

金融時系列における カルマンフィルタ入門

佐久間吉行

November 8, 2025

自己紹介

- ・早稲田大学大学院理工学研究科 数学専攻 修士（理学）1990年
- ・アクチュアリー会正会員 1999年
- ・一橋大学大学院国際企業戦略科 博士課程修了 博士（経営） 2020年
- ・某私大 非常勤講師 経済数学担当 2020年～
- ・AIスタートアップ企業 Geotrans に転職予定 2025年11月～
 - 金融畠で銀行などに所属していました.
 - 2025年5月に還暦を迎えました.

ベイズの定理の復習

ベイズの定理（復習）

ベイズの定理

事象 A が起こる前提で事象 B が起こる確率 $P(B | A)$ は

$$P(B | A) = \frac{P(A \cap B)}{P(A)}.$$

結合確率は

$$P(A \cap B) = P(B | A)P(A) = P(A | B)P(B)$$

なので入れ替えても同様に成立する。

独立性と逆のベイズ則

- A と B が独立なら $P(B | A) = P(B)$, $P(A | B) = P(A)$.
- 逆向きのベイズ則 : $P(A | B) = \frac{P(B | A)P(A)}{P(B)}$, $P(B | A) = \frac{P(A | B)P(B)}{P(A)}$.

密度関数を用いた表現

条件付き確率と密度関数

確率変数 h と観測 y の結合密度を用いて条件付き確率を表現すると,

$$f(h | y) \equiv \frac{f(h, y)}{f(y)} \quad (f(h, y): \text{結合密度}, \ f(y): \text{周辺密度}).$$

ベイズ則の密度版

$$f(h | y) = \frac{f(y | h) f(h)}{\int_{-\infty}^{\infty} f(y | h) f(h) dh} = \frac{f(y | h) f(h)}{f(y)}.$$

分母 $f(y) = \int_{-\infty}^{\infty} f(y | h) f(h) dh$ は周辺密度と呼び計算可能である。今後は

$$f(h | y) = \frac{f(y | h) f(h)}{f(y)}$$

を多用する。

モンティホール問題とベイズ解法①

問題設定

- 3つの扉のうち1つに当たり、残り2つははずれ
- あなたが最初に1つ（例えば扉A: Fix）を選ぶ
- ホストは残りのはずれの扉（例えば扉B: Fix）を開ける
- ホストははずれを開けた後、あなたに残る扉（扉C）への変更を提案
- 扉を変えるべきか？

ベイズによる解法

$$P(C \text{ 当たり} \mid B \text{ を開く}) = \frac{P(B \text{ を開く} \mid C \text{ 当たり}) P(C \text{ 当たり})}{P(B \mid A) P(A) + P(B \mid C) P(C)} = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{2}{3}.$$

- 事前確率：どの扉も $\frac{1}{3} = P(A) = P(B) = P(C)$ で当たる
- 尤度：ホストは当たり扉以外を必ず開く
- 結果：変更すると勝率は $\frac{2}{3}$

モンティホール問題とベイズ解法②

考える手順1

$$P(C \text{ 当たり} | B \text{ を開く}) = \frac{P(B \text{ を開く} | C \text{ 当たり}) P(C \text{ 当たり})}{Q(B \text{ を開く})}$$

注: $P(C \text{ 当たり})$ と $Q(B \text{ を開く})$ は異なる確率なので記号を変えます。

考える手順2

$$P(B \text{ を開く}, A \text{ 当たり}) = P(B \text{ を開く} | A \text{ 当たり}) P(A \text{ 当たり}) = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$$

$$P(B \text{ を開く}, B \text{ 当たり}) = P(B \text{ を開く} | B \text{ 当たり}) P(B \text{ 当たり}) = 0 \times \frac{1}{3} = 0$$

$$P(B \text{ を開く}, C \text{ 当たり}) = P(B \text{ を開く} | C \text{ 当たり}) P(C \text{ 当たり}) = 1 \times \frac{1}{3} = \frac{1}{3}$$

考える手順3

$$Q(B \text{ を開く}) = P(B \text{ を開く}, A \text{ 当たり}) + P(B \text{ を開く}, B \text{ 当たり}) + P(B \text{ を開く}, C \text{ 当たり}) = \frac{1}{2}$$

おまけ：ギャラリー



阿良々木暦



羽川翼



老倉育

- 最近のラノベでは理学部数学科や統計科学部に入学する主人公が多い。ex 阿良々木暦（化物語）、梓川咲太（青ブタ）
- 何でもは知らないわよ。知ってることだけ（羽川翼）・・・←結構深い。分かることと分からることの線引きは実は難しい。何が分からぬいかを明確にできるといろいろ見えてくる。

4 扇版モンティホール問題とベイズ解法

問題設定

- 扇は4つあり、そのうち1つだけに当たり（車）、残り3つははずれ（ヤギ）。
- あなたは最初に扇 A を選ぶ。
- ホストは必ずはずれ扇を2つ（ここでは B と D）を開ける。
- 残った扇 C に当たりがある確率はいくらか？

ベイズによる解法

$$P(C \text{ 当たり} \mid \{B, D\} \text{ を開く}) = \frac{P(\{B, D\} \mid C) P(C)}{\sum_{X \in \{A, B, C, D\}} P(\{B, D\} \mid X) P(X)}.$$

初期確率は全扇均等： $P(X) = \frac{1}{4}$ 。ホストは当たり扇や選択扇は開けず、はずれを2つ選ぶので

$$P(\{B, D\} \mid A) = \frac{1}{\binom{3}{2}} = \frac{1}{3}, \quad P(\{B, D\} \mid C) = 1, \quad P(\{B, D\} \mid B) = P(\{B, D\} \mid D) = 0.$$

よって

$$P(C \mid \{B, D\}) = \frac{1 \cdot \frac{1}{4}}{\frac{1}{3} \cdot \frac{1}{4} + 1 \cdot \frac{1}{4}} = \frac{1}{1 + \frac{1}{3}} = \frac{3}{4}.$$

4 扇版モンティホール問題とベイズ解法③

考える手順1

$$P(C \text{ 当たり} \mid \{B,D\} \text{ を開く}) = \frac{P(\{B,D\} \text{ を開く} \mid C \text{ 当たり}) P(C \text{ 当たり})}{Q(\{B,D\} \text{ を開く})}$$

