1 カルマンフィルタ

線形ガウス状態空間モデル：

$$\begin{cases} \boldsymbol{y}_t = X_t\boldsymbol{\beta} + Z_t\boldsymbol{\alpha}_t + G_t\boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 I) \\ \boldsymbol{\alpha}_{t+1} = W_t\boldsymbol{\beta} + T_t\boldsymbol{\alpha}_t + H_t\boldsymbol{u}_t \\ \boldsymbol{\alpha}_0 = \boldsymbol{0} \end{cases}$$

を考える。カルマンフィルタとは、観測地 $Y_n = \{Y_1, ..., Y_n\}$ の情報を用いて状態変数 $\boldsymbol{\alpha}_{t+1}$ の条件付き分布 $\pi(\boldsymbol{\alpha}_{t+1}|Y_t, \boldsymbol{\theta})$ を逐次的に求める方法である。まず、$t-1$ 期までの観測地が与えられた時の $\boldsymbol{\alpha}_t$ の分布を

$$\boldsymbol{\alpha}_t|\boldsymbol{\theta}, Y_{t-1} \sim (\boldsymbol{a}_t, \sigma^2 P_t)$$

とする。この時、t 期までの観測地が与えられた時の $\boldsymbol{\alpha}_{t+1}$ の分布も同様に、

$$\boldsymbol{\alpha}_{t+1}|\boldsymbol{\theta}, Y_t \sim (\boldsymbol{a}_{t+1}, \sigma^2 P_{t+1})$$

である。この時、$\boldsymbol{a}_{t+1}$ と $\boldsymbol{a}_t$、P_{t+1} と P_t には次のような逐次的な関係がある。

1. $\boldsymbol{a}_1 = W_0\boldsymbol{\beta},\ P_1 = H_0\ {}^T H_0$ とする。
2. $t = 1, ..., n$ に対し、

$$\boldsymbol{a}_{t+1} = W_t\boldsymbol{\beta} + T_t\boldsymbol{a}_t + K_t\boldsymbol{e}_t, \quad P_{t+1} = T_tP_t\ {}^T L_t + H_t\ {}^T J_t$$

ただし、

$$\begin{aligned} \boldsymbol{e}_t &= \boldsymbol{y}_t - X_t\boldsymbol{\beta} - Z_t\boldsymbol{a}_t \\ D_t &= Z_tP_t\ {}^T Z_t + G_t\ {}^T G_t \\ K_t &= (T_tP_t\ {}^T Z_t + H_t\ {}^T G_t)D_t^{-1} \\ L_t &= T_t - K_tZ_t \\ J_t &= H_t - K_tG_t \end{aligned}$$

解釈：状態方程式より単に $\boldsymbol{a}_{t+1} = W_t\boldsymbol{\beta} + T_t\boldsymbol{a}_t$ と推測できる。データ $\boldsymbol{y}_t$ が得られた時、単純な推定値 $W_t\boldsymbol{\beta}+T_t\boldsymbol{a}_t$ と実現値 $\boldsymbol{y}_t$ との乖離を $\boldsymbol{e}_t$ とし、$\boldsymbol{e}_t$ によって推定値を修正する。どれだけ修正するかは撹乱項の分散に依存する (K_t)。

$\boldsymbol{e}_t \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 D_t)$ なので、データ n 個の尤度関数は

$$f(Y_n|\boldsymbol{\theta}) = \Pi_{t=1}^n \frac{1}{\sqrt{(2\pi)^{m_y}\sigma^2|D_t|}} \exp(-\frac{1}{2\sigma^2}\ {}^T\boldsymbol{e}_tD_t^{-1}\boldsymbol{e}_t)$$

対数をとって、

$$\log f(Y_n|\boldsymbol{\theta}) = -\frac{nm_y}{2}\log(2\pi) - \frac{nm_y}{2}\log\sigma^2 - \frac{1}{2}\sum_{t=1}^{n}\log|D_t| - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\ {}^T\boldsymbol{e}_t D_t^{-1}\boldsymbol{e}_t$$

この対数尤度関数を最大化することでパラメータの最尤推定量を求めることができる。また、対数尤度に対数事前確率密度関数を加えたものを最大化するモードを求めて、モードを平均とする正規分布を提案分布とする MH アルゴリズムを行うこともできる。

2 平滑化

Y_n が得られた時、$t+1,...,n$ 期のデータも用いてより高精度の $\boldsymbol{a}_t$ を求めることが平滑化である。

2.1 状態平滑化 (kalman smoother)

