

時系列解析（12）

－非線形・非ガウス型状態空間モデル－

東京大学 数理・情報教育研究センター

北川 源四郎

概要

- 復習
 - 非線形・非ガウス型状態空間モデル
 - 非線形・非ガウス型フィルタ・平滑化
- 非線形・非ガウス型モデリング（例）
- ガウス和フィルタ

非線形・非ガウス型状態空間モデル

$$\begin{aligned}x_n &= f(x_{n-1}, v_n) \\ y_n &= h(x_n, w_n)\end{aligned}$$

x_n : 状態

y_n : 時系列

v_n : システムノイズ

$$v_n \sim q(v)$$

w_n : 観測ノイズ

$$w_n \sim r(w)$$

非ガウス型フィルタ・平滑化

一期先予測

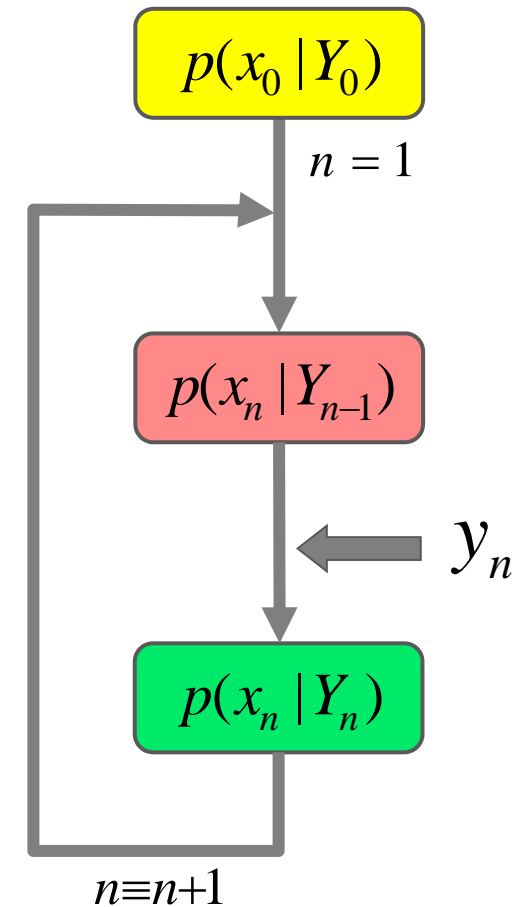
$$p(x_n | Y_{n-1}) = \int_{-\infty}^{\infty} p(x_n | x_{n-1}) p(x_{n-1} | Y_{n-1}) dx_{n-1}$$

フィルタ

$$p(x_n | Y_n) = \frac{p(y_n | x_n) p(x_n | Y_{n-1})}{p(y_n | Y_{n-1})}$$

平滑化

$$p(x_n | Y_N) = p(x_n | Y_n) \int_{-\infty}^{\infty} \frac{p(x_{n+1} | x_n) p(x_{n+1} | Y_N)}{p(x_{n+1} | Y_n)} dx_{n+1}$$



分布の近似

0. 線形・正規モデル近似

カルマンフィルタ・平滑化

1. 正規分布近似

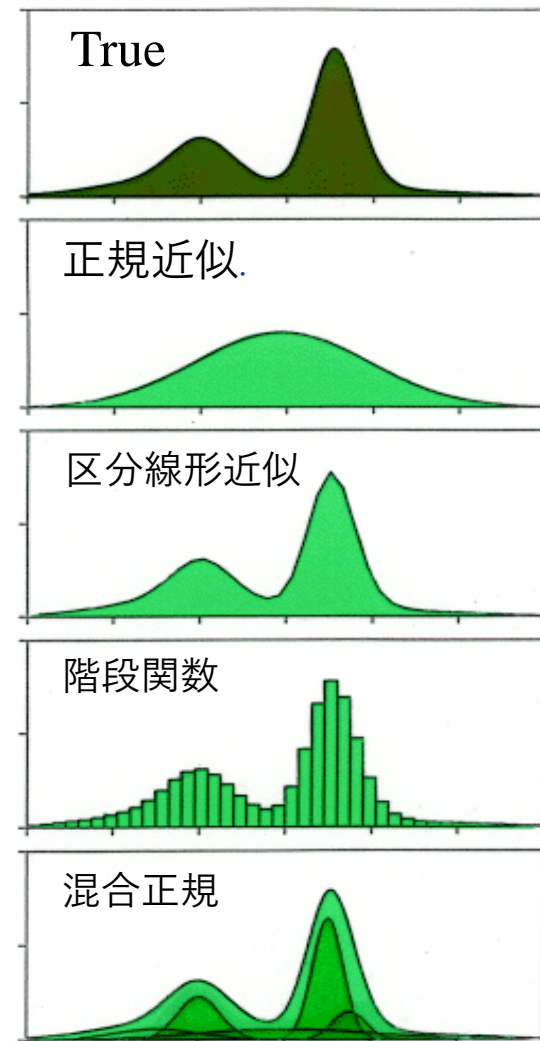
拡張カルマンフィルタ・平滑化

2. 区分線形（階段関数）近似

非ガウス型フィルタ・平滑化

3. 混合正規分布近似

ガウス和フィルタ・平滑化



非ガウス型フィルタ・平滑化の応用例

1. 非ガウス型平滑化

- 非ガウス型システムノイズ（ジャンプの検出）
- 非ガウス型観測ノイズ（時変分散の推定）

2. 非線形平滑化

3. 一般化状態空間モデル

- 非定常離散系列

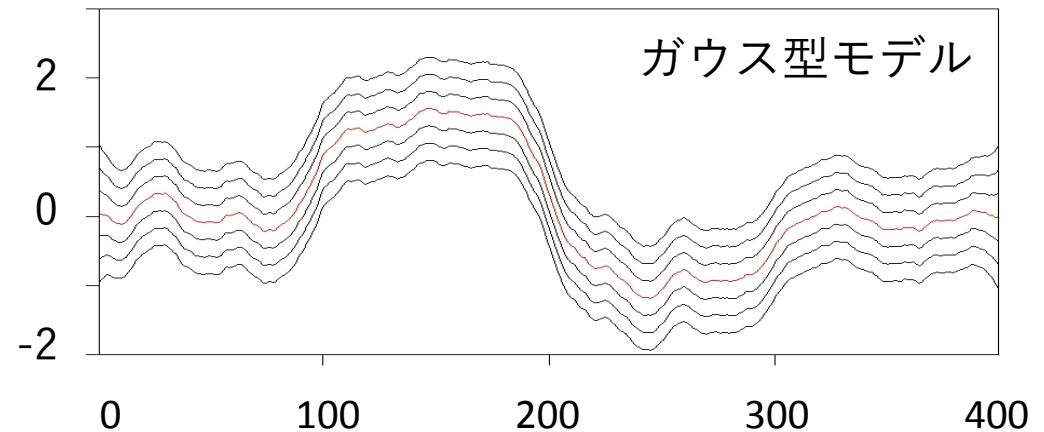
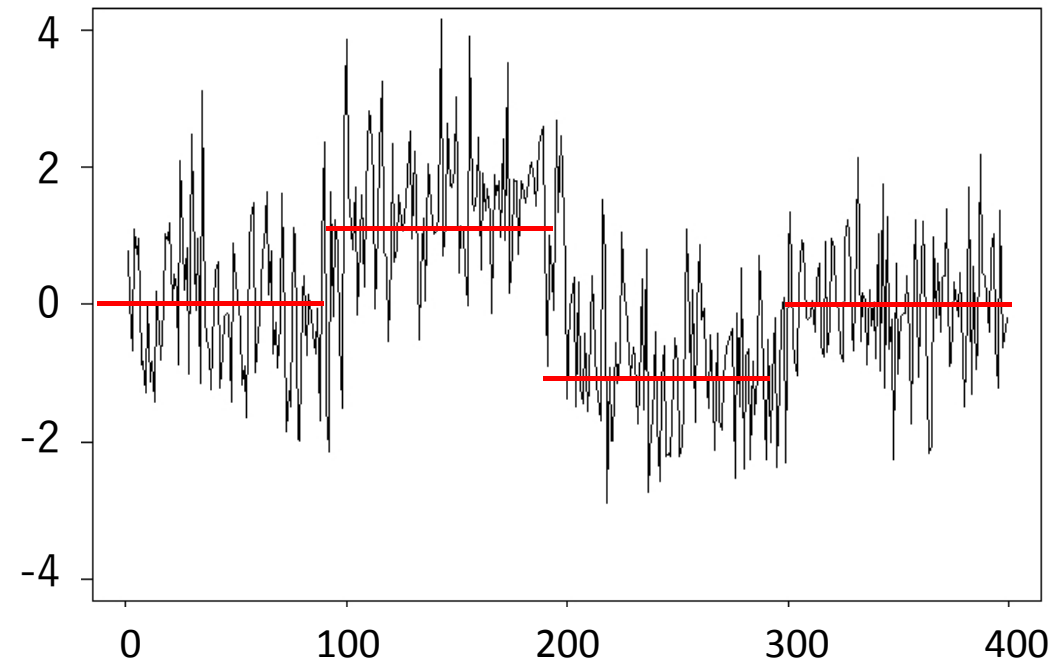
4. 自己組織型フィルタ・平滑化

- 状態とパラメータの同時推定

トレンドのジャンプの検出

$$y_n \sim N(\mu_n, 1)$$

$$\mu_n = \begin{cases} 0 & 1 \leq n \leq 100 \\ 1 & 101 \leq n \leq 200 \\ -1 & 201 \leq n \leq 300 \\ 0 & 301 \leq n \leq 400 \end{cases}$$



トレンドモデル (1次)

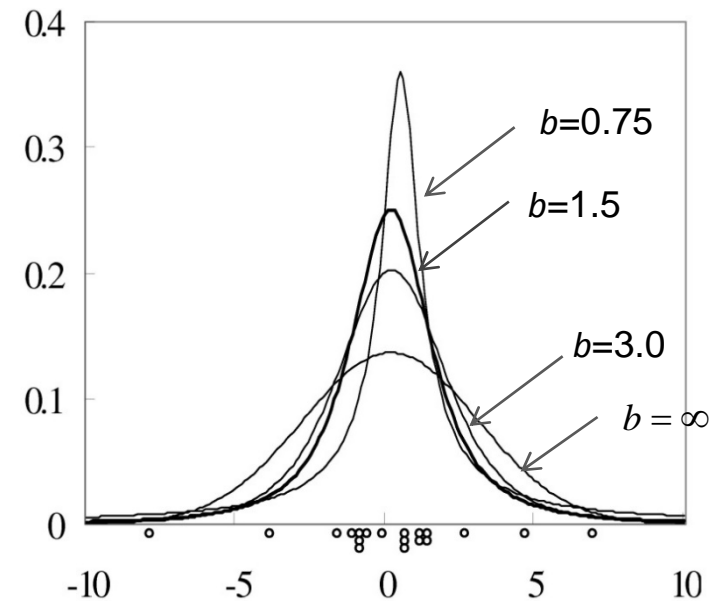
$$t_n = t_{n-1} + v_n$$

$$y_n = t_n + w_n$$

ノイズ分布

- 観測ノイズ $w_n \sim N(0, \sigma^2)$
- システムノイズ

$$p(v_n) = \begin{cases} \frac{1}{\sqrt{2\pi\tau^2}} \exp\left\{-\frac{v_n^2}{2\tau^2}\right\} \\ \frac{\tau^{2b-1}\Gamma(b)}{\Gamma(\frac{1}{2})\Gamma(b-\frac{1}{2})} \frac{1}{(v_n^2 + \tau^2)^b} \end{cases}$$