考える手順2

$$P(\{B,D\} \text{ を開く}, A \text{ 当たり}) = P(\{B,D\} \text{ を開く} \mid A \text{ 当たり}) P(A \text{ 当たり}) = \frac{1}{\binom{3}{2}} \times \frac{1}{4} = \frac{1}{12}$$

$$P(\{B,D\} \text{ を開く}, B \text{ 当たり}) = P(\{B,D\} \text{ を開く} \mid B \text{ 当たり}) P(B \text{ 当たり}) = 0 \times \frac{1}{4} = 0$$

$$P(\{B,D\} \text{ を開く}, C \text{ 当たり}) = P(\{B,D\} \text{ を開く} \mid C \text{ 当たり}) P(C \text{ 当たり}) = 1 \times \frac{1}{4} = \frac{1}{4}$$

$$P(\{B,D\} \text{ を開く}, D \text{ 当たり}) = P(\{B,D\} \text{ を開く} \mid D \text{ 当たり}) P(D \text{ 当たり}) = 0 \times \frac{1}{4} = 0$$

考える手順3

$$Q(\{B,D\}) = \sum_{X \in \{A,B,C,D\}} P(\{B,D\}, X) = \sum_{X \in \{A,B,C,D\}} P(\{B,D\} \mid X) P(X) = \frac{1}{3}$$

状態空間モデル

状態空間モデル（カルマン・フィルタ）とは何か

線形ガウス状態空間モデル（カルマン・フィルタ）

- 観測モデル

$$y_t = Z_t h_t + H_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I)$$

- 状態モデル

$$h_{t+1} = T_t h_t + R_t^{1/2} \eta_t, \quad \eta_t \sim \mathcal{N}(0, I)$$

- 観測データは必ずノイズを含む。カルマン・フィルタはこの「誤差に埋もれた観測」から、背後にある真の状態 h_t を逐次的に推定する仕組み
- 「予測」と「観測」をバランスよく融合することで、より正確で安定した推定値を得られる
- 数学的に閉じた形で計算できるため、予測・更新・平滑化の各ステップを効率的に実行可能
- アルゴリズムは再帰的で高速。時刻依存のモデル設定も可能なので、構造変化を含む非定常データにも柔軟に対応できる
- GPS 測位、株価・天候予測、信号処理など幅広い応用分野を持つのが大きな強み

活用事例①：経済指標の波形分解 (1/4)

観測式と状態式（トレンド=ローカルレベル）

$$\begin{aligned}y_t &= \mu_t + \gamma_t + \varepsilon_t, & \varepsilon_t &\sim \mathcal{N}(0, \sigma_\varepsilon^2) \\ \mu_{t+1} &= \mu_t + \eta_t, & \eta_t &\sim \mathcal{N}(0, \sigma_\eta^2) \quad (\text{レベルのランダム歩行}) \\ \gamma_{t+1} &= -\sum_{j=0}^{s-2} \gamma_{t-j} + \omega_t, & \omega_t &\sim \mathcal{N}(0, \sigma_\omega^2) \quad (\text{季節：周期 } s, \text{ 和制約})\end{aligned}$$

カルマン平滑化による分解（図の4パネルに対応）

Original : y_t

Trend : $\hat{\mu}_t = \mathbb{E}[\mu_t | y_{1:T}]$

Seasonal : $\hat{\gamma}_t = \mathbb{E}[\gamma_t | y_{1:T}]$

Residual : $\hat{\varepsilon}_t = y_t - \hat{\mu}_t - \hat{\gamma}_t$

活用事例①：経済指標の波形分解 (2/4)

状態と基本式 (ローカルレベル+季節, 周期 s)

観測式

$$y_t = Z \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$$

状態方程式

$$\alpha_{t+1} = T \alpha_t + R \eta_t, \quad \eta_t = \begin{pmatrix} \eta_t^{(\mu)} \\ \omega_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, Q\right)$$

状態ベクトル (横ベクトル表記で省スペース)

$$\alpha_t = (\mu_t, \gamma_t, \gamma_{t-1}, \dots, \gamma_{t-s+2})^\top \in \mathbb{R}^s$$

μ_t : ローカルレベル (RW) / γ_t : 季節 ($\sum_{j=0}^{s-1} \gamma_{t-j} = 0$ を維持)

活用事例①：経済指標の波形分解 (3/4)

行列の定義（カルマン型の行列表記）

4

季節ブロック S

$$S = \begin{pmatrix} -1 & -1 & \cdots & -1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$(s-1) \times (s-1)$ で季節の和制約
 $\sum_{j=0}^{s-1} \gamma_{t-j} = 0$ を維持。

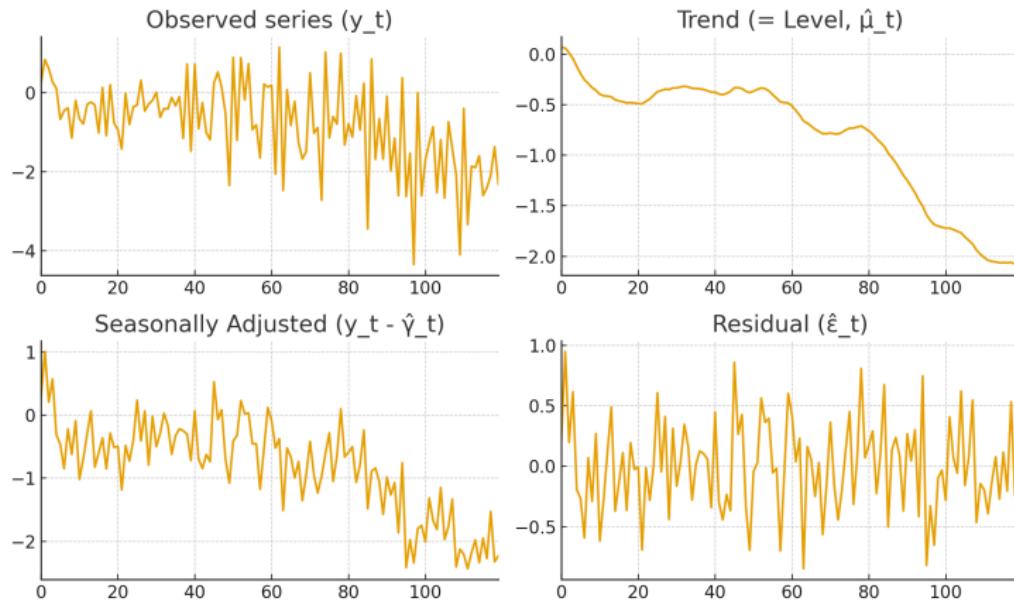
観測側

$$Z = (1 \ 1 \ 0 \ \cdots \ 0), \quad H = \sigma_\varepsilon^2$$

状態遷移とノイズ注入

$$T = \begin{pmatrix} 1 & \mathbf{0}_{1 \times (s-1)} \\ \mathbf{0}_{(s-1) \times 1} & S \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\omega^2 \end{pmatrix}$$

活用事例①：経済指標の波形分解 (4/4)



上段：観測値

中段：トレンド 下段：季節性 最下段：ノイズ

左上：Original (y_t) 右上：Trend ($\hat{\mu}_t$) 左下：Seasonal ($\hat{\gamma}_t$) 右下：Residual ($\hat{\varepsilon}_t$)

活用事例②：時変ベータの推定

モデル設定

- 観測変数：個別株超過収益率 $y_t = r_{i,t} - r_{f,t}$
- 説明変数：市場超過収益率 $x_t = r_{M,t} - r_{f,t}$ (観測値)
- 状態変数：時変ベータ (市場リスクの感応度) β_t (推定対象)

$$\begin{cases} y_t = \beta_t x_t + \varepsilon_t, & \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \\ \beta_{t+1} = \beta_t + \eta_t, & \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \end{cases}$$

カルマン・フィルタを用いて逐次推定することで、 β_t の時系列推定値とその不確実性をリアルタイムに取得可能.