1. カルマンフィルタを $t=1,...,n$ に対して行う。
2. $t=n,...,1$ に対し、

   $$\boldsymbol{r}_{t-1} = \ {}^T Z_t D_t^{-1}\boldsymbol{e}_t + \ {}^T L_t \boldsymbol{r}_t,\ U_{t-1} = \ {}^T Z_t D_t^{-1} Z_t + \ {}^T L_t U_t L_t$$

   と置く。
3. この時状態変数の事後平均、事後共分散行列は

   $$E[\boldsymbol{\alpha}_t|\boldsymbol{\theta}, Y_n] = \boldsymbol{a}_t + P_t\boldsymbol{r}_{t-1},\quad Var(\boldsymbol{\alpha}_t|\boldsymbol{\theta}, Y_n) = \sigma^2(P_t - P_t U_{t-1} P_t)$$

2.2 撹乱項平滑化

1. カルマンフィルタを $t=1,...,n$ について行う。
2. $\boldsymbol{r}_n = \boldsymbol{0},\ U_n = O$ と置く。
3. $t=n,..,1$ に対し、

   $$\boldsymbol{r}_{t-1} = \ {}^T Z_t D_t^{-1}\boldsymbol{e}_t + \ {}^T L_t \boldsymbol{r}_t,\ U_{t-1} = \ {}^T Z_t D_t^{-1}, Z_t + \ {}^T L_t U_t L_t$$

   と置く。
4. 撹乱項の事後平均および事後共分散行列は

   $$\begin{aligned} E[\boldsymbol{u}_t|\boldsymbol{\theta}, Y_n] &= \ {}^T G_t D_t^{-1}\boldsymbol{e}_t + \ {}^T J_t \boldsymbol{r}_t \\ Var(\boldsymbol{u}_t|\boldsymbol{\theta}, Y_n) &= \sigma^2(I - \ {}^T G_t D_t^{-1} G_t - \ {}^T J_t U_t J_t \end{aligned}$$

   となる。

2.3 シミュレーション平滑化 (de Jong and Shephard)

シミュレーション平滑化は錯乱項をサンプリングしてからいっぺんに $\boldsymbol{\alpha}_t$ を求める方法である。状態方程式の撹乱項 $G_t\boldsymbol{u}_t$ または観測方程式の撹乱項 $H_t\boldsymbol{u}_t$ をサンプリングするために、F_t を定数行列として $\boldsymbol{\eta}_t = F_t\boldsymbol{u}_t$ をサンプリングする方法を示す。

1. カルマンフィルタを実行し、$\{\boldsymbol{e}_t, D_t, J_t, L_t\}_{t=1}^{n}$ を保存する。

2. $\boldsymbol{r}_n = \boldsymbol{0},\ U_n = O$ として $t = n....,1$ に対し

$$
\begin{aligned}
C_t &= F_t(I - {}^TG_tD_t^{-1}G_t - {}^TJ_tU_tJ_t)\,{}^TF_t \\
\boldsymbol{\kappa}_t &\sim \mathcal{N}(\boldsymbol{0}, \sigma^2 C_t) \\
V_t &= F_t({}^TG_tD_t^{-1}Z_t + {}^TJ_tU_tL_t) \\
\boldsymbol{r}_{t-1} &= {}^TZ_tD_t^{-1}\boldsymbol{e}_t + {}^TL_t\boldsymbol{r}_t - {}^TV_tC_t^{-1}\boldsymbol{\kappa}_t \\
U_{t-1} &= {}^TZ_tD_t^{-1}Z_t + {}^TL_tU_tL_t + {}^TV_tC_t^{-1}V_t
\end{aligned}
$$

と $\{C_t, \boldsymbol{\kappa}_t, \boldsymbol{r}_t, U_t, V_t\}_{t=0}^{n}\}$ を求め、

$$\boldsymbol{\eta}_t = F_t({}^TG_tD_t^{-1}\boldsymbol{e}_t + {}^TJ_t\boldsymbol{r}_t)\boldsymbol{\kappa}_t,\quad t = 0, ..., n$$

を保存する。得られた $\boldsymbol{\eta} = (\boldsymbol{\eta}_0, ..., \boldsymbol{\eta}_n)$ は $\boldsymbol{\eta}$ の事後分布からの確率標本となる。

$F_t = H_t$ とすれば、状態方程式の撹乱項 $\boldsymbol{\eta}_t = H_t\boldsymbol{u}_t,\ (t = 1, ..., n-1)$ が得られるので、$\boldsymbol{\alpha}_{t+1}$ を計算できる。