正規分布

ピアソン分布族

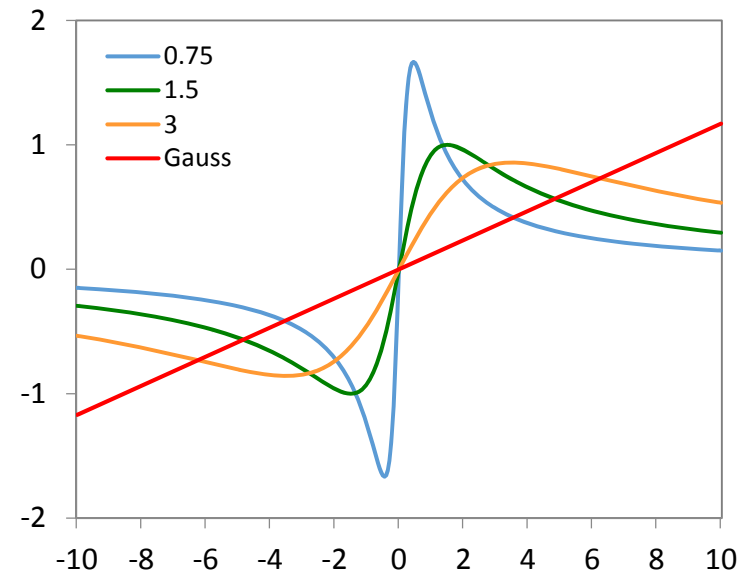
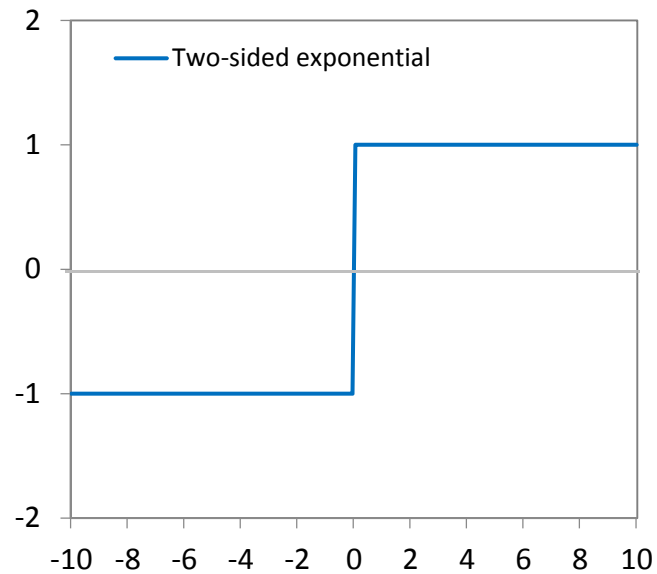
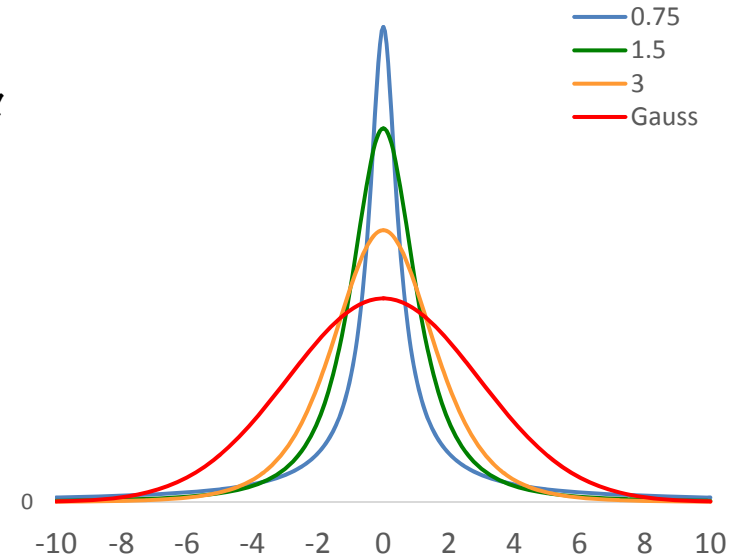
ピアソン分布族とロバスト推定

$$p_b(v) = \frac{C}{(v^2 + \tau^2)^b}$$

b : 形状パラメータ

$$\frac{d \log p_b(v)}{dv} = \frac{2bx}{(v^2 + \tau^2)}$$

影響関数



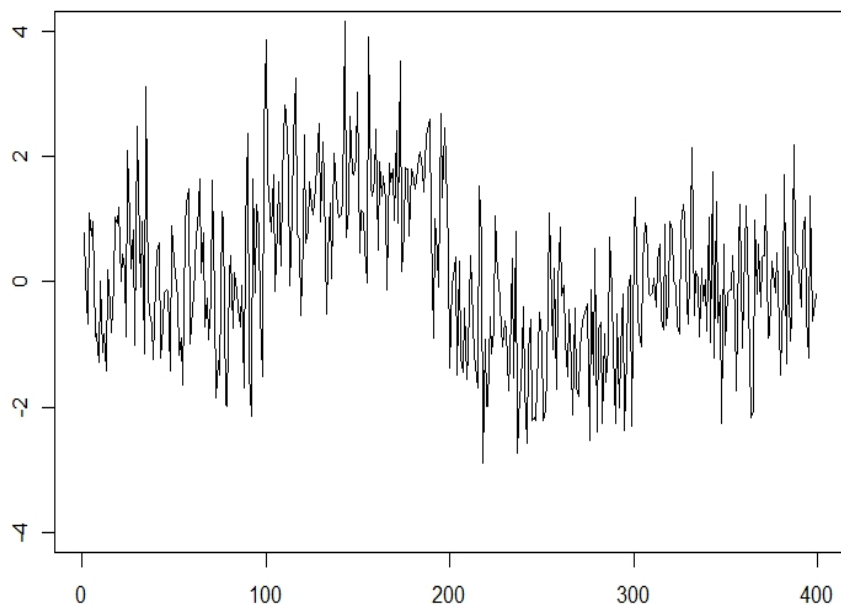
ピアソン分布族の比較

| b | τ^2 | σ^2 | 対数尤度 | AIC | 分布形 |
|----------|------------------------|------------|---------|----------------|--------|
| 0.60 | 0.211×10^{-9} | 1.042 | -597.19 | 1198.38 | |
| 0.75 | 0.299×10^{-7} | 1.043 | -597.39 | 1198.78 | |
| 1.00 | 0.353×10^{-4} | 1.045 | -597.99 | 1198.98 | Cauchy |
| 1.50 | 0.303×10^{-2} | 1.045 | -599.13 | 1202.26 | t_2 |
| 3.00 | 0.406×10^{-1} | 1.046 | -600.40 | 1204.80 | t_5 |
| ∞ | 0.140×10^{-1} | 1.048 | -600.69 | 1205.38 | Normal |

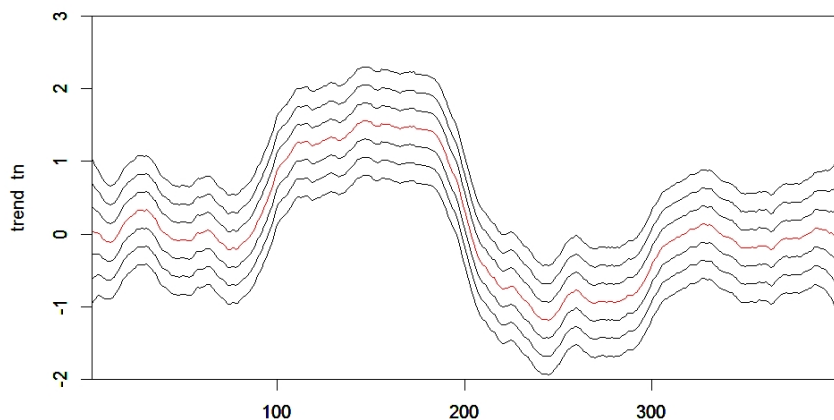
$$b = \frac{3}{m} \quad (m = 0, \dots, 5)$$

$$b = \frac{k+1}{2} \quad (k : t \text{分布の自由度})$$

非ガウスモデルによる構造変化の自動検出



ガウス型モデル



`ngsmth(y, noisev = 1, tau2 = 1.4e-02, noisew = 1, sigma2 = 1.048)`

トレンドモデル

$$t_n = t_{n-1} + v_n$$

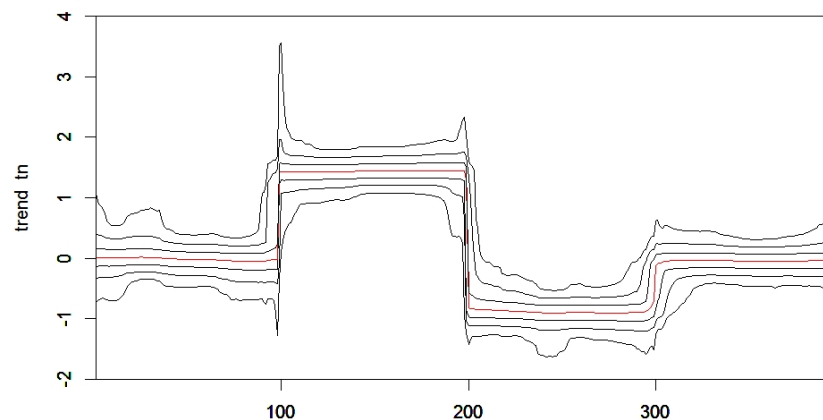
$$y_n = t_n + w_n$$

ノイズ分布

$v_n \sim p(0, \tau^2, b)$ ピアソン分布

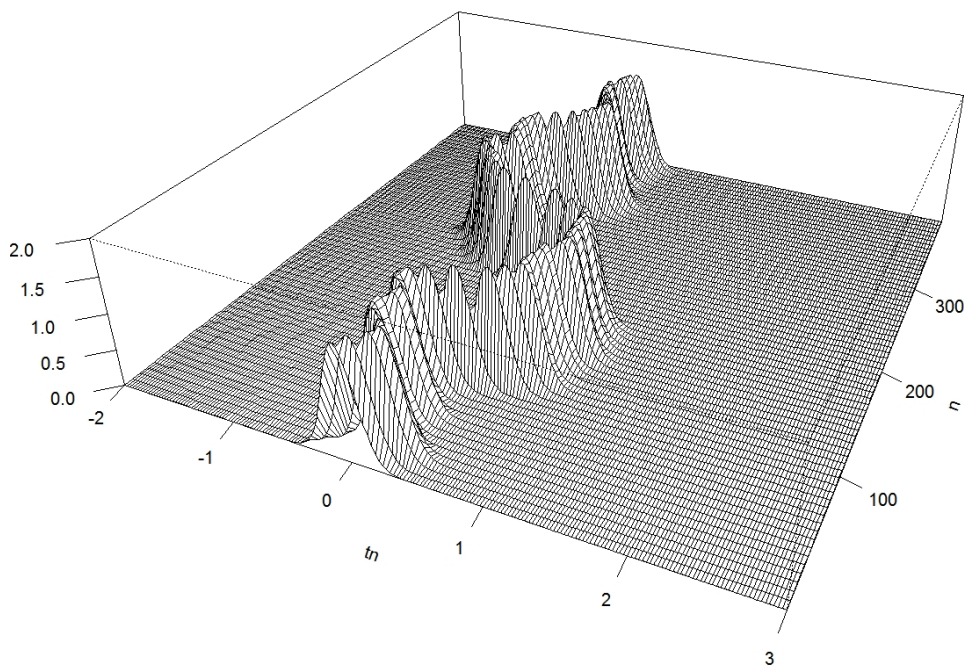
$w_n \sim N(0, \sigma^2)$ 正規分布

非ガウス型モデル



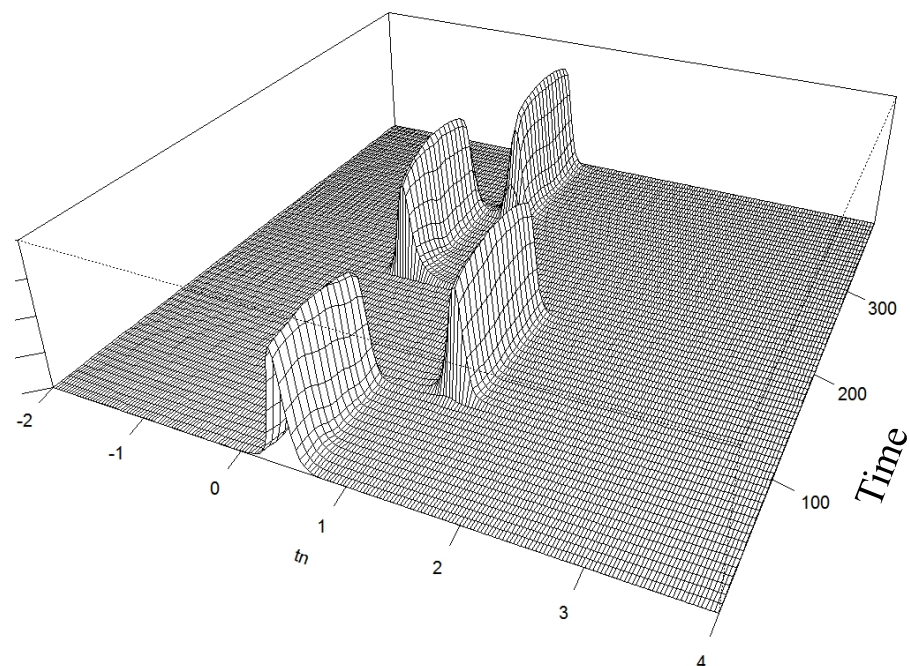
`ngsmth(y, noisev = 2, tau2 = 2.11e-10, bv = 0.6, noisew = 1, sigma2 = 1.042)`

非ガウス型平滑化による事後分布



ガウス分布モデル

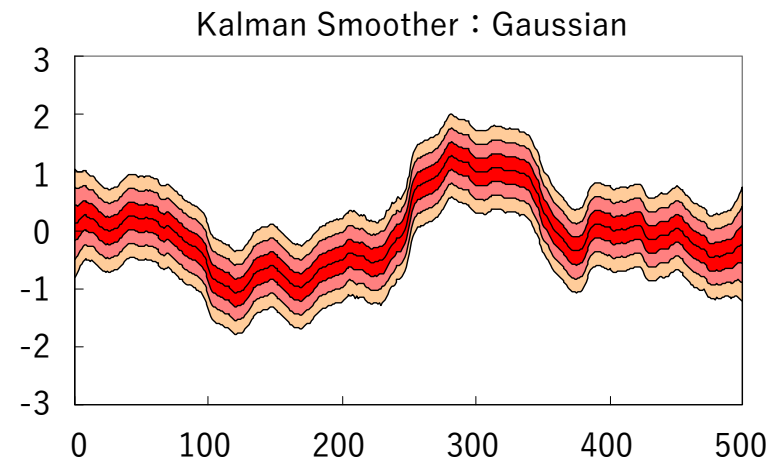
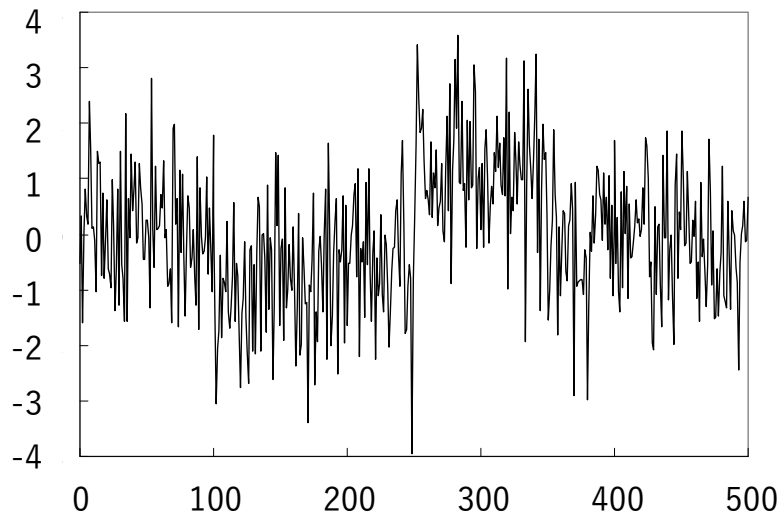
```
s1 <- ngsmth(y, noisev = 1, tau2 = 1.4e-02, noisew = 1, sigma2 = 1.048)
par(mar=c(0,2,0,0)+0.1)
plot(s1, "smt", theta = 25, phi = 30, expand = 0.25, col="white")
```