活用事例③：ファクターエクスポージャーの動的推定_1

モデル設定

ポートフォリオ超過収益率 $r_t = r_{p,t} - r_{f,t}$ を、市場超過収益率 $(r_{M,t} - r_{f,t})$ 、サイズ因子 SMB_t 、スタイル因子 HML_t などの観測ファクター $F_t = [F_{1,t}, F_{2,t}, F_{3,t}]^\top$ と、動的なファクターエクスポージャー(感応度) θ_t によって説明：

$$\begin{cases} r_t = F_t^\top \theta_t + \varepsilon_t, & \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \\ \theta_{t+1} = \theta_t + \eta_t, & \eta_t \sim \mathcal{N}(0, \Sigma_\eta). \end{cases}$$

- カルマン・フィルタを用いて、時点 t ごとのエクスポージャー θ_t を推定
- エクスポージャーのリアルタイム変化を把握し、リスク管理・ポジション調整に活用

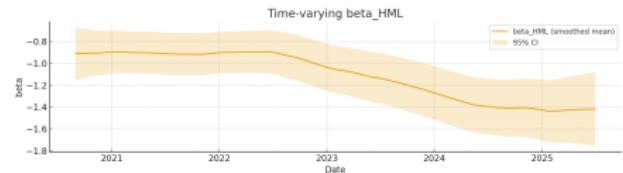
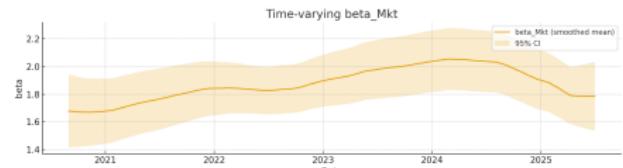
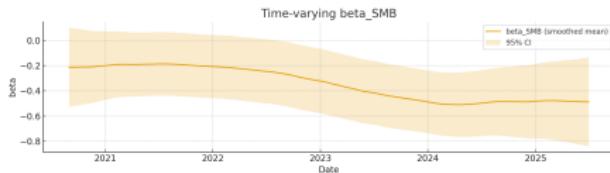
活用事例③：ファクターエクスポートの動的推定_2

- 市場因子 $r_{M,t} - r_{f,t}$ ：
 - 全銘柄の期待収益率の等価加重や時価総額加重で算出
- サイズ因子 SMB_t ：(小型株 vs 大型株)
 - SMB_t は小型株が強いとプラス、大型株が強いとマイナスに動く、 β マイナスなら大型株が強いときリターンが出ることを示唆
 - 毎年の会計年度末の各銘柄の時価総額 (株価 × 発行済株式数)
 - 上場全企業の時価総額を算出し 2 分割 (B,S にタグ付け),
 - $SMB_t = \frac{1}{3}((R_{S/H,t} + R_{S/M,t} + R_{S/L,t}) - (R_{B/H,t} + R_{B/M,t} + R_{B/L,t}))$
- スタイル因子 HML_t ：(バリュー株 vs グロース株)
 - HML_t はバリュー株が強いとプラス、グロース株が強いとマイナスに動く、 β マイナスならグロース株が強いときリターンが出ることを示唆
 - 每年の会計年度末の簿価と時価を用いて作成。翌年度の 7 月から 6 月までを作成。
 - 上場全企業の簿価/時価を算出し 3 分割 (H,M,L にタグ付け),
 - $HML_t = \frac{1}{2}((R_{S/H,t} + R_{B/H,t}) - (R_{S/L,t} + R_{B/L,t}))$
- 時変ベータを観測することで、特定の銘柄や投信の収益の源泉がみてとれる。

活用事例③：ファクターエクスポートの動的推定_3

因子	指標	正ポジション	正ポジション 上昇期	逆ポジション	逆ポジション 上昇期
市場因子	市場リターン	市場全体と相 関強い株	景気拡大期	ディフェンシ ブ銘柄・生活 関連銘柄	景気後退期
サイズ因子	時価総額	小型株	景気回復期	大型株	不況期
スタイル因子	簿価/時価	バリュー株 (割安株)	景気回復期	グロース株 (成長株)	景気拡大期／ 低金利・技術 革新期

活用事例③：NVDA 株価とファーマフレンチ 3 因子モデルの推定結果



- 市場ベータは一貫して高めで、期間後半にやや低下傾向。
- SMB はマイナスにシフト（大型優位の傾向）。
- HML はマイナス方向が強まり、NVDA の成長（グロース）性を反映。

線形ガウス状態空間モデル

対象モデル

$$\begin{cases} h_{t+1} = A_t h_t + Q_t^{1/2} \varepsilon_t, \\ y_t = C_t h_t + R_t^{1/2} \eta_t, \end{cases} \quad \varepsilon_t \sim \mathcal{N}(0, I), \quad \eta_t \sim \mathcal{N}(0, I).$$

$$W_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + R_t)^{-1},$$

$$\hat{h}_{t|t} = \hat{h}_{t|t-1} + W_t (y_t - C_t \hat{h}_{t|t-1}), \quad \hat{h}_{t+1|t} = A_t \hat{h}_{t|t},$$

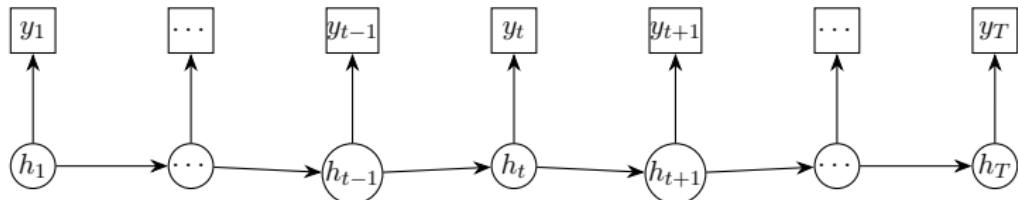
$$P_{t|t} = P_{t|t-1} - W_t C_t P_{t|t-1}, \quad P_{t+1|t} = A_t P_{t|t} A_t^T + Q_t.$$

