カルマンフィルタ

樋口さぶろお https://hig3.net

龍谷大学大学院 理工学研究科 数理情報学専攻

理論物理学特論　L12 (2022-12-14 Wed)

最終更新: Time-stamp: "2022-12-14 Wed 14:29 JST hig"

今日の目標

- 時系列モデルのカルマンフィルタを説明できる

L10-Q4

Quiz 解答: 移動平均モデル MA(1) の例

1. $X_2 = \epsilon_2 + \theta_1 \epsilon_1$.
2. $\mathrm{E}[(\epsilon_2 + \theta_1 \epsilon_1)^2] = \sigma^2 (1 + \theta_1)^2$.
3. $\mathrm{E}[(\epsilon_2 + \theta_1 \epsilon_1)(\epsilon_3 + \theta_1 \epsilon_2)] = \sigma^2 \theta_1$.
4. 0.

L11-Q1

Quiz 解答:AR(4) の状態空間モデル

1. $\boldsymbol{X}_t = \begin{bmatrix} x_t \\ x_{t-1} \\ x_{t-2} \\ x_{t-3} \end{bmatrix}$

2. $\boldsymbol{x}_t = F\boldsymbol{x}_{t-1} + G\boldsymbol{v}_t, y_t = HVx_t + 0 \cdot w_t$ とすると

$$F = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, H = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}.$$

定義 (状態空間モデル (の簡単な場合))

- 観測できない量
  - $\mathbf{X}_t$: 状態ベクトル (2 次元「状態空間」の元)
  - $\mathbf{v}_t$: システムノイズベクトル (2 次元)
  - w_t: 観測ノイズ (1 次元) $\mathrm{V}[w_t] = \sigma^2$.
  - $\mathbf{v}_t, w_t$ は互いに独立なホワイトノイズ
- 観測できる量
  - Y_t: 観測値 (1 次元)
- 定数
  - F, G: 2×2 行列
  - H: 1×2 行列

$$\text{システム (状態) モデル} \quad \mathbf{X}_t = F\mathbf{X}_{t-1} + G\mathbf{v}_t$$

$$\text{観測モデル} \quad Y_t = H\mathbf{X}_t + w_t$$

カルマンフィルタ

- 状態空間モデルで, パラメタ $F, G, \sigma_v^2, \sigma^2$ を既知として, 観測された $Y_1, \ldots, Y_t$ から, 状態 $X_{t'}$ を推定する方法.
  - 平滑化 $t' < t$
  - フィルタ $t' = t$
  - 予測 $t' > t$
- 雑音として, 正規ホワイトノイズを仮定: $v_t \sim \mathrm{N}(0, \sigma_v^2 E)$ $w_t \sim \mathrm{N}(0, \sigma^2)$.
- 計算時間が少ない, 消費メモリーが少ない
- オンラインなアルゴリズム (↔ バッチなアルゴリズム)
  - $Y_1, \ldots, Y_{t-1}$ からの推測結果を持っているところに, データ Y_t が追加されたとき, 全てのデータからの計算し直しではなく, 現在の推測結果からの (少ない計算での) 修正が行える.
- Python library pykalman `https://pykalman.github.io/`
- カルマンフィルタを備える, 状態空間モデルのライブラリ `https://www.statsmodels.org/dev/generated/statsmodels.tsa.statespace.structural.UnobservedComponents.html`

相手にする分布

状態の周辺事後分布

$$f(x_{t'}|y_{1:t}) \overset{\text{略記}}{=} (t'|t)$$

- $y_{1:t}$ は $(y_1, \ldots, y_t)$ の略記.
- 周辺と言う名は t' 以外の $x_{t''}$ について周辺化した (和を取った) から.

分布の間の漸化式

	x_0	x_1	x_2	x_3	x_4
$y_{1:0} = \emptyset$	(0|0)	(1|0)	(2|0)	(3|0)	(4|0)
$y_{1:1}$	(0|1)	(1|1)	(2|1)	(3|1)	(4|1)
$y_{1:2}$	(0|2)	(1|2)	(2|2)	(3|2)	(4|2)
$y_{1:3}$	(0|3)	(1|3)	(2|3)	(3|3)	(4|3)
$y_{1:4}$	(0|4)	(1|4)	(2|4)	(3|4)	(4|4)

単純化: ローカルレベルモデル

定義 (直線上を歩く AirTag をつけた猫)

- 観測できない量
  - x_t: 猫の位置
  - $v_t \sim \mathrm{N}(0, \sigma_v^2)$: 猫の気まぐれな移動
  - $w_t \sim \mathrm{N}(0, \sigma^2)$: AirTag の誤差
- 観測できる量
  - y_t: AirTag の位置表示

$$\text{システム (状態) モデル} \quad x_t = F \cdot x_{t-1} + v_t$$
$$\text{観測モデル} \quad y_t = H \cdot x_t + w_t$$

定数 $F = H = 1$

ベイズ推定の復習と準備

L12-Q1

Quiz(正規分布のベイズ推定)

確率変数 X の確率密度関数 (' 事前分布') を

$$f_{\mathrm{X}}(x) = \frac{1}{\sqrt{2\pi C}} \mathrm{e}^{-\frac{(x-m)^2}{2C}}$$

とする.
確率変数 Y は $Y = X + \epsilon$, $\epsilon \sim \mathrm{N}(0, \sigma^2)$ で定まる.