ピアソン分布族モデル

```
s2 <- ngsmth(y, noisev = 2, tau2 = 3.53e-5, bv = 1.0, noisew = 1,
sigma2 = 1.045)
plot(s2, "smt", theta = 25, phi = 30, expand = 0.25, col="white")
```

システムノイズ分布の比較



トレンドモデル

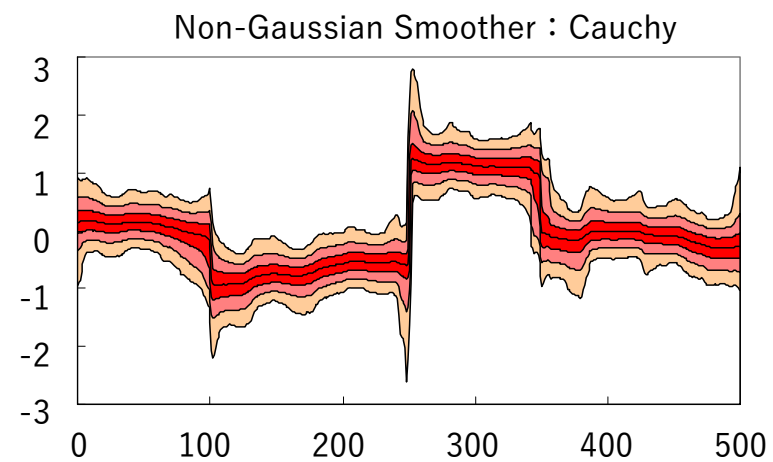
$$t_n = t_{n-1} + v_n$$

$$y_n = t_n + w_n$$

ノイズ分布

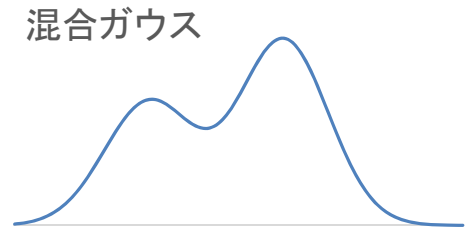
$$v_n \sim q(v)$$

$$w_n \sim N(0, \sigma^2)$$

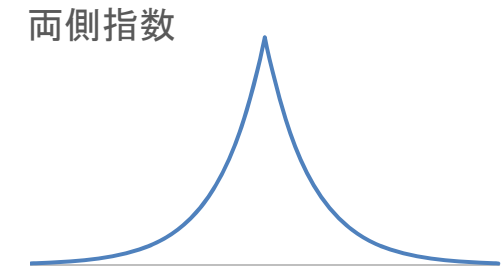


その他のノイズ分布

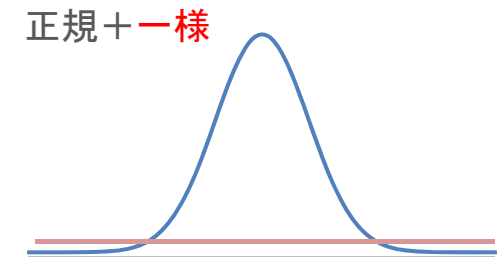
混合ガウス分布



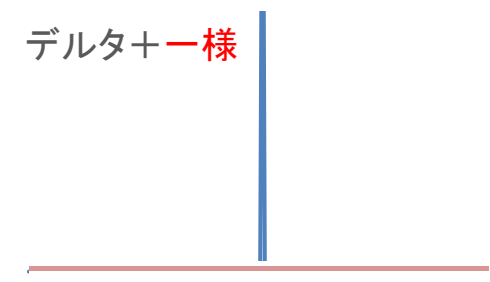
両側指数分布 (ラプラス分布)



正規分布と一様分布の混合

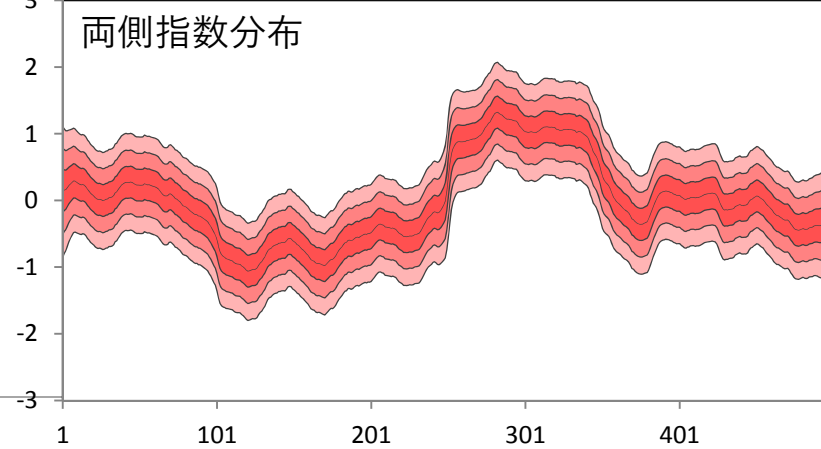
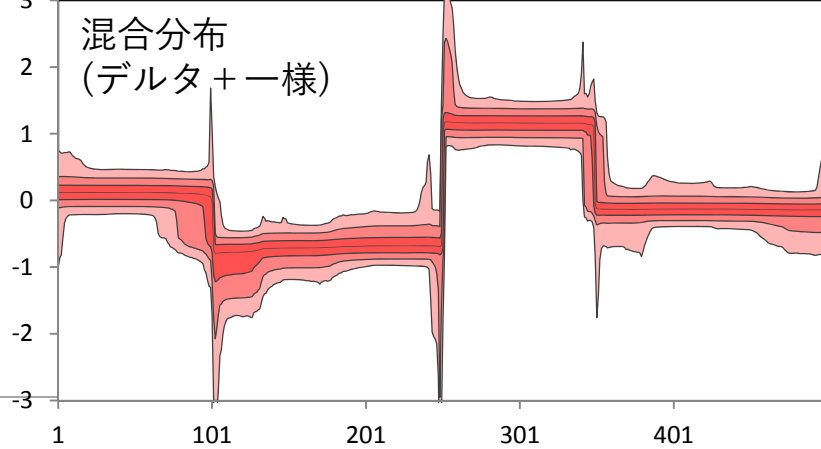
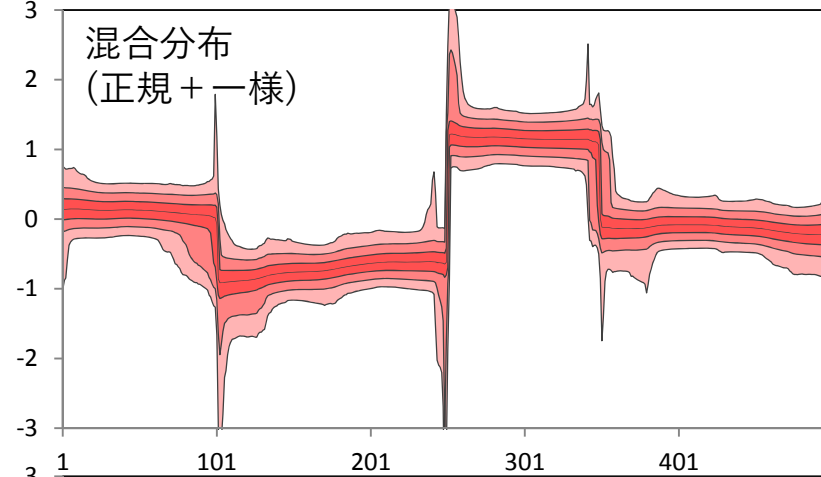
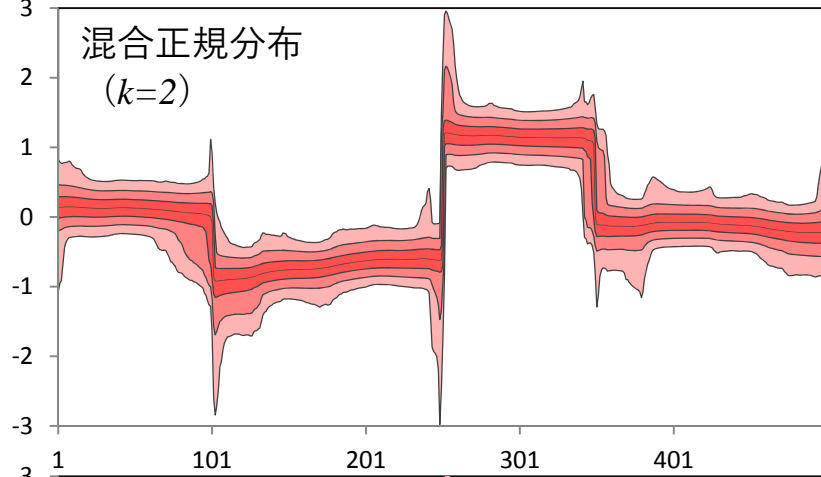
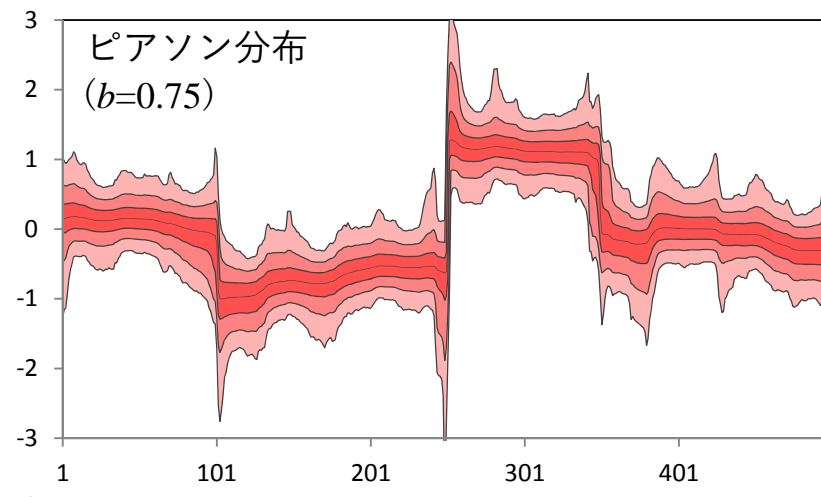
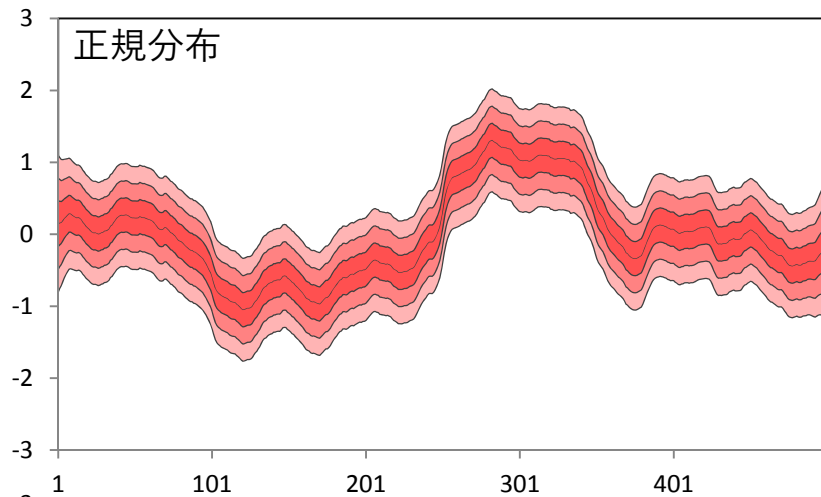


デルタ分布と一様分布の混合



モデル比較

| 分布 | tau2 | alpha | FF |
|----------|-----------------------|-------|-----------------|
| 正規 | 1.22×10^{-2} | | -748.100 |
| Pearson | 2.20×10^{-7} | | -742.492 |
| 混合正規 | 4.30×10^{-4} | 0.991 | -741.440 |
| 両側指数 | 1.22×10^1 | | -747.566 |
| 正規 + 一様 | 4.33×10^{-4} | 0.992 | -742.584 |
| デルタ + 一様 | — | 0.892 | -742.837 |