これらは「状態の平均 \hat{x} と共に分散行列 P の更新式」ですが、式だけを見ると「何を」「なぜ」更新しているのか全くわかりません。

しかし、ベイズの定理から導出する「予測分布 \Leftrightarrow フィルタ分布」の相互の推移で考えると式の意味が見えてきます。

状態空間モデル：予測・フィルタリング・平滑化分布の定義

状態空間モデルの遷移図



- $y^t = \{y_1, y_2, \dots, y_t\}$ (時刻 1 から t までの観測系列)
- $y_{t+1}^T = \{y_{t+1}, \dots, y_T\}$ (時刻 $t + 1$ から T までの観測系列)
- $h^t = \{h_1, h_2, \dots, h_t\}$ (時刻 1 から t までの隠れ状態系列)

モデルの定義：線形ガウス状態空間モデル

観測方程式：

$$y_t = Z_t h_t + H_t^{1/2} \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1).$$

状態方程式：

$$h_{t+1} = T_t h_t + Q_t^{1/2} \eta_t, \quad \eta_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1).$$

ここで, ε_t と η_s は任意の t, s に対して互いに独立.

初期値：

$$h_1 \sim \mathcal{N}(a_1, P_1).$$

本モデルは「線形ガウス」かつ「マルコフ性」を仮定した状態空間モデルです.

- カルマンフィルタ：線形ガウスモデルに特化した最適推定手法
- 状態空間モデル：より一般には非線形／非ガウスにも拡張可能

マルコフ性による条件付き独立性

本モデルのマルコフ性から、次の同値性が成り立ちます。

$$y_t \mid h^t, y^{t-1} = y_t \mid h_t$$

観測方程式では、過去の状態や観測を与えられても、 y_t に影響を与えるのは当時刻の状態 h_t のみです。

$$h_{t+1} \mid h^t, y^t = h_{t+1} \mid h_t$$

状態方程式では、今までの状態・観測を与えられても、次時刻の状態 h_{t+1} に影響を与えるのは直前の状態 h_t のみです。

この性質により、カルマンフィルタの各ステップ（予測・更新・平滑化）が逐次的かつ効率的に実装できます。

観測データ y_t の条件付き分布

後に使うため、観測データ $y^{t-1} \equiv y_{1:t-1}$ が与えられたときの時刻 t での観測データ y_t の条件付き分布を求めておく：

$$f(y_t | y^{t-1}) = \int f(y_t | h_t) f(h_t | y^{t-1}) dh_t.$$

- 観測モデルの密度関数 $f(y_t | h_t)$ と 予測分布 $f(h_t | y^{t-1})$ を畳み込むことで得られる

観測データ y_t の条件付き分布の導出詳細

観測データ y_t の条件付き分布

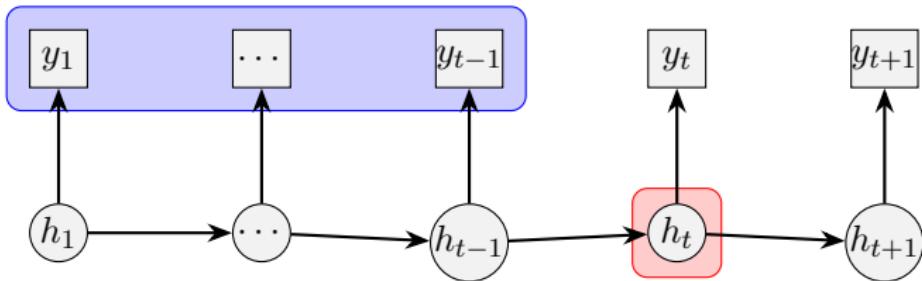
観測データ $y^{t-1} \equiv y_{1:t-1}$ が与えられたときの時刻 t での観測データ y_t の条件付き分布は

$$f(y_t | y^{t-1}) = \int f(y_t | h_t) f(h_t | y^{t-1}) dh_t.$$

$$\begin{aligned} f(y_t | y^{t-1}) &= \int f(y_t, h_t | y^{t-1}) dh_t \\ &= \int f(y_t | h_t, y^{t-1}) f(h_t | y^{t-1}) dh_t \\ &= \int f(y_t | h_t) f(h_t | y^{t-1}) dh_t. \end{aligned}$$

- 第1行：結合確率の積分による周辺化
- 第2行：条件付き確率の乗法定理
- 第3行：観測モデルのマルコフ性

状態空間モデルの遷移図（予測分布用ハイライト）



- 青く塗られた領域：既知の観測データ $\{y_1, \dots, y_{t-1}\}$
- 赤く塗られた領域：予測対象の隠れ状態 h_t
- 矢印：状態遷移モデル・観測モデル

予測分布 (Predictive Distribution)

観測データ $y^{t-1} \equiv y_{1:t-1}$ が得られたときの、時刻 t の状態 h_t の事前分布を「予測分布」と呼ぶ。

$$\underbrace{f(h_t | y_{1:t-1})}_{\text{予測分布}} = \int \underbrace{f(h_t | h_{t-1})}_{\text{状態遷移モデル}} \underbrace{f(h_{t-1} | y_{1:t-1})}_{\text{前期までのフィルタ分布}} dh_{t-1}.$$

予測分布の導出

予測分布

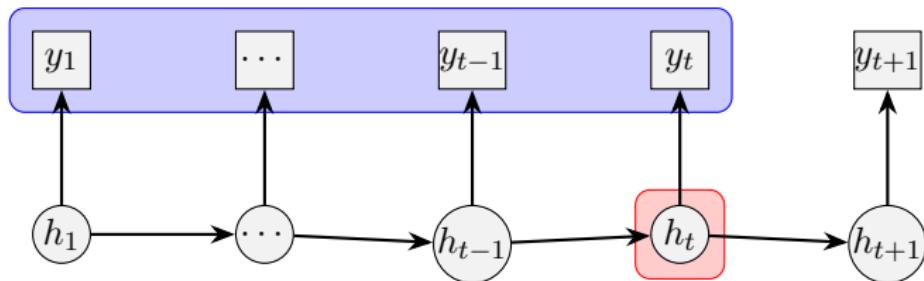
観測データ $y_{1:t-1}$ のもとでの時刻 t の状態 h_t の事前分布は

$$f(h_t | y_{1:t-1}) = \int f(h_t | h_{t-1}) f(h_{t-1} | y^{t-1}) dh_{t-1}.$$

$$\begin{aligned} f(h_t | y^{t-1}) &= \int f(h_t, h_{t-1} | y^{t-1}) dh_{t-1} \\ &= \int f(h_t | h_{t-1}, y^{t-1}) f(h_{t-1} | y^{t-1}) dh_{t-1} \\ &= \int f(h_t | h_{t-1}) f(h_{t-1} | y^{t-1}) dh_{t-1}. \end{aligned}$$