1. $X = x_0$ という条件のもとで, Y の条件付き分布を求めよう.
2. Y として, y_0 が得られたという条件のもとで, X の条件付き分布 ('事後分布') を求めよう.

L12-Q2

Quiz(正規分布の母平均値のベイズ推定)

真の値 x を測定器で測定して値 y を得る.
測定器の出力 y は, 母平均値を x, 母分散を σ^2 とする正規分布にしたがう.

1. 1 個の測定値 y_1 を得たとき, 真の値の (ベイズで言う事後) 分布を答えよう.
2. さらにもう 1 個の測定値 y_2 を得たとき, 真の値の (ベイズ更新後の事後) 分布を答えよう.

ノイズが正規分布なので, (いちばん最初の事前分布が正規なら) 次の分布はすべて正規分布になる. 正規分布の加法性

略記 $f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\cdot\sigma^2}}\mathrm{e}^{-\frac{1}{2\cdot\sigma^2}(x-\mu)^2}$.

$$\text{フィルタ分布} f(x_t|y_{1:t}) = \frac{1}{\sqrt{2\pi\cdot C_t}}\mathrm{e}^{-\frac{1}{2\cdot C_t}(x_t-m_t)^2}$$

$$\text{一期先予測分布} f(x_t|y_{1:t-1}) = \frac{1}{\sqrt{2\pi\cdot \hat{C}_t}}\mathrm{e}^{-\frac{1}{2\cdot \hat{C}_t}(x_t-\hat{m}_t)^2}$$

$$\text{一期先予測尤度} f(y_t|y_{1:t-1})$$

$C_t, \hat{C}_t, m_t, \hat{m}_t$ を漸化式で計算して, $f(x_t|y_{1:t})$ を求めていける.

状態の更新

フィルタ分布 $f(x_{t-1}|y_{1:t-1}) = \frac{1}{\sqrt{2\pi\cdot C_{t-1}}}\mathrm{e}^{-\frac{1}{2\cdot C_{t-1}}(x_{t-1}-m_{t-1})^2}$ を持っているときに, システムモデルの式を使って, 一期先予測分布 $f(x_t|y_{1:t-1}) = \frac{1}{\sqrt{2\pi\cdot \hat{C}_t}}\mathrm{e}^{-\frac{1}{2\cdot \hat{C}_t}(x_t-\hat{m}_t)^2}$ を求める計算

$$
\begin{aligned}
f(x_t|y_{1:t-1}) &= \int dx_{t-1} f(x_t|x_{t-1}) f(x_{t-1}|y_{1:t-1}) \\
&= \int dx_{t-1} f(x_t - F x_{t-1}; 0, \sigma_v^2) f(x_{t-1}; m_{t-1}, C_{t-1}) \\
&\overset{\text{平方完成}}{=} f(x_t; \hat{m}_t, \hat{C}_t),
\end{aligned}
$$

$\hat{m}_t = \mathrm{E}[F x_{t-1} + v_t] = m_{t-1}$, $\hat{C}_t = \mathrm{V}[F x_{t-1} + v_t] = C_{t-1} + \sigma_v^2$
分布の間の漸化式で言うと, 横向きの矢印 $(t-1|t-1) \to (t|t-1)$

フィルタリング I

ここで正規分布のベイズ推定を使う
一期先予測分布 $f(x_t|y_{1:t-1}) = \frac{1}{\sqrt{2\pi \cdot \hat{C}_t}} \mathrm{e}^{-\frac{1}{2\cdot\hat{C}_t}(x_t - \hat{m}_t)^2}$ を持っているときに, 最新データ y_t が来たので, これを考慮してベイズ更新して, フィルタ分布 $f(x_t|y_{1:t})$ を求める計算.

$$
\begin{aligned}
f(x_t|y_{1:t}) =& f(x_t|y_{1:t-1}) \frac{f(y_t|x_t)}{f(y_t|y_{1:t-1})} \\
=& f(x_t; m_t, C_t)
\end{aligned}
$$