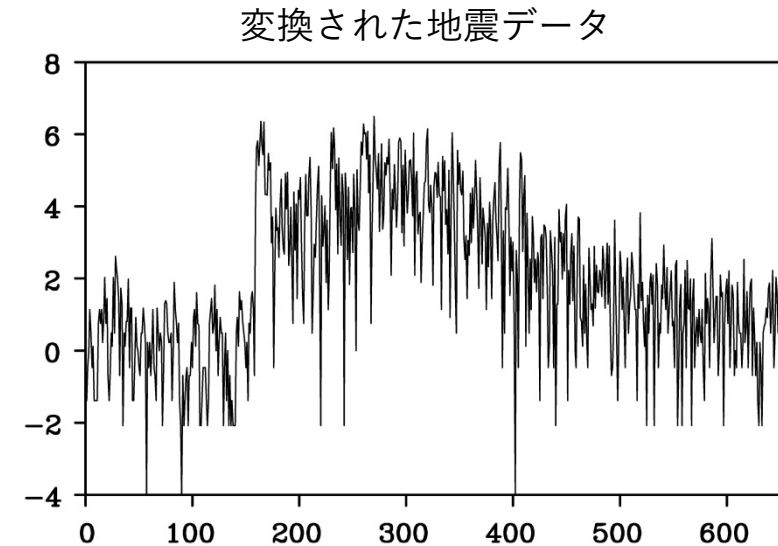


時変分散の推定

$$u_m = \log\{(y_{2m-1}^2 + y_{2m}^2) / 2\}$$

$$t_m = t_{m-1} + v_m$$

$$u_m = t_m + w_m$$

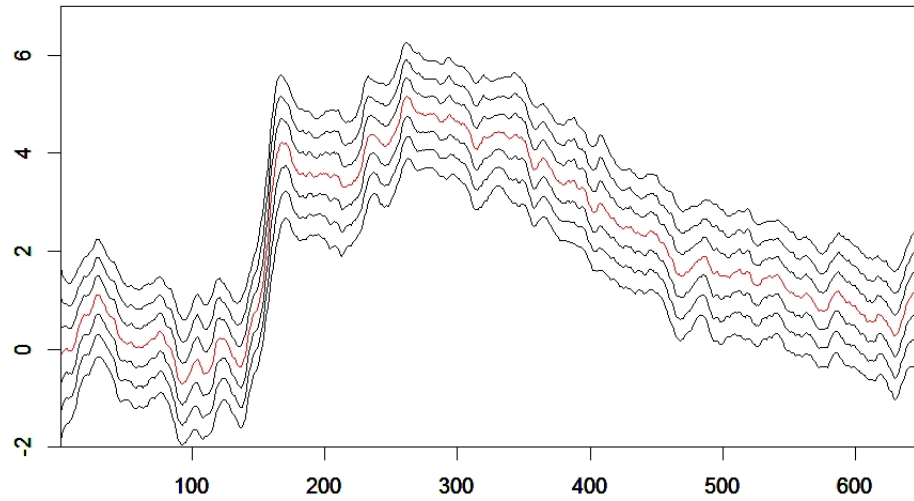


| システムノイズ | 観測ノイズ | AIC |
|------------------|----------------------|---------|
| $N(0, 0.04909)$ | $N(-\zeta, \pi^2/6)$ | 2165.10 |
| $C(0, 0.000260)$ | $\exp\{w - e^w\}$ | 2056.97 |

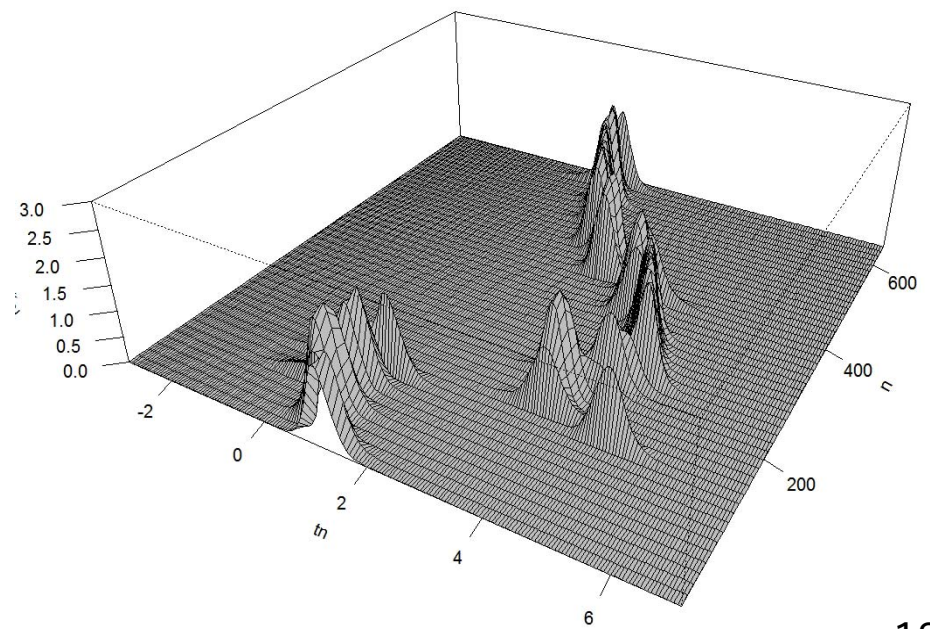
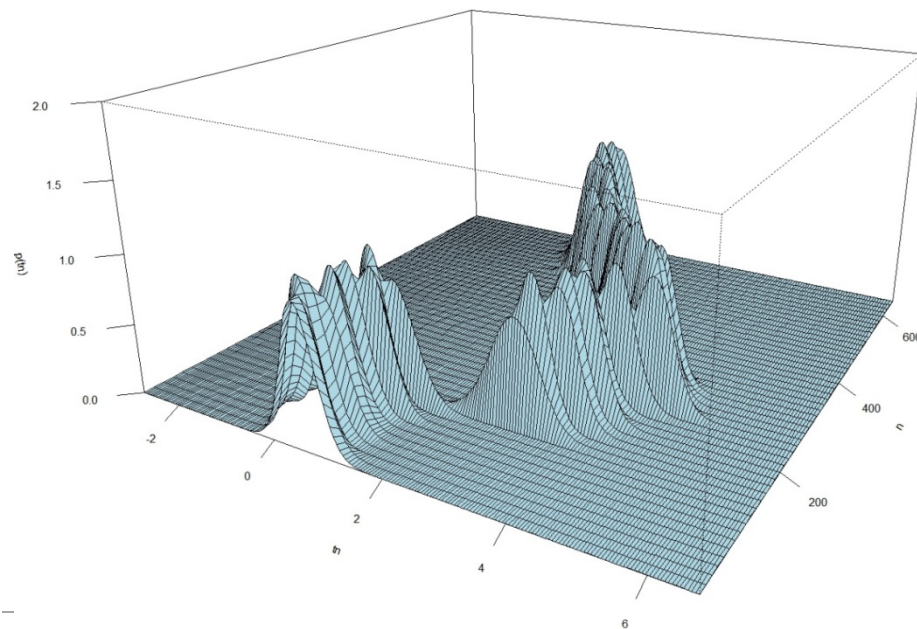
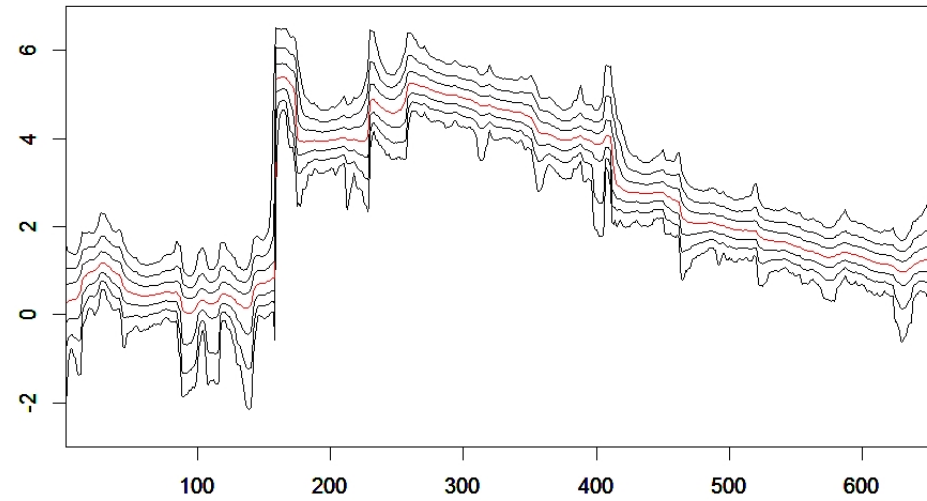
$h(w) = \exp\{w - e^w\}$: 二重指数分布