- 第1行：結合確率の積分による周辺化
- 第2行：条件付き確率の乗法定理
- 第3行：マルコフ性（状態遷移モデルのみ依存）

状態空間モデルの遷移図（フィルタリング分布用ハイライト）



- 青く塗られた領域：フィルタリングで用いる観測データ $\{y_1, \dots, y_t\}$
- 赤く塗られた領域：推定対象の隠れ状態 h_t
- 矢印 : 状態遷移モデル・観測モデルによる更新プロセス

フィルタリング分布 (Filtering Distribution)

観測データ $y^t \equiv y_{1:t}$ が得られたときの、時刻 t の状態 h_t の事後分布を「フィルタリング分布」と呼ぶ。

$$\underbrace{f(h_t | y_{1:t})}_{\text{フィルタ分布}} = \frac{\underbrace{f(y_t | h_t)}_{\substack{\text{観測モデル} \\ \text{予測分布}}} \underbrace{f(h_t | y_{1:t-1})}_{\substack{\text{証拠 (正規化定数)}}}}{\underbrace{f(y_t | y_{1:t-1})}_{\text{証拠 (正規化定数)}}}.$$

フィルタリング分布の導出

フィルタリング分布

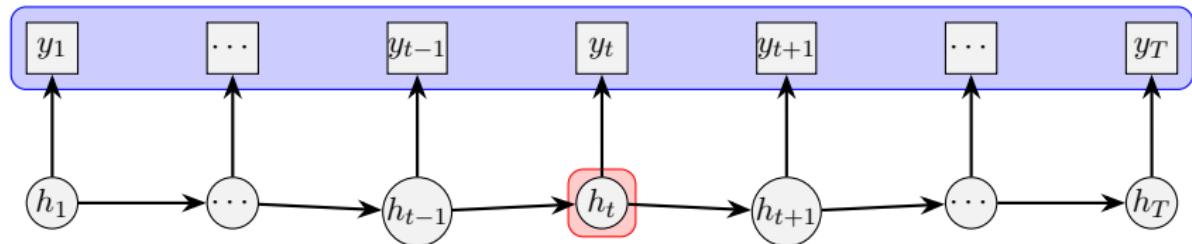
観測データ $y^t \equiv y_{1:t} = \{y_s\}_{s=1}^t$ のもとでの時刻 t の状態 h_t の事後分布は

$$f(h_t | y_{1:t}) = \frac{f(y_t | h_t) f(h_t | y_{1:t-1})}{f(y_t | y_{1:t-1})}.$$

$$\begin{aligned} f(h_t | y_{1:t}) &= f(h_t | y_t, y_{1:t-1}) = \frac{f(h_t, y_t, y_{1:t-1})}{f(y_t, y_{1:t-1})} = \frac{f(h_t, y_t | y_{1:t-1})}{f(y_t | y_{1:t-1})} \\ &= \frac{f(y_t | h_t, y_{1:t-1}) f(h_t | y_{1:t-1})}{f(y_t | y_{1:t-1})} \\ &= \frac{f(y_t | h_t) f(h_t | y_{1:t-1})}{f(y_t | y_{1:t-1})}. \end{aligned}$$

- 第3行：マルコフ性

状態空間モデルの遷移図（平滑化分布用ハイライト）



- 青く塗られた領域：平滑化分布で用いる観測データ $\{y_1, \dots, y_T\}$
- 赤く塗られた領域：推定対象の状態 h_t
- 矢印 : 状態遷移モデル・観測モデルによる更新プロセス

平滑化分布 (Smoothing Distribution)

観測データ全期間 $y_{1:T}$ が得られたときの、時刻 t の状態 h_t の事後分布を「平滑化分布」と呼ぶ。

$$\underbrace{f(h_t \mid y_{1:T})}_{\text{平滑化分布}} = \underbrace{f(h_t \mid y_{1:t})}_{\text{フィルタ分布}} \times \int \underbrace{f(h_{t+1} \mid h_t)}_{\text{状態遷移モデル}} \underbrace{\frac{f(h_{t+1} \mid y_{1:T})}{f(h_{t+1} \mid y_{1:t})}}_{\text{平滑化分布／予測分布}} dh_{t+1}.$$

平滑化分布の導出 ①

平滑化分布

観測データ全期間 $y_{1:T}$ のもとでの時刻 t の状態 h_t の事後分布は

$$f(h_t \mid y_{1:T}) = f(h_t \mid y_{1:t}) \int \frac{f(h_{t+1} \mid y_{1:T}) f(h_{t+1} \mid h_t)}{f(h_{t+1} \mid y_{1:t})} dh_{t+1}.$$

h_t は、 h_{t+1}, y^t が与えられると分布が決まる。残りの y_{t+1}^T は h_t 既知なので落とすことができる。これにより、以下が成り立つ：

$$\begin{aligned} f(h_t \mid h_{t+1}, y^T) &= f(h_t \mid h_{t+1}, y^t, y_{t+1}^T) = f(h_t \mid h_{t+1}, y^t) \\ &= \frac{f(h_t, h_{t+1} \mid y^t)}{f(h_{t+1} \mid y^t)} = \frac{f(h_{t+1} \mid h_t, y^t) f(h_t \mid y^t)}{f(h_{t+1} \mid y^t)} \\ &= \frac{f(h_{t+1} \mid h_t) f(h_t \mid y^t)}{f(h_{t+1} \mid y^t)}. \end{aligned}$$

平滑化分布の導出 ②

上式を用いて平滑化分布を求める：

$$\begin{aligned} f(h_t | y^T) &= \int f(h_t, h_{t+1} | y^T) dh_{t+1} \\ &= \int f(h_t | h_{t+1}, y^T) f(h_{t+1} | y^T) dh_{t+1} \\ &= \int \frac{f(h_{t+1} | h_t) f(h_t | y^t)}{f(h_{t+1} | y^t)} f(h_{t+1} | y^T) dh_{t+1} \\ &= f(h_t | y^t) \int \frac{f(h_{t+1} | y^T) f(h_{t+1} | h_t)}{f(h_{t+1} | y^t)} dh_{t+1}. \end{aligned}$$

以上により，平滑化分布 $f(h_t | y^T)$ の密度関数は，

- 一期前の平滑化分布 $f(h_{t+1} | y^T)$
- 状態方程式（状態遷移モデル） $f(h_{t+1} | h_t)$
- 事前に求めた予測分布 $f(h_{t+1} | y^t)$
- 事前に求めたフィルタリング分布 $f(h_t | y^t)$

の各密度関数から計算できる。

平滑化分布の導出 ③

後で使用するので、平滑化分布の導出の中で出てきた以下を記憶する。

$$f(h_t \mid h_{t+1}, y^t) = \frac{f(h_{t+1} \mid h_t) f(h_t \mid y^t)}{f(h_{t+1} \mid y^t)}.$$

