有限次元のコンパクトアトラクタを持つ系を想定した定式化が基本である．そのような系の自由度の有限次元性に由来する有限次元観測による軌道の追跡可能性を議論する．

第Ⅳ部

データ同化アルゴリズム

第 10 章

フィルタリングアルゴリズム

10.1 本章の要点

本章では、フィルタリング問題に対するアルゴリズムについて紹介する。まず理論的に取り扱いやすい線形ガウス系という状態空間モデルの特別な場合を考える。この場合、状態推定分布がガウス分布に閉じ、「最適な」フィルタリングアルゴリズムとしてカルマンフィルタが得られる。線形ガウス系以外にも適用できるように、カルマンフィルタを拡張したアルゴリズムがいくつか提案されているがいずれも最適ではなく、フィルタリング分布を近似できない。一方で、モンテカルロ法に基づく粒子フィルタは、粒子数無限大でフィルタリング分布を再現できるが計算コストが高い。さらに、これらのアルゴリズムに対して限られた計算リソースや観測値からより良い推定を行うための改良技術についても紹介する。

10.2 カルマンフィルタとその拡張

10.2.1 カルマンフィルタと特徴づけ

例 7.1 の状態空間モデルは一般に以下のように書ける。状態空間と観測空間の次元を $N_u, N_y \in \mathbb{N}$ とし、 $(F_n)_{n \in \mathbb{N}} \subset \mathbb{R}^{N_u \times N_u}$ および $(H_n)_{n \in \mathbb{N}} \subset \mathbb{R}^{N_y \times N_u}$ として、

$$\mathbf{u}_n = F_n \mathbf{u}_{n-1} + \boldsymbol{\xi}_n, \quad \mathbf{y}_n = H_n \mathbf{u}_n + \boldsymbol{\eta}_n. \quad (10.1)$$

ここで、モデルノイズ $\boldsymbol{\xi}_n \sim \mathcal{N}(0, Q_n)$ は共分散行列 $Q_n \succeq 0$ を持つガウスノイズ、観測ノイズ $\boldsymbol{\eta}_n \sim \mathcal{N}(0, R_n)$ は共分散行列 $R_n \succ 0$ を持つガウスノイズとする。線形ガウス系 (10.1) は (7.1) の特殊な場合に対応する。初期状態の不確実性は $\mathbf{u}_0 \sim \mathcal{N}(\boldsymbol{\varpi}_0, P_0)$ に従うと仮定し、 $\boldsymbol{\varpi}_0 \in \mathbb{R}^{N_u}, P_0 \in \mathbb{R}^{N_u \times N_u}$ 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 [4]、線形ガウス系 (10.1) におけるフィルタ分布 (定義 7.2) を厳密かつ明示的に表し、逐次的な計算を可能にする。逐次データ同化 (定義 7.2) の各ステップで分布のガウス性が保たれるため、フィルタ分布は平均と共分散行列によって完全に特徴付けられる。実際、予測ステップでは命題 4.1 と命題 4.2 により解

析分布がガウス分布ならば予測分布もガウス分布となることがわかる．さらに，解析ステップでは命題 4.3 により事後分布（解析分布）もガウス分布となることがわかる．次のようにまとめられる．

定義 10.1 (カルマンフィルタ)．時刻 $n-1$ でのフィルタ分布 $\mathcal{N}(\varpi_{n-1}, P_{n-1})$ が与えられているとする．カルマンフィルタのアルゴリズムは以下の通り．

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

$$\widehat{\varpi}_n = F_n \varpi_{n-1}, \quad (10.2)$$

$$\widehat{P}_n = F_n P_{n-1} F_n^\top + Q_n. \quad (10.3)$$

(II) (解析: $\widehat{\varpi}_n, \widehat{P}_n, \mathbf{y}_n \rightarrow \varpi_n, P_n$) 事後分布の平均と共分散を計算する：

$$\varpi_n = \widehat{\varpi}_n + K_n(\mathbf{y}_n - H_n \widehat{\varpi}_n), \quad (10.4)$$

$$P_n = (I_{N_u} - K_n H_n) \widehat{P}_n. \quad (10.5)$$

ここで， K_n はカルマンゲインを表し，以下で与えられる．

$$K_n = \widehat{P}_n H_n^\top (H_n \widehat{P}_n H_n^\top + R_n)^{-1}. \quad (10.6)$$

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

真の状態 \mathbf{u}_n に対するカルマンフィルタによる推定値として，フィルタ分布を代表する平均 ϖ_n を用いることができる．このとき，共分散 P_n は推定の不確実性を表す．

例 10.1 (単振動 +KF)．例 7.1 を考えると， $N_u = 2, N_y = 1$ である． $q_1, q_2, r > 0$ に対して，

$$F_n = M_{\Delta t}^{n_{\text{interval}}}, \quad H_n = H, \quad Q_n = \text{diag}(q_1^2, q_2^2), \quad R_n = r^2 I_2$$