時変分散の推定

Gauss & Gauss

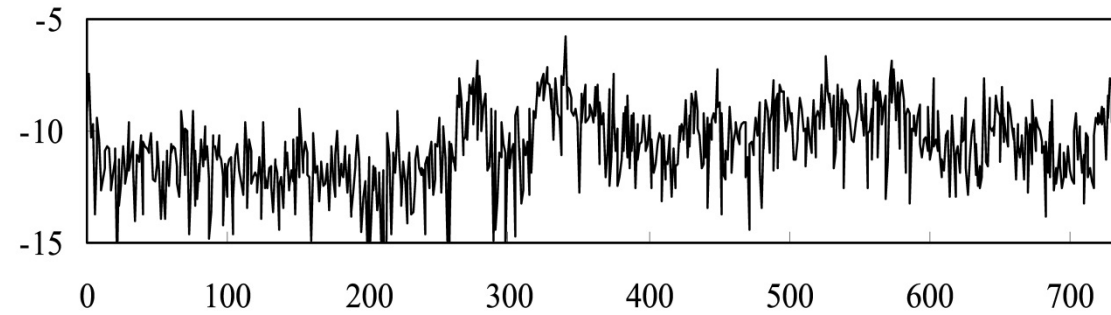


Cauchy & Double Exponential

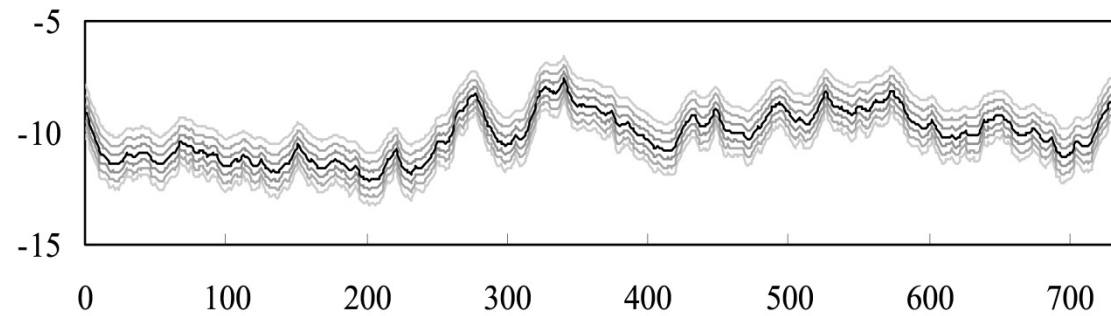


日経225平均株価データ

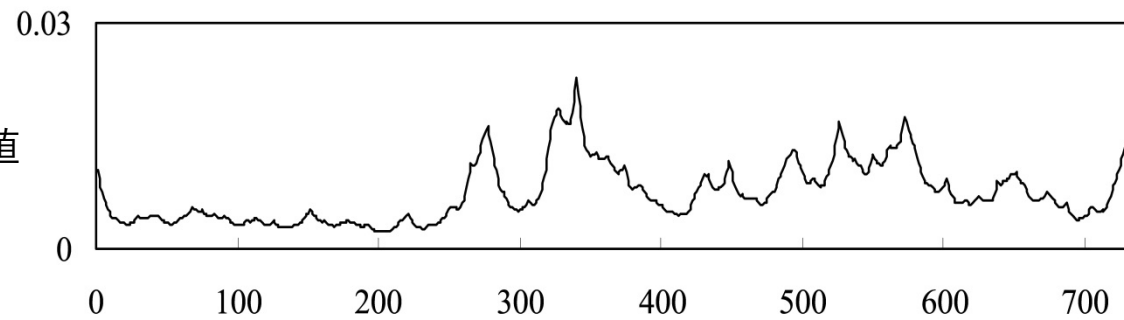
変換後



$\log \sigma_n^2$ の事後分布



ボラティリティの推定値
 σ_n



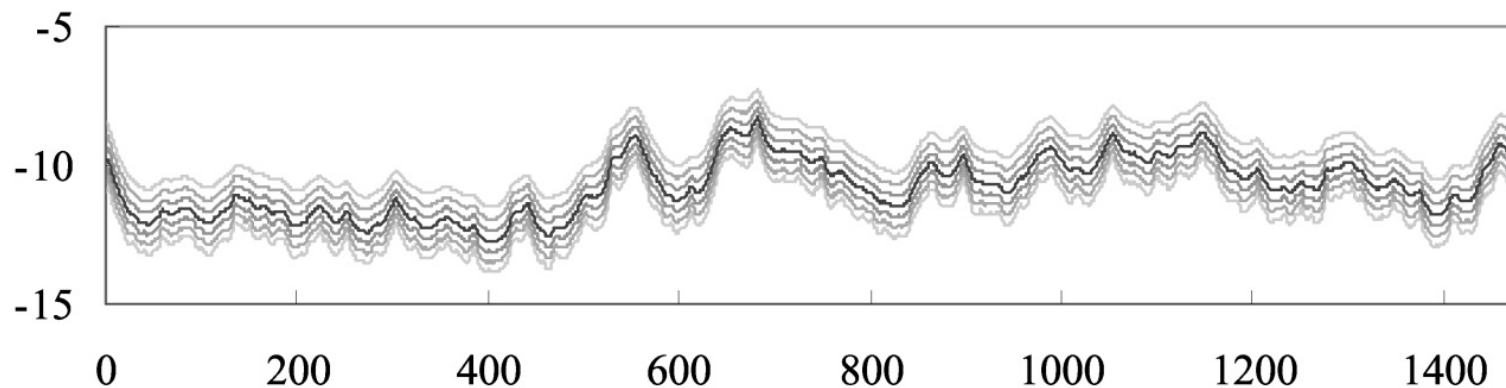
時変分散の直接推定

非線形モデルを用いる場合、データの変換($u=\log(y^2)$ など)は、必ずしも必要ではない

$$t_n = t_{n-1} + v_n$$

$$y_n \sim N(0, e^{t_n})$$

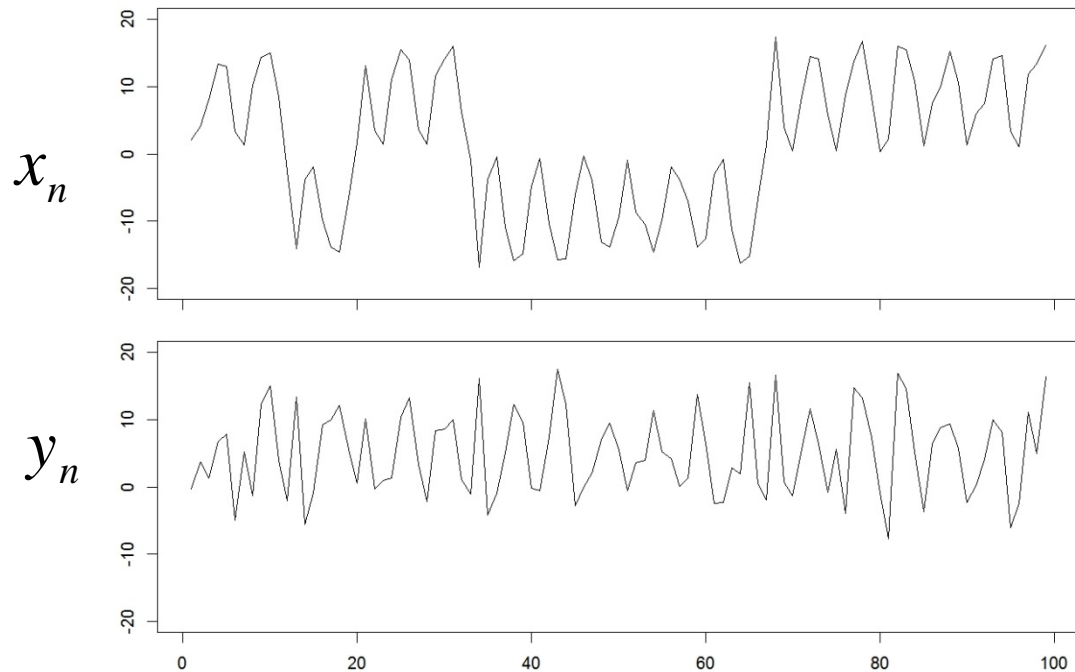
$$y_n = e^{t_n} w_n$$



非線形モデル

$$x_n = \frac{1}{2}x_{n-1} + \frac{25x_{n-1}}{1+x_{n-1}^2} + 8\cos(1.2n) + v_n$$

$$y_n = \frac{x_n^2}{20} + w_n \quad v_n \sim N(0,1), w_n \sim N(0,10)$$



信号

観測値

信号復元

- 観測モデル $h(x) = x^2$
- y_n から x_n の推定は難しい（単一時点のデータからは正負の判定は不可能）
- 時系列情報を使えば推定可能

拡張カルマンフィルタ

予測

$$x_{n|n-1} = f_n(x_{n-1|n-1}) + v_n$$

$$V_{n|n-1} = F_n V_{n-1|n-1} F_n^T + Q$$

フィルタ

$$K_n = V_{n|n-1} H_n (H_n^T V_{n|n-1} H_n + R)^{-1}$$

$$x_{n|n} = x_{n|n-1} + K_n (y_n - h_n(x_{n|n-1}))$$

$$V_{n|n} = V_{n|n-1} - K_n H_n^T V_{n|n-1}$$

$$f_n(x) = \frac{1}{2}x + \frac{25x}{1+x^2} + 8\cos(1.2n)$$

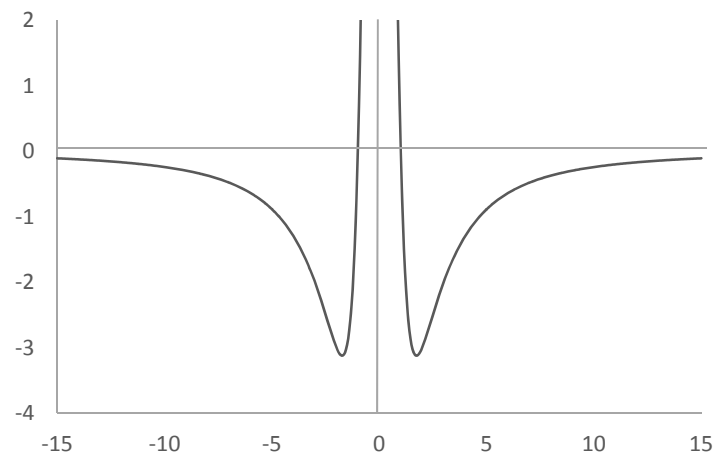
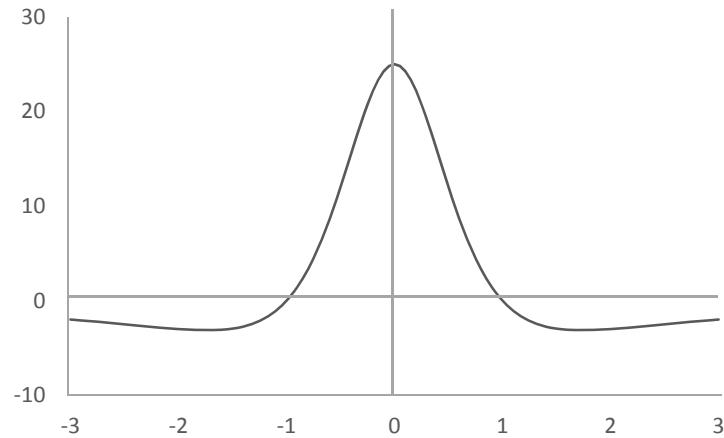
$$h(x) = \frac{x^2}{20}$$

$$F_n = \left. \frac{\partial f_n(x)}{\partial x} \right|_{x=x_{n-1|n-1}} = \frac{25(1-x_{n-1|n-1}^2)}{(1+x_{n-1|n-1}^2)^2}$$

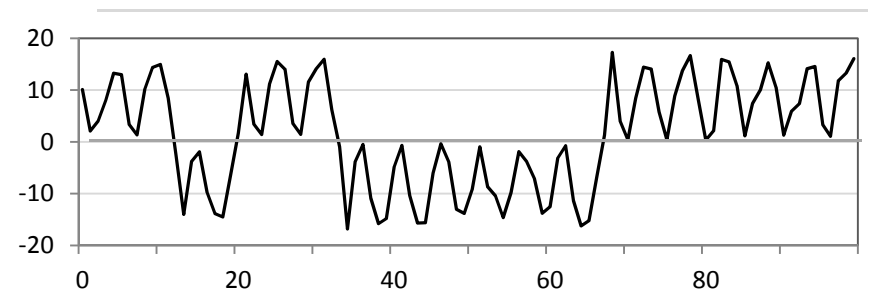
$$H_n = \left. \frac{\partial h(x)}{\partial x} \right|_{x=x_{n|n-1}} = \frac{x_{n|n-1}}{10}$$

拡張カルマンフィルタ・平滑化による結果

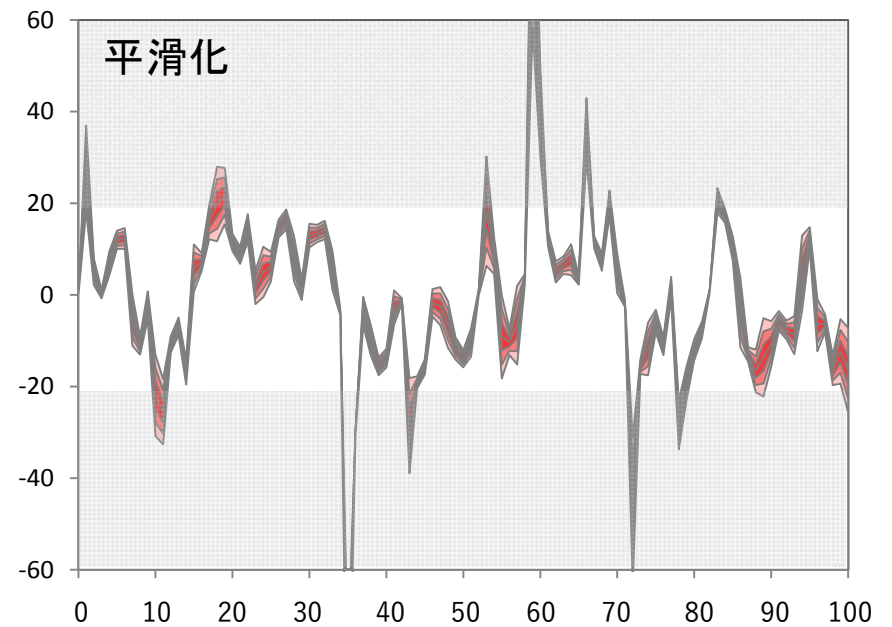
$$F(x) = \frac{df(x)}{dx}$$



True State



平滑化



非ガウス型フィルタ（予測分布の数値計算）

$$x_n = \frac{1}{2} x_{n-1} + \frac{25x_{n-1}}{1 + x_{n-1}^2} + 8\cos(1.2n) + v_n$$

$$p(x_n | Y_{n-1}) = \int_{-\infty}^{\infty} p(x_n | x_{n-1}) p(x_{n-1} | Y_{n-1}) dx_{n-1}$$

$$p(x_n | Y_n) = \frac{p(y_n | x_n) p(x_n | Y_{n-1})}{p(y_n | Y_{n-1})}$$

For $i = 1$ to k

$$f_i = 0$$

For $i = 1$ to k

(a) $y_0 = \min\{f(x_{i-1}), f(x_i)\}$

(b) $y_3 = \max\{f(x_{i-1}), f(x_i)\}$

(c) $i_0 = [(y_0 - z_0)/\Delta x]$, $i_1 = [(y_3 - z_0)/\Delta x]$

(d) for $j = i_0 + 1$ to i_1

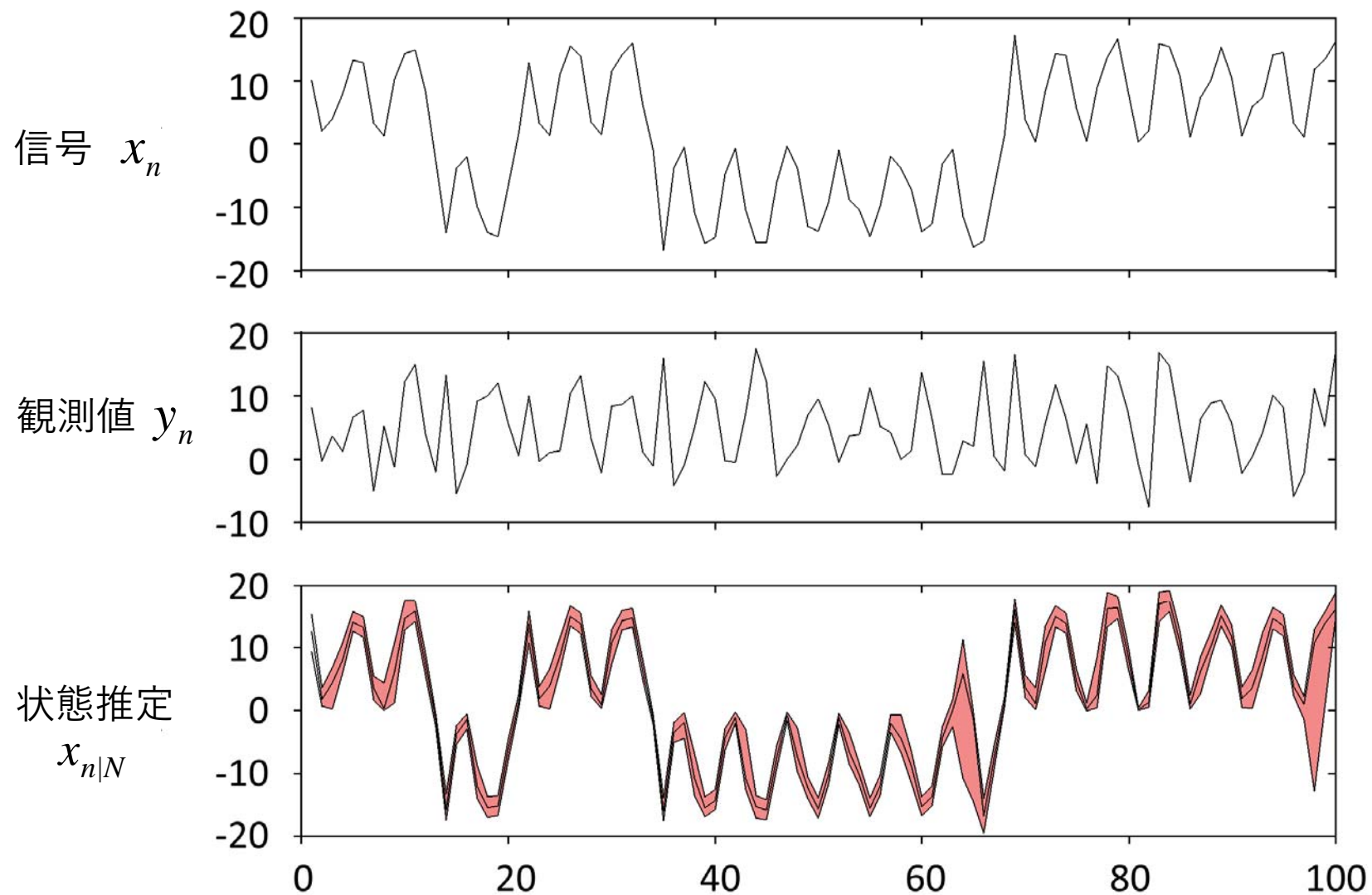
- $y_1 = \max\{y_0, z_0 + (j-1)\Delta x\}$