分布の違いと使い分け

- 予測分布: 未来の状態を予測
- フィルタ分布: 現在の状態をその時点で推定
- 平滑化分布: 将来情報も使って過去の状態を再推定

状態空間モデルが線形ガウスモデル（カルマンフィルタ）の場合：

- モデルが、平均・分散で決まる正規分布で表現できる。
- 平均・分散が行列演算だけで更新できる。

線形ガウス状態空間モデルの導出とパラメータ推計 (ローカル・レベル・モデル)

ローカル・レベル・モデル：定義と分布

系列 $y^T = \{y_t\}_{t=1}^T$ が以下に従うとする。初期値 a_1, P_1 は既知とする。

モデル構造（観測方程式／状態方程式／初期値）

$$\begin{cases} y_t = h_t + \sigma \varepsilon_t, & \varepsilon_t \sim N(0, 1), \quad t = 1, \dots, T, \\ h_{t+1} = h_t + \tau \eta_t, & \eta_t \sim N(0, 1), \quad t = 1, \dots, T, \\ h_1 | y_0 \sim N(a_1, P_1). \end{cases}$$

求めたい分布：（導出はカルマン・フィルタの更新式に基づく）

予測分布 $h_t | y^{t-1} \sim N(\mu_{t|t-1}, \sigma_{t|t-1}^2) \equiv N(a_t, P_t), \quad t = 2, \dots, T,$

フィルタ分布 $h_t | y^t \sim N(\mu_{t|t}, \sigma_{t|t}^2), \quad t = 1, \dots, T,$

平滑化分布 $h_t | y^T \sim N(\mu_{t|T}, \sigma_{t|T}^2), \quad t = 1, \dots, T.$

ローカル・レベル・モデル（補足）

- ローカル・レベル・モデルは、時系列データ y_t を「隠れたレベル h_t 」とノイズで構成
- 時刻ごとに「レベル」がゆっくり変動する「ランダムウォーク」モデル
- 観測モデルと状態モデルだから、予測・フィルタ・平滑化分布を得られる
- カルマン・フィルタを用いれば、
 - 逐次的にレベル h_t の「時系列推定値」と「不確実性」を算出
 - 未知のパラメータ (σ^2, τ^2) は最尤推定などで推定可能
- 金融や経済データの「トレンド抽出」や「構造変化検出」によく使われる

ローカル・レベル・モデルの分布の導出結果

フィルタリング分布：

$$h_t \mid y_t \sim N(\mu_{t|t}, \sigma_{t|t}^2) = N\left(a_t + \frac{P_t}{P_t + \sigma^2}(y_t - a_t), P_t \left(1 - \frac{P_t}{P_t + \sigma^2}\right)\right)$$

予測分布：

$$h_{t+1} \mid y^t \sim N(a_{t+1}, P_{t+1}) = N\left(\mu_{t|t}, \tau^2 + \sigma_{t|t}^2\right)$$

平滑化分布：

$$\begin{aligned} h_t \mid y^T &\sim N(\mu_{t|T}, \sigma_{t|T}^2) \\ &= N\left(\mu_{t|t} + \frac{\sigma_{t|t}^2}{P_{t+1}}(\mu_{t+1|T} - a_{t+1}), \sigma_{t|t}^2 + \frac{\sigma_{t|t}^4}{P_{t+1}^2}(\sigma_{t+1|T}^2 - P_{t+1})\right) \end{aligned}$$

尤度関数とパラメータ推定

未知パラメータ σ, τ の推定：得られた結果をもとに，系列 $y^T = \{y_t\}_{t=1}^T$ の尤度関数 $f(y^T)$ を明示的に記述できる（チェーンルールとマルコフ性）：

$$f(y^T) = \prod_{t=1}^T f(y_t \mid y^{t-1}) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi(P_t + \sigma^2)}} \exp\left(-\frac{(y_t - a_t)^2}{2(P_t + \sigma^2)}\right) \quad (\text{ii})$$

更新式（時点 t におけるフィルタリングと予測の更新）：

$$\begin{aligned} a_{t+1} &= a_t + \frac{P_t}{P_t + \sigma^2}(y_t - a_t), \\ P_{t+1} &= \tau^2 + P_t \left(1 - \frac{P_t}{P_t + \sigma^2}\right) \end{aligned} \quad (\text{iii})$$

初期値 a_1, P_1 は与えられていると仮定する。

モデルのパラメータ σ, τ を推定したい場合は，(ii) の尤度関数が最大化されるように σ, τ を選べばよい。

フィルタリング分布①：定義と積分準備

ベイズの定理に基づき (以下はフィルタリング分布の計算過程で導出したことを思い出そう),

$$f(y_t | h_t) f(h_t | y^{t-1}) = f(h_t | y^t) f(y_t | y^{t-1}).$$

ここからフィルタリング分布 $f(h_t | y^t)$ を導出する.

まず, 以下の分布が仮定されている :

$$y_t | h_t \sim N(h_t, \sigma^2)$$

$$h_t | y^{t-1} \sim N(a_t, P_t),$$

したがって,

$$f(h_t | y^{t-1}) = \frac{1}{\sqrt{2\pi P_t}} \exp\left(-\frac{(h_t - a_t)^2}{2P_t}\right),$$

$$f(y_t | h_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - h_t)^2}{2\sigma^2}\right).$$

フィルタリング分布②：積分による導出と結果

上記は以下の形に変形できる：

$$\begin{aligned} f(y_t | h_t) f(h_t | y^{t-1}) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - h_t)^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi P_t}} \exp\left(-\frac{(h_t - a_t)^2}{2P_t}\right) \\ &= \sqrt{\frac{\sigma^2 + P_t}{2\pi \sigma^2 P_t}} \exp\left(-\frac{(\sigma^2 + P_t)\left(h_t - \frac{a_t \sigma^2 + y_t P_t}{\sigma^2 + P_t}\right)^2}{2\sigma^2 P_t}\right) \\ &\quad \times \frac{1}{\sqrt{2\pi(\sigma^2 + P_t)}} \exp\left(-\frac{(y_t - a_t)^2}{2(\sigma^2 + P_t)}\right). \end{aligned}$$

ここで, $f(y_t | y^{t-1})$ は以下で導出できる. 上の式をそのまま代入して h_t で積分すれば良い.

$$\begin{aligned} f(y_t | y^{t-1}) &= \int f(y_t | h_t) f(h_t | y^{t-1}) dh_t \\ &= \frac{1}{\sqrt{2\pi(\sigma^2 + P_t)}} \exp\left(-\frac{(y_t - a_t)^2}{2(\sigma^2 + P_t)}\right) \end{aligned}$$