分布の間の漸化式で言うと, 下向きの矢印
$(t|t-1)$
$\downarrow$
$(t|t)$

フィルタリング II

$$C_t^{-1} = \hat{C}_t^{-1} + (\sigma^2)^{-1} = (C_{t-1} + \sigma_v^2)^{-1} + (\sigma^2)^{-1}$$

$$\begin{aligned}
m_t =& \frac{\hat{C}_t^{-1}}{\hat{C}_t^{-1} + (\sigma^2)^{-1}} \hat{m}_t + \frac{(\sigma^2)^{-1}}{\hat{C}_t^{-1} + (\sigma^2)^{-1}} y_t \\
=& \frac{(C_{t-1} + \sigma_v^2)^{-1}}{(C_{t-1} + \sigma_v^2)^{-1} + (\sigma^2)^{-1}} m_{t-1} + \frac{(\sigma^2)^{-1}}{(C_{t-1} + \sigma_v^2)^{-1} + (\sigma^2)^{-1}} y_t \\
K_t =& \frac{(\sigma^2)^{-1}}{(C_{t-1} + \sigma_v^2)^{-1} + (\sigma^2)^{-1}} \\
=& \hat{m}_t + \frac{\hat{C}_t}{\sigma^2 + \hat{C}_t} (y_t - \hat{m}_t)
\end{aligned}$$

K_t はカルマンゲイン (カルマン利得). 新着データ y_t で得られる情報の比率を表す.

(一般次元の) カルマンフィルタリング

(m_{t-1}, C_{t-1}) から (m_t, C_t) を得る手続き

システムの状態の更新
$$\hat{m}_t = G m_{t-1}$$
$$\hat{C}_t = G C_{t-1} G_t^{\mathrm{t}} + V$$

カルマンゲインの計算
$$f_t = F \hat{m}_t Q_t = \qquad F \hat{C} F^{\mathrm{t}} + W$$
$$K_t = \hat{C}_t F^{\mathrm{t}} Q_t^{-1}$$

ベイズ更新
$$m_t = \hat{m}_t + K_t (y_t - f_t)$$
$$C_t = (I - K_t F) \hat{C}_t$$

予測分布

(長期) 予測分布の更新

$$((t+k)\text{の予測}) = \int \mathrm{d}x_{t+k-1}\,\text{システムの状態更新} \times ((t+k-1)\text{の予測})$$

$$f(x_{t+k}|y_{1:t}) = \int \mathrm{d}x_{t+k-1} f(x_{t+k}|x_{t+k-1}) f(x_{t+k-1}|y_{1:t})$$

右向き矢印 $(t+k-1|t) \to (t+k|t)$

	x_0	x_1	x_2	x_3	x_4
$y_{1:0}=\emptyset$	(0|0)	(1|0)	(2|0)	(3|0)	(4|0)
$y_{1:1}$	(0|1)	(1|1)	(2|1)	(3|1)	(4|1)
$y_{1:2}$	(0|2)	(1|2)	(2|2)	(3|2)	(4|2)
$y_{1:3}$	(0|3)	(1|3)	(2|3)	(3|3)	(4|3)
$y_{1:4}$	(0|4)	(1|4)	(2|4)	(3|4)	(4|4)

平滑化

平滑化分布の更新 ($t' < t$)
フィルタリング分布 × 補正

$$f(x_{t'}|y_{1:t}) = f(x_{t'}|y_{1:t'}) \times \int \mathrm{d}x_{t'+1} \frac{f(x_{t'+1}|x_{t'})}{f(x_{t'+1}|y_{1:t})} f(x_{t'+1}|y_{1:t})$$

左向き矢印 $(t'|t) \leftarrow (t'+1|t)$

	x_0	x_1	x_2	x_3	x_4
$y_{1:0} = \emptyset$	(0|0)	(1|0)	(2|0)	(3|0)	(4|0)
$y_{1:1}$	(0|1)	(1|1)	(2|1)	(3|1)	(4|1)
$y_{1:2}$	(0|2)	(1|2)	(2|2)	(3|2)	(4|2)
$y_{1:3}$	(0|3)	(1|3)	(2|3)	(3|3)	(4|3)
$y_{1:4}$	(0|4)	(1|4)	(2|4)	(3|4)	(4|4)

状態空間モデルの扱いやすい場合 2: 隠れマルコフモデル

- 「$\mathbf{X}_t$ が離散型確率変数の場合」を考えたい. $X_t = 1, 2, \cdots, n$.
- システムモデルは「 $\mathbf{X}_t$ が確率的に時間発展する」ことを表すので, マルコフ連鎖で書くのが自然. M は推移確率行列. 計算科学 B, 確率モデル及び実習

$$
\begin{aligned}
P(X_t = 1) &= M_{11}P(X_{t-1} = 1) + M_{12}P(X_{t-1} = 2) + \cdots + M_{1n}P(X_{t-1} = n) \\
P(X_t = 2) &= M_{21}P(X_{t-1} = 1) + M_{22}P(X_{t-1} = 2) + \cdots + M_{2n}P(X_{t-1} = n) \\
&\vdots \\
P(X_t = n) &= M_{n1}P(X_{t-1} = 1) + M_{n2}P(X_{t-1} = 2) + \cdots + M_{nn}P(X_{t-1} = n)
\end{aligned}
$$

- 観測できない (隠れた)$X = 1, 2, \cdots, n$ という状態があり, ある法則で時間発展する. 観測される Y_t の値は X から決まる.