- $y_2 = \min\{z_0 + y_3, j\Delta x\}$

- $f_j = f_j + (y_2 - y_1)/(y_3 - y_0) f_i$

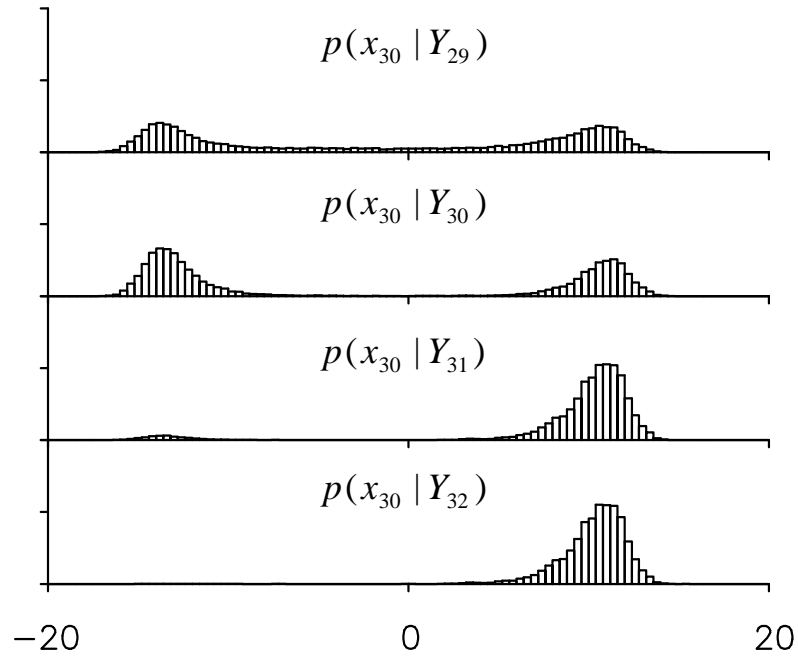
$$p_i = q_{i-1} f_1 + \dots + q_{i-k} f_k$$

非ガウス型平滑化による信号復元



非ガウス型分布の効果

x_{30} の推定

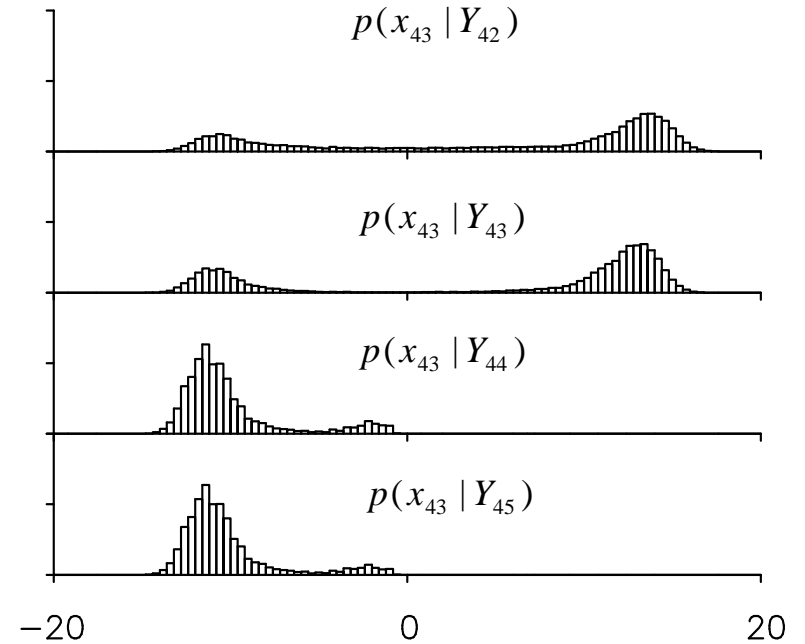


予測分布

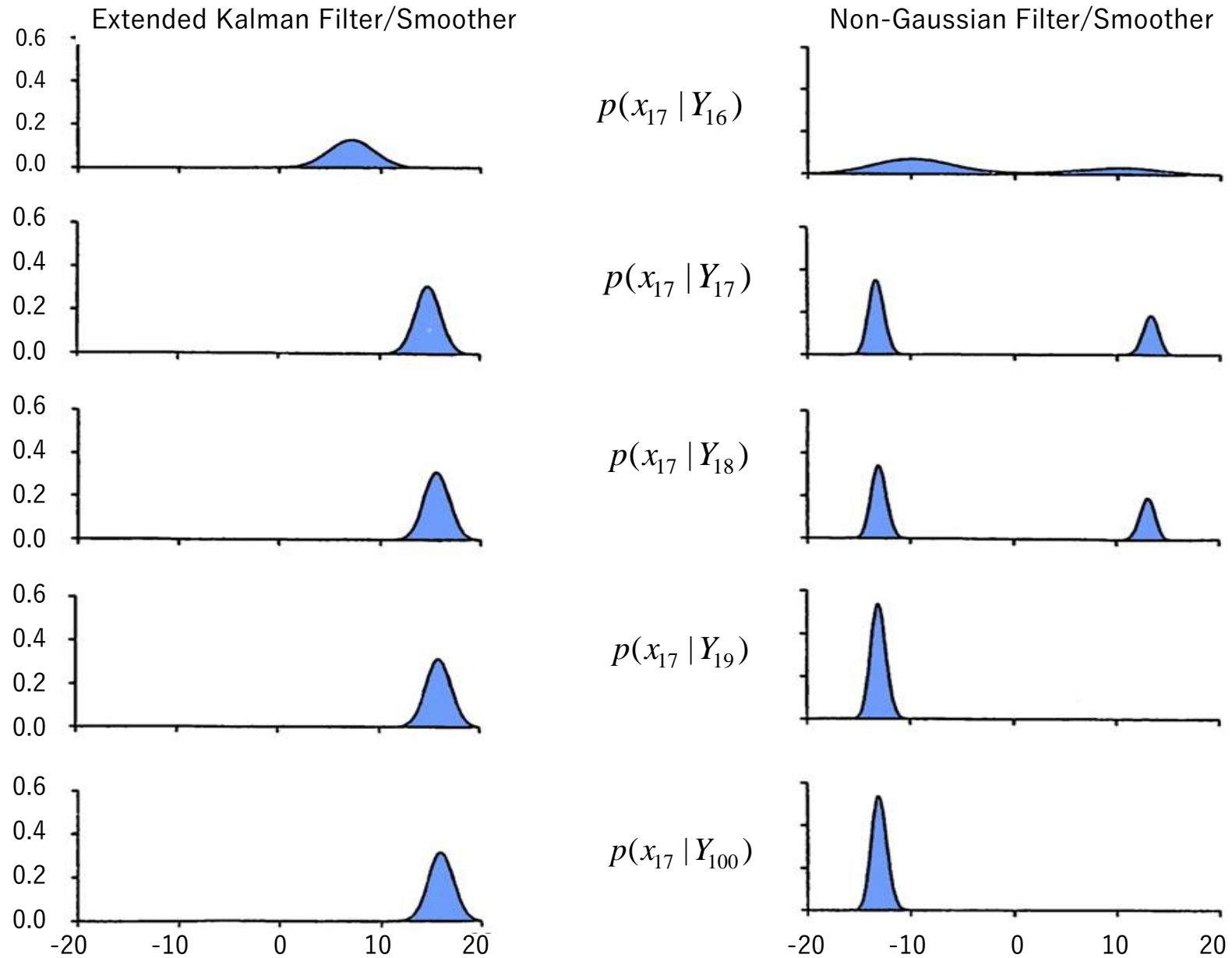
フィルタ

平滑分布

x_{43} の推定

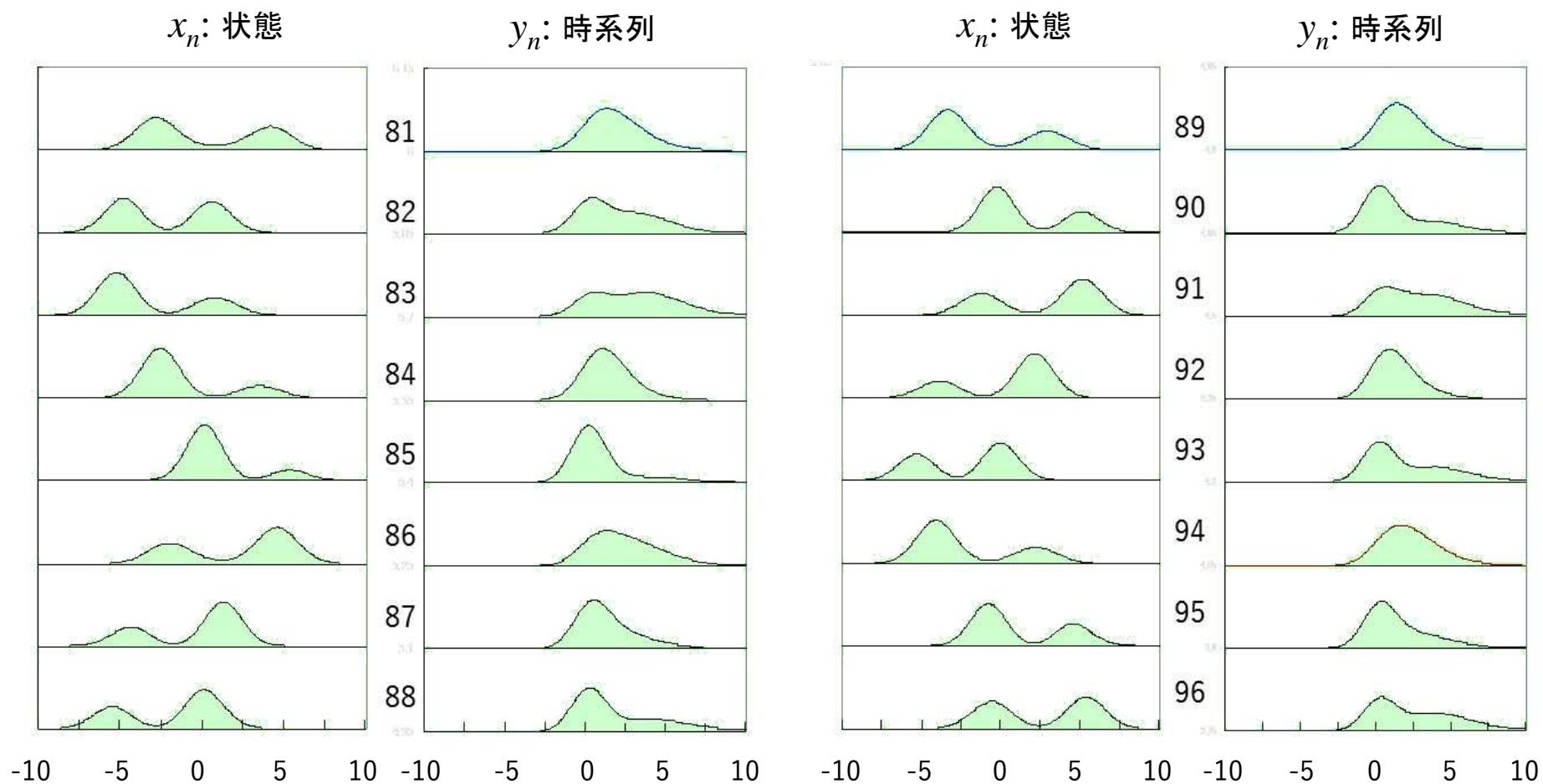


状態推定分布の変化



状態と時系列の予測分布

$t=80$ から予測開始



一般化状態空間モデル

状態空間モデルの本質

$$p(x_n | x_{n-1}, \dots, x_0, y_{n-1}, \dots, y_1) = F(x_n | x_{n-1})$$

$$p(y_n | x_n, \dots, x_0, y_{n-1}, \dots, y_1) = H(y_n | x_n)$$

一般化状態空間モデル

$$x_n \sim F(\cdot | x_{n-1})$$

$$y_n \sim H(\cdot | x_n)$$

観測値 and/or 状態が離散変数でも同様に適用できる

• Hidden Markov Model

状態空間モデルの拡張

線形・ガウス型

$$\begin{aligned}x_n &= Fx_{n-1} + Gv_n \\y_n &= Hx_n + w_n\end{aligned}$$

非線形・非ガウス型

$$\begin{aligned}x_n &= f(x_{n-1}, v_n) \\y_n &= h(x_n, w_n)\end{aligned}$$

一般型

$$\begin{aligned}x_n &\sim F(\cdot | x_{n-1}) \\y_n &\sim H(\cdot | x_n)\end{aligned}$$