フィルタリング分布③：積分による導出と結果（続き）

$f(y_t | h_t)$ は以下で得られるので、

$$f(h_t | y^t) = \frac{f(y_t | h_t) f(h_t | y^{t-1})}{f(y_t | y^{t-1})}.$$

上記の式を代入することで、以下が得られる。

$$f(h_t | y^t) = \sqrt{\frac{\sigma^2 + P_t}{2\pi \sigma^2 P_t}} \exp\left(-\frac{(\sigma^2 + P_t)\left(h_t - \frac{a_t \sigma^2 + y_t P_t}{\sigma^2 + P_t}\right)^2}{2 \sigma^2 P_t}\right).$$

フィルタリング分布は以下の定義であり、

$$h_t | y^t \sim N\left(\mu_{t|t}, \sigma_{t|t}^2\right)$$

フィルタリング分布の密度関数は以下の定義であった。

$$f(h_t | y^t) = \frac{1}{\sqrt{2\pi \sigma_{t|t}^2}} \exp\left(-\frac{(h_t - \mu_{t|t})^2}{2 \sigma_{t|t}^2}\right).$$

フィルタリング分布④：積分による導出と結果（続き）

以上より、

$$\begin{aligned}\mu_{t|t} &= \frac{a_t \sigma^2 + y_t P_t}{\sigma^2 + P_t} \\ &= a_t + \frac{P_t}{P_t + \sigma^2} (y_t - a_t) \\ \sigma_{t|t}^2 &= \frac{P_t \sigma^2}{P_t + \sigma^2}\end{aligned}$$

これは、カルマンゲイン

$$K_t = \frac{P_t}{P_t + \sigma^2}$$

を用いて表すと：

$$\mu_{t|t} = a_t + K_t (y_t - a_t), \quad \sigma_{t|t}^2 = (1 - K_t) P_t$$

(4) 予測分布①：定義と積分準備

ベイズの定理に基づき (以下は平滑化分布の計算過程で導出したことを思い出そう),

$$f(h_{t+1} | h_t) f(h_t | y^t) = f(h_t | h_{t+1}, y^t) f(h_{t+1} | y^t).$$

ここから予測分布 $f(h_{t+1} | y^t)$ を導出する.

まず, 以下の分布が仮定されている :

$$\begin{aligned} h_{t+1} | h_t &\sim N(h_t, \tau^2) \\ h_t | y^t &\sim N(\mu_{t|t}, \sigma_{t|t}^2), \end{aligned}$$

したがって,

$$\begin{aligned} f(h_t | y^t) &= \frac{1}{\sqrt{2\pi\sigma_{t|t}^2}} \exp\left(-\frac{(h_t - \mu_{t|t})^2}{2\sigma_{t|t}^2}\right), \\ f(h_{t+1} | h_t) &= \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(h_{t+1} - h_t)^2}{2\tau^2}\right). \end{aligned}$$

(4) 予測分布②：積分による導出と結果

$$\begin{aligned} f(h_{t+1} | h_t) f(h_t | y^t) &= \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(h_{t+1}-h_t)^2}{2\tau^2}\right) \frac{1}{\sqrt{2\pi\sigma_{t|t}^2}} \exp\left(-\frac{(h_t-\mu_{t|t})^2}{2\sigma_{t|t}^2}\right) \\ &= \sqrt{\frac{\tau^2 + \sigma_{t|t}^2}{2\pi\tau^2\sigma_{t|t}^2}} \exp\left(-\frac{(\tau^2 + \sigma_{t|t}^2)\left(h_t - \frac{\mu_{t|t}\tau^2 + h_{t+1}\sigma_{t|t}^2}{\tau^2 + \sigma_{t|t}^2}\right)^2}{2\tau^2\sigma_{t|t}^2}\right) \\ &\quad \times \frac{1}{\sqrt{2\pi(\tau^2 + \sigma_{t|t}^2)}} \exp\left(-\frac{(h_{t+1} - \mu_{t|t})^2}{2(\tau^2 + \sigma_{t|t}^2)}\right). \end{aligned}$$

ここで、予測分布 $h_{t+1} | y^t$ の密度関数は以下で導出できる。上の式をそのまま代入して h_t で積分すれば良い。

$$\begin{aligned} f(h_{t+1} | y^t) &= \int_{-\infty}^{\infty} f(h_{t+1} | h_t) f(h_t | y^t) dh_t \\ &= \frac{1}{\sqrt{2\pi(\tau^2 + \sigma_{t|t}^2)}} \exp\left(-\frac{(h_{t+1} - \mu_{t|t})^2}{2(\tau^2 + \sigma_{t|t}^2)}\right). \end{aligned}$$

(4) 予測分布③ (続き)

$f(h_{t+1} | y^t)$ は以下となる一方で,

$$f(h_{t+1} | y^t) = \frac{1}{\sqrt{2\pi(\tau^2 + \sigma_{t|t}^2)}} \exp\left(-\frac{(h_{t+1} - \mu_{t|t})^2}{2(\tau^2 + \sigma_{t|t}^2)}\right).$$

予測分布 $h_{t+1} | y^t$ の定義は, 以下より,

$$h_{t+1} | y^t \sim N(a_{t+1}, P_{t+1}).$$

$f(h_{t+1} | y^t)$ は,

$$f(h_{t+1} | y^t) = \frac{1}{\sqrt{2\pi P_{t+1}}} \exp\left(-\frac{(h_{t+1} - a_{t+1})^2}{2P_{t+1}}\right).$$

よって, 以下を得る.

$$a_{t+1} = \mu_{t|t}, \quad P_{t+1} = \tau^2 + \sigma_{t|t}^2$$

(4) まとめ：カルマン・フィルタとカルマン・ゲイン

予測分布 $h_t | y^{t-1} \sim N(a_t, P_t)$ が所与のとき, フィルタ分布の更新式は,

$$\mu_{t|t} = a_t + \frac{P_t}{P_t + \sigma^2}(y_t - a_t), \quad \sigma_{t|t}^2 = \frac{P_t \sigma^2}{P_t + \sigma^2}$$

フィルタ分布 $h_t | y^t \sim N(a_t, P_t)$ が所与のとき, 予測分布の更新式は,

$$a_{t+1} = \mu_{t|t}, \quad P_{t+1} = \tau^2 + \sigma_{t|t}^2$$

よって, $K_t = \frac{P_t}{P_t + \sigma^2}$ とするとき, 両方をあわせた更新式は以下となる.

$$a_{t+1} = \mu_{t|t} = a_t + K_t(y_t - a_t), \quad P_{t+1} = \tau^2 + P_t(1 - K_t)$$

補足: K_t を カルマン・ゲイン と呼ぶ. 観測値と予測の重み付けのバランスを制御する係数となる. カルマン・ゲインが大きいほど更新時の予測値は観測の影響を受ける.