とおき，フィルタ分布の初期値 $\varpi_0 = (0, 0)^\top$ ， $P_0 = 4I_2$ を与えると，カルマンフィルタが適用できる．真の状態として，初期値 $(x_0, v_0) = (2.0, 0.0)$ に対する (7.3) の解析解を用いる．Heun 法による遷移行列 (7.7)， $q_1 = 0.01$ ， $q_2 = \sqrt{0.1}$ ， $r = 1.0$ として， $\Delta t = 0.1$ ， $n_{\text{interval}} = 2$ ($\tau = 0.2$) の場合に時刻 $T = 20$ までカルマンフィルタを実行する．図 10.1 は真の状態，観測値及び推定値の時系列である．

注意 10.1 (カルマンフィルタの制約)．カルマンフィルタは理論的に明快であり，フィルタ分布を厳密に計算する．しかし，高次元または複雑なシステム（例えば気象モデル）に適用する際には以下の 2 つの重要な制約がある．

- (1) カルマンフィルタは Ψ が線形であることを仮定するが，気象モデルは通常非線形である．
- (2) 状態空間の次元が非常に大きくなる（例えば $N_u \sim 10^9$ ）場合がある．これにより，共分散行列は $10^9 \times 10^9$ サイズとなり，現在のコンピュータではメモリに格納するのは不可能である．

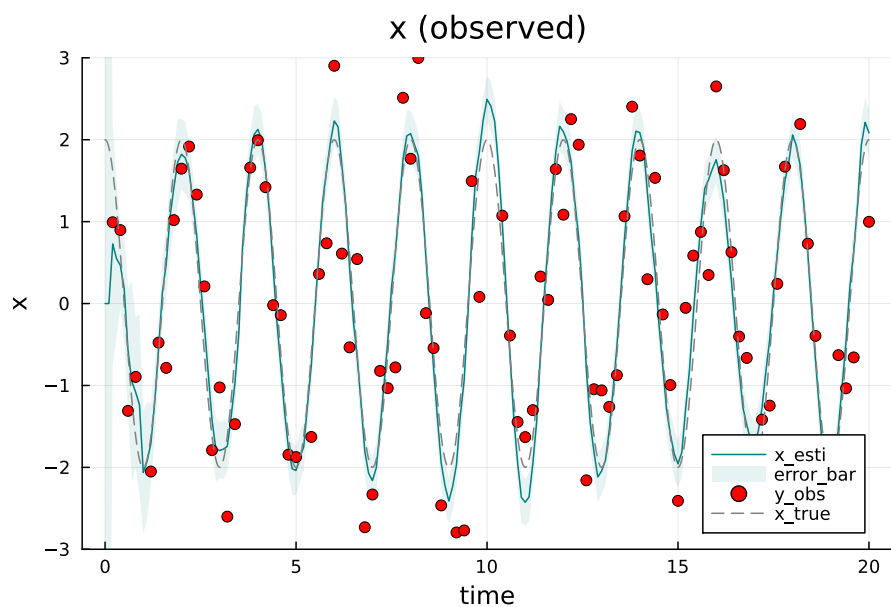


図 10.1 単振動に対するカルマンフィルタ. 観測されている変数である x の時系列 ('x_true'), 観測値 ('y_obs') 及びカルマンフィルタによる推定値 ('x_esti') がプロットされている. また, 標準偏差 2 倍分のエラーバー ('error_bar') を半透明色で表示している.

10.3 粒子フィルタ

10.4 改良技術

第 11 章

スモーディングアルゴリズム

11.1 本章の要点

解析時刻より未来の観測値を用いるスモーディング問題を解くためのアルゴリズムを紹介する．変分的な手法とベイズ的な手法の 2 つを扱う．

11.2 変分法

11.3 MCMC

第 V 部

数学解析

第 VI 部

応用例と実装

プログラムは https://github.com/KotaTakeda/da_math_intro に置いてある.

参考文献

- [1] C. M. BISHOP, *Pattern Recognition and Machine Learning*, Information Science and Statistics, Springer.
- [2] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, second edition, corrected reprint ed.
- [3] A. H. JAZWINSKI, *Stochastic Processes and Filtering Theory*, Dover publications, Inc., dover edition ed.
- [4] R. E. KALMAN, *A New Approach to Linear Filtering and Prediction Problems*, 82, pp. 35–45.
- [5] K. J. H. LAW, A. M. STUART, AND K. C. ZYGALAKIS, *Data Assimilation: A Mathematical Introduction*, Springer.
- [6] S. REICH AND C. COTTER, *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press.
- [7] T. SULLIVAN, *Introduction to Uncertainty Quantification*, vol. 63 of Texts in Applied Mathematics, Springer International Publishing.
- [8] 中野慎也, データ同化, 共立出版.