関数：非線形
分布：非ガウス型

条件付分布
離散状態・離散観測値

一般化フィルタ・平滑化

一期先予測

$$p(x_n | Y_{n-1}) = \int_{-\infty}^{\infty} p(x_{n-1} | Y_{n-1}) dF(x_n | x_{n-1})$$

フィルタ

$$p(x_n | Y_n) = \frac{H(y_n | x_n) p(x_n | Y_{n-1})}{p(y_n | Y_{n-1})}$$

平滑化

$$p(x_n | Y_N) = p(x_n | Y_n) \int_{-\infty}^{\infty} \frac{p(x_{n+1} | Y_N)}{p(x_{n+1} | Y_n)} dF(x_{n+1} | x_n)$$

非定常離散分布モデル

- 観測モデルの例

$$P(y_n | p_n, m_n) = \binom{m_n}{y_n} p_n^{y_n} (1 - p_n)^{m_n - y_n} \quad \text{二項分布 } (0 \leq p_n \leq 1)$$

$$P(y_n | \lambda_n) = \frac{e^{-\lambda_n} \lambda_n^{y_n}}{y_n!} \quad \text{ポアソン分布 } (\lambda_n \geq 0)$$

- システムモデルの例

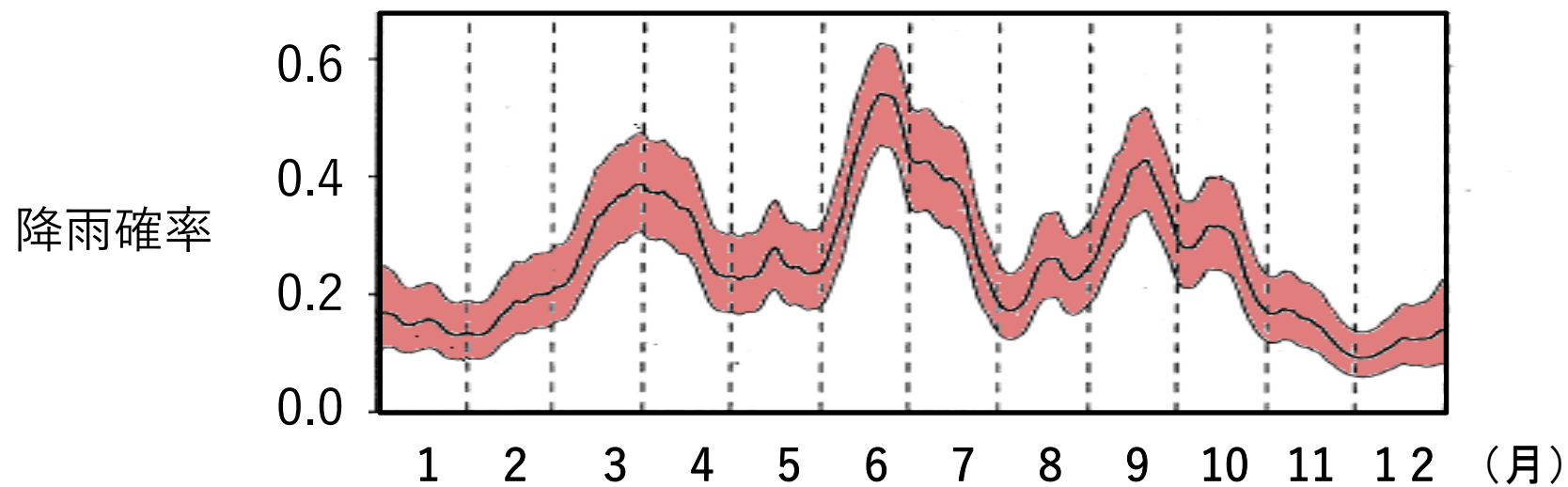
$$x_n = \begin{bmatrix} t_n \\ t_{n-1} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} t_{n-1} \\ t_{n-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} v_n \quad -\infty < t_n < \infty$$

$$p_n = \frac{e^{t_n}}{1 + e^{t_n}} \quad \text{二項分布}$$

$$\lambda_n = e^{t_n} \quad \text{ポアソン分布}$$

降雨データ (東京)

y_n : 2年間のうち雨が降った回数



非線形状態空間モデルのパラメータ推定の問題点

状態空間モデルのパラメータ推定

- (1) フィルタによって対数尤度を計算
- (2) 非線形最適化によって(1)を繰り返し、最尤推定量を求める

- 計算量が大きい
- 対数尤度が誤差を含む (粒子フィルタの場合)

状態とパラメータの同時推定

非線形平滑化 $\{y_1, \dots, y_N\} \Rightarrow z_n$

拡大した状態ベクトル

$$z_n = \begin{bmatrix} x_n \\ \theta \end{bmatrix} \Rightarrow x_n \text{ and } \theta$$

z_n を推定できれば状態とパラメータを同時に推定できる

$$p(x_n | Y_N) = \int p(x_n, \theta | Y_N) d\theta$$

$$p(\theta | Y_N) = \int p(x_n, \theta | Y_N) dx_n$$

z_n に関する状態空間モデル

$$\begin{aligned}x_n &= F(x_{n-1}, v_n) \\ y_n &= H(x_n, w_n)\end{aligned} \quad \theta_n = \theta_{n-1}$$

$$z_n = \begin{bmatrix} x_n \\ \theta_n \end{bmatrix} \quad u_n = \begin{bmatrix} v_n \\ \varepsilon_n \end{bmatrix}$$

z_n に関する状態空間モデル

$$\begin{aligned}z_n &= F^*(z_{n-1}, u_n) \\ y_n &= H^*(z_n, w_n)\end{aligned} \quad F^*(z_{n-1}, v_n) = \begin{bmatrix} F(x_{n-1}, v_n) \\ \theta_{n-1} \end{bmatrix}$$
$$H^*(z_n, w_n) = H(x_n, w_n)$$

状態とパラメータの同時推定

“Although this extended Kalman filter approach appears perfectly straightforward, experiences has shown that with the usual state space model, it does not work well in practice.”

Anderson and Moore (1979) p284

自己組織型フィルタ・平滑化

状態空間モデル

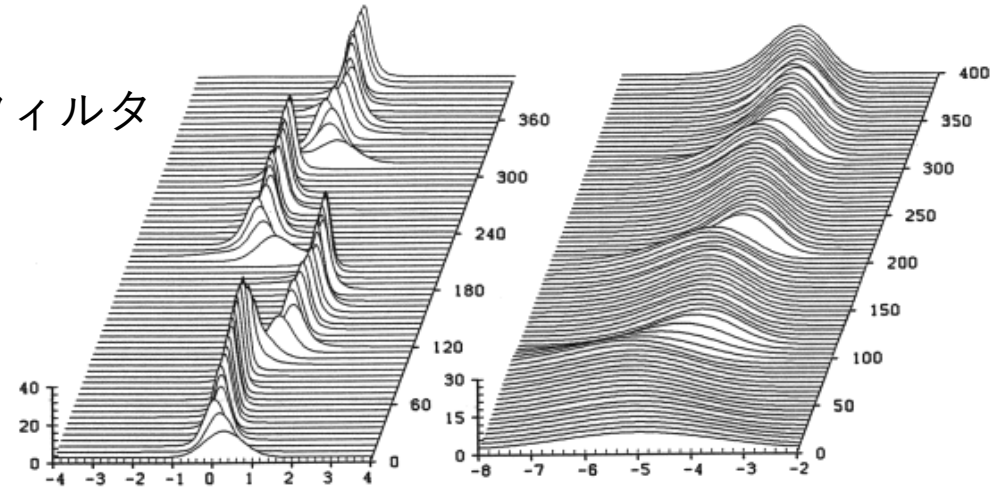
$$x_n = x_{n-1} + v_n$$

$$y_n = x_n + w_n$$

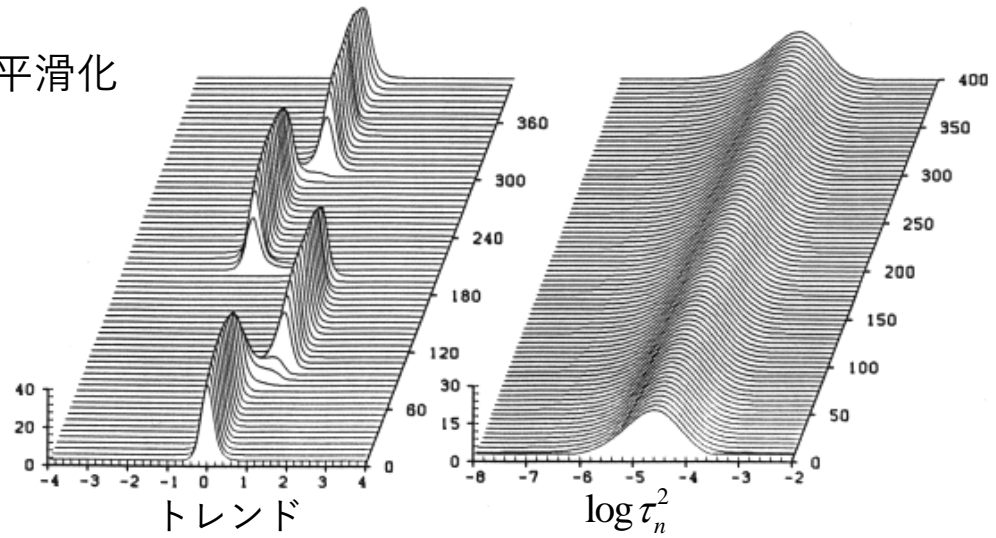
コーシー分布

$$p(v_n) = \frac{\tau}{\pi} \frac{1}{(v_n^2 + \tau^2)}$$

フィルタ



平滑化



拡大した状態空間モデル

$$\begin{bmatrix} x_n \\ \log \tau_n^2 \end{bmatrix} = \begin{bmatrix} x_{n-1} \\ \log \tau_{n-1}^2 \end{bmatrix} + \begin{bmatrix} \tau_{n-1} \\ 0 \end{bmatrix} v_n$$

$$y_n = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_n \\ \log \tau_n^2 \end{bmatrix} + w_n$$

自己組織型フィルタ・平滑化

状態空間モデル

$$x_n = x_{n-1} + v_n$$

$$y_n = x_n + w_n$$

$$p(v; \tau^2, b) = \frac{\Gamma(b)\tau^{2b-1}}{\Gamma(1/2)\Gamma(b-1/2)} \frac{1}{(v^2 + \tau^2)^b}$$

拡大した状態空間モデル

$$\begin{bmatrix} x_n \\ \theta_{n,1} \\ \theta_{n,2} \end{bmatrix} = \begin{bmatrix} x_{n-1} \\ \theta_{n-1,1} \\ \theta_{n-1,2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} v_n$$

$$y_n = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_n \\ \theta_{n,1} \\ \theta_{n,2} \end{bmatrix} + w_n$$