(5) 尤度関数とパラメータ推定①：尤度の導出

フィルタ済み分布 $h_t | y^{t-1} \sim N(a_t, P_t)$ が与えられるとき, $y_t | y^{t-1}$ の密度関数は：

$$\begin{aligned} f(y_t | y^{t-1}) &= \int f(y_t | h_t) f(h_t | y^{t-1}) dh_t \\ &= \frac{1}{\sqrt{2\pi(P_t + \sigma^2)}} \exp\left(-\frac{(y_t - a_t)^2}{2(P_t + \sigma^2)}\right) \end{aligned}$$

よって, 観測系列 $y^T = \{y_t\}_{t=1}^T$ の尤度関数は：

$$f(y^T) = \prod_{t=1}^T f(y_t | y^{t-1}) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi(P_t + \sigma^2)}} \exp\left(-\frac{(y_t - a_t)^2}{2(P_t + \sigma^2)}\right) \quad (\text{i})$$

(5) 尤度関数とパラメータ推定②：更新式と最尤推定

パラメータ σ, τ の推定のために、尤度関数 (i) を最大化する。
ただし、初期値 a_1, P_1 は与えられていると仮定。

フィルタの更新式は以下の通り。

$$a_{t+1} = a_t + \frac{P_t}{P_t + \sigma^2}(y_t - a_t), \quad P_{t+1} = \tau^2 + P_t \left(1 - \frac{P_t}{P_t + \sigma^2}\right) \quad (\text{ii})$$

推定方法：

- 尤度関数 $f(y^T)$ を最大化する σ, τ を選ぶ（最尤推定）

注意： $f(y^T)$ は a_t, P_t を通じて σ, τ の関数である。

(6) 補足：正規分布の積の公式①

ベイズの定理により以下が成り立つ。

$$f(\mu_1 \mid x) f(x) = f(x \mid \mu_1) f(\mu_1).$$

今、正規分布 $\mu_1 \mid x \sim N(x, \sigma_1^2), x \sim N(\mu_2, \sigma_2^2)$ の密度関数を以下で与える。

$$f(\mu_1 \mid x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mu_1 - x)^2}{2\sigma_1^2}\right), \quad f(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right).$$

このとき、2つの密度関数の積をとると

$$\begin{aligned} f(\mu_1 \mid x) f(x) &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mu_1 - x)^2}{2\sigma_1^2}\right) \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right) \\ &= \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2\pi\sigma_1^2\sigma_2^2}} \exp\left(-\frac{(\sigma_1^2 + \sigma_2^2)\left(x - \frac{\mu_2\sigma_1^2 + \mu_1\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)^2}{2\sigma_1^2\sigma_2^2}\right) \\ &\quad \times \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right). \end{aligned}$$

(6) 補足：正規分布の積の公式②

$$\begin{aligned}f(\mu_1) &= \int_{-\infty}^{\infty} f(\mu_1 | x) f(x) dx \\&= \int_{-\infty}^{\infty} \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2\pi \sigma_1^2 \sigma_2^2}} \exp\left(-\frac{(\sigma_1^2 + \sigma_2^2)\left(x - \frac{\mu_2 \sigma_1^2 + \mu_1 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right)^2}{2\sigma_1^2 \sigma_2^2}\right) dx \\&\quad \times \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \\&= \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right).\end{aligned}$$

よって $\mu_1 \sim N(\mu_2, \sigma_1^2 + \sigma_2^2)$ となり、その密度関数は

$$f(\mu_1) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right).$$

(6) 補足：正規分布の積の公式③

残りの $x | \mu_1$ の分布は,

$$f(x | \mu_1) = \frac{f(\mu_1 | x) f(x)}{f(\mu_1)}.$$

よって

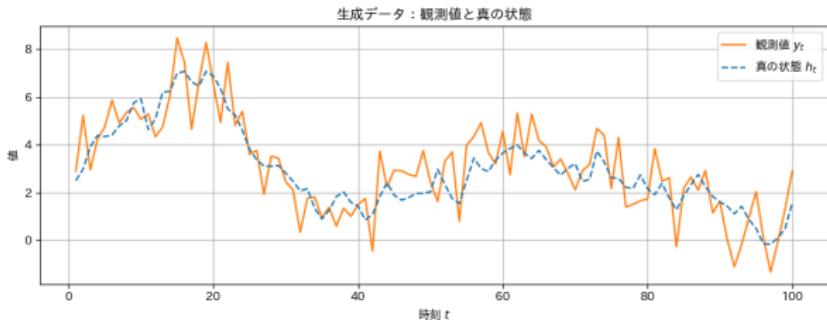
$$x | \mu_1 \sim N\left(\frac{\mu_2 \sigma_1^2 + \mu_1 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}, \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right),$$

密度関数は

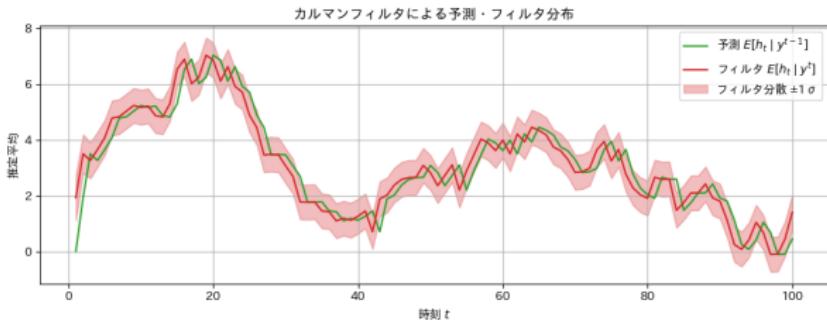
$$f(x | \mu_1) = \sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2\pi \sigma_1^2 \sigma_2^2}} \exp\left(-\frac{(\sigma_1^2 + \sigma_2^2)(x - \frac{\mu_2 \sigma_1^2 + \mu_1 \sigma_2^2}{\sigma_1^2 + \sigma_2^2})^2}{2\sigma_1^2 \sigma_2^2}\right).$$

サンプルコードからの作画 (ローカル・レベル・モデル)

サンプルコードからの作画(サンプルコード参照)



観測値と真の状態の推移



予測分布・フィルタ分布の期待値とフィルタ分布の 1σ 範囲推移

参考文献

- ① 時系列解析入門 北川源四郎 著 岩波書店
- ② 予測にいかす統計モデリングの基本 樋口知之 著 講談社
- ③ Rで学ぶデータサイエンス 3 ベイズ統計データ解析 姜興起 著
共立出版
- ④ カルマンフィルタ 野村俊一 著 共立出版
- ⑤ 時系列解析 (下 (非定常/応用定常過程編)) J.D. ハミルトン 著
シーエービー出版