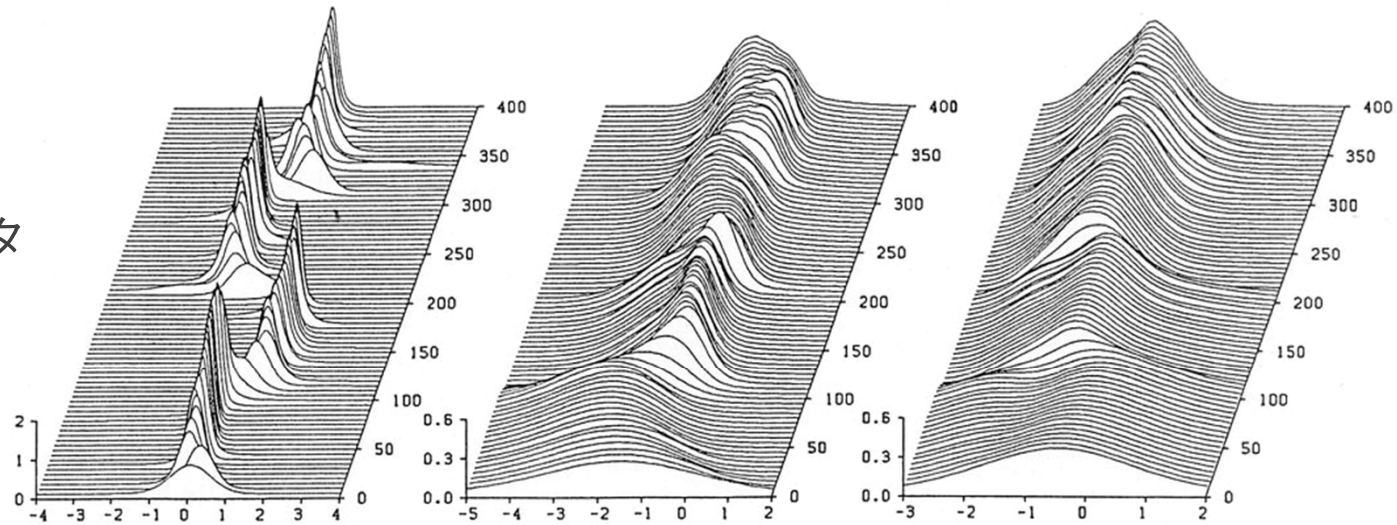
$$\theta_{n,1} = \log \tau_n^2 - 3\theta_{n,2}$$

$$\theta_{n,2} = \log(b_n - 1/2)$$

最適な τ^2 は b によって異なる

自己組織型フィルタ・平滑化

フィルタ

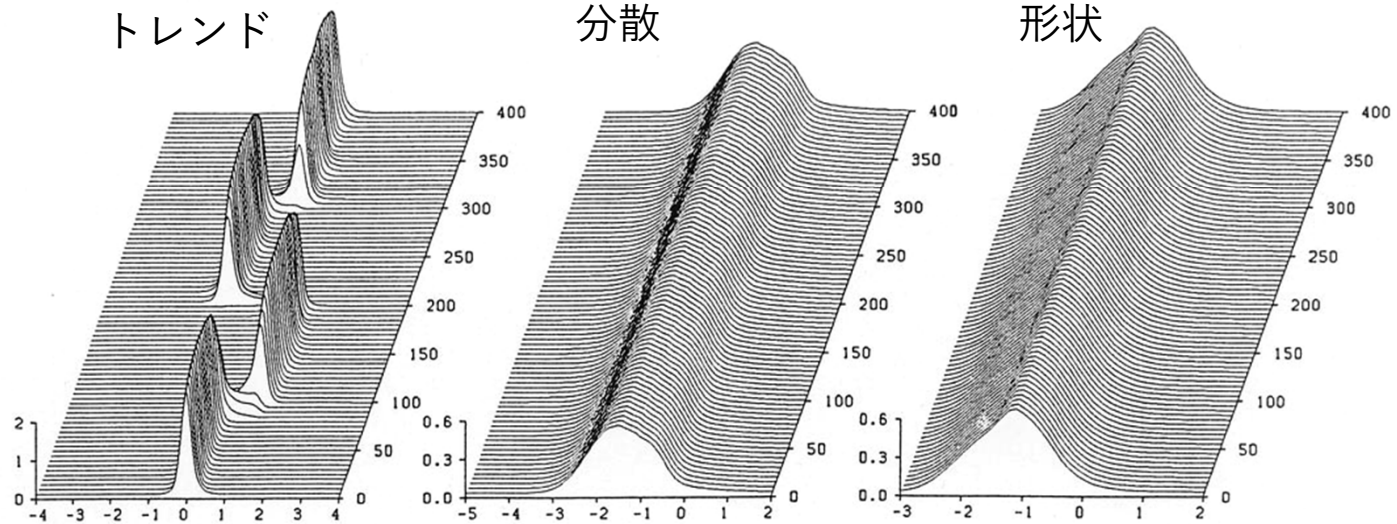


トレンド

分散

形状

平滑化



時変パラメータへの拡張

$$\theta \quad \Rightarrow \quad \theta_n$$

$$z_n = \begin{bmatrix} x_n \\ \theta_n \end{bmatrix}$$

パラメータの時間変化のモデル

$$\theta_n = \theta_{n-1} + \zeta_n, \quad \zeta_n \sim N(0, \xi^2)$$

- 高次元になる場合，計算は粒子フィルタで
- 粒子フィルタにはシステムノイズが必要

非ガウス型フィルタの長所・欠点

- 数値計算誤差を除き，厳密解が得られる
- 高次元の状態に適用できない

$$\text{計算量} \sim (k_1 \times \dots \times k_m)(k_{q_1} + \dots + k_{q_\ell})n \sim \ell k^{m+1} n$$

m 状態の次元

k_i 状態の分点数

ℓ システムノイズの次元

k_{q_j} システムノイズの分点数

| k | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----|------|-----|------|------|----|------|------|
| 10 | 100 | 1K | 10K | 100K | 1M | 10M | 100M |
| 30 | 1K | 30K | 1M | 30M | 1G | 30G | 1T |
| 100 | 10K | 1M | 100M | 10G | 1T | 100T | 10P |
| 300 | 100K | 30M | 10T | 3P | – | – | – |

分布の近似

0. 線形・正規モデル近似

カルマンフィルタ・平滑化

1. 正規分布近似

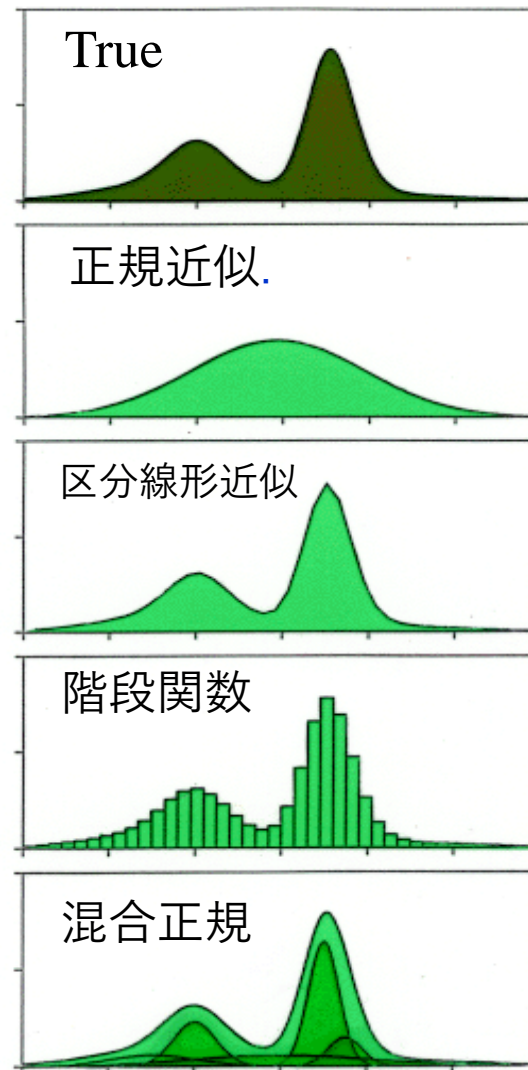
拡張カルマンフィルタ・平滑化

2. 区分線形（階段）近似

非ガウス型フィルタ・平滑化

3. 混合正規分布近似

ガウス和フィルタ・平滑化



ガウス和フィルタ・平滑化

モデル

$$x_n = F_n x_{n-1} + G_n v_n$$

$$y_n = H_n x_n + w_n$$

ノイズ分布

$$p(v_n) \sim \sum_{i=1}^q \alpha_i \varphi_i(v_n)$$

$$p(w_n) \sim \sum_{j=1}^r \beta_j \varphi_j(w_n)$$

予測分布とフィルタ分布

$$p(x_n | Y_{n-1}) \sim \sum_{k=1}^{m_n} \gamma_k \varphi_k(x_n | Y_{n-1})$$

$$p(x_n | Y_n) \sim \sum_{\ell=1}^{\ell_n} \delta_\ell \varphi_\ell(x_n | Y_n)$$

$$\varphi_k(x_n | Y_{n-1}) \equiv N(x_{n|n-1}^{(k)}, V_{n|n-1}^{(k)}),$$

$$\varphi_\ell(x_n | Y_n) \equiv N(x_{n|n}^{(\ell)}, V_{n|n}^{(\ell)})$$

ガウス和フィルタ

$$Y_n \triangleq \{y_1, \dots, y_n\}$$

一期先予測

$$\begin{aligned} p(x_n | Y_{n-1}) &= \sum_{i=1}^q \sum_{k=1}^{\ell_{n-1}} \alpha_i \gamma_{k,n-1} \varphi(x_n | x_{n|n-1}^{ik}, V_{n|n-1}^{ik}) \\ &\equiv \sum_{j=1}^{m_n} \delta_{jn} \varphi(x_n | x_{n|n-1}^j, V_{n|n-1}^j) \\ m_n &= q\ell_{n-1}, \quad \delta_{jn} = \alpha_i \gamma_{k,n-1} \end{aligned}$$

フィルタ

$$\begin{aligned} p(x_n | Y_n) &= \sum_{j=1}^r \sum_{k=1}^{m_n} \gamma_{jk,n} \varphi(x_n | x_{n|n}^{jk}, V_{n|n}^{jk}) \\ &\equiv \sum_{i=1}^{\ell_n} \gamma_{i,n} \varphi(x_n | x_{n|n}^i, V_{n|n}^i) \\ \ell_n &= rm_n, \quad \gamma_{jk,n} = \beta_j \delta_{k,n} \varphi(y_n | x_{n|n-1}^{jk}, V_{n|n-1}^{jk}) \end{aligned}$$

Alspach and Sorenson (1972). Nonlinear Bayesian estimation using Gaussian sum approximations, IEEE Trans. AC, AC-17, 439-448.

Kitagawa (1989). Non-Gaussian seasonal Adjustment. *Comp. & Math. with Appl*, 18(6-7), 503-514.

Kitagawa (1994). The two-filter formula for smoothing and implementation of the Gaussian-sum smoother. *AISM*, 46(4), 605-623.

ガウス和フィルタの長所

ガウス和フィルタのパラメータ $\delta_{i,n}, x_{n|n-1}^i, V_{n|n-1}^i, \gamma_{k,n}, x_{n|n}^k, V_{n|n}^k$ はカルマンフィルタを並列に実行することによって計算できる。

並列カルマンフィルタ

予測

$$x_{n|n-1}^{ik} = F_n x_{n-1|n-1}^k$$

$$V_{n|n-1}^{ik} = F_n V_{n-1|n-1}^k F_n^T + G_n Q_n^i G_n^T$$

フィルタ

$$K_n^{jk} = V_{n|n-1}^k H_n^T (H_n V_{n|n-1}^k H_n^T + R_j)^{-1}$$

$$x_{n|n}^{jk} = x_{n|n-1}^k + K_n^{jk} (y_n - H_n x_{n|n-1}^k)$$

$$V_{n|n}^{jk} = (I - K_n^{jk} H_n) V_{n|n-1}^k$$

ガウス和フィルタ

予測

$$p(x_n | Y_{n-1}) = \sum_{i=1}^q \sum_{k=1}^{\ell_{n-1}} \alpha_i \gamma_{k,n-1} \varphi(x_n | x_{n|n-1}^{ik}, V_{n|n-1}^{ik})$$

フィルタ

$$p(x_n | Y_n) = \sum_{j=1}^r \sum_{k=1}^{m_n} \gamma_{jk,n} \varphi(x_n | x_{n|n}^{jk}, V_{n|n}^{jk})$$

2方向フィルタによる平滑化

フィルタ

$$\begin{aligned} p(x_n | Y_n) &= p(x_n | Y_{n-1}, y_n) \\ &\propto p(x_n, y_n | Y_{n-1}) \\ &= p(y_n | x_n) p(x_n | Y_{n-1}) \end{aligned}$$

平滑化

$$\begin{aligned} p(x_n | Y_N) &= p(x_n | Y_{n-1}, Y_N^n) \\ &\propto p(x_n, Y_N^n | Y_{n-1}) \\ &= p(Y_N^n | x_n) p(x_n | Y_{n-1}) \end{aligned}$$

$p(Y_N^n | x_n)$ は後ろ向きフィルタによって求められる

$$\begin{aligned} Y_n &= \{y_1, \dots, y_n\} \\ Y_N^n &= \{y_n, \dots, y_N\} \end{aligned}$$

$$\begin{aligned} p(x_n | Y_{n-1}) &= \sum_{j=1}^{m_n} \delta_{j,n} \varphi(x_n | x_{n|n-1}^j, V_{n|n-1}^j) \\ p(Y_N^n | x_n) &= \sum_{i=1}^{\ell_n} \gamma_{in} \varphi(x_n | z_{n|n}^i, U_{n|n}^i) \end{aligned}$$



$$p(x_n | Y_N) = \sum_{i=1}^{\ell_n} \sum_{j=1}^{m_n} \gamma_{in} \delta_{jn} \varphi(x_n | x_{n|n-1}^{ij}, V_{n|n-1}^{ij})$$

$$J_n^{ij} = V_{n|n-1}^j H_n^T (H_n V_{n|n-1}^j H_n^T + U_k)^{-1}$$

$$x_{n|n}^{ij} = x_{n|n-1}^j + J_n^{ij} (z_{n|n}^i - x_{n|n-1}^j)$$

$$V_{n|n}^{ij} = (I - J_n^{ij}) V_{n|n-1}^j$$

ガウス和フィルタの問題点

ガウス成分数が時刻とともに爆発的に増加する

$$m_n = qr \times m_{n-1} = (qr)^n \rightarrow \infty \quad \text{as } n \rightarrow \infty$$

- ガウス成分数の削減が不可欠
- 解決法

各時刻 n で $m_n \Rightarrow m$ と削減する (m は適当な整数)

ガウス成分数の縮減

$$p(x_n | Y_{n-1}) \sim \sum_{k=1}^{m_n} \gamma_k \varphi_k(x_n | Y_{n-1}) \Rightarrow \sum_{k=1}^m \gamma'_k \varphi'_k(x_n | Y_{n-1})$$

$$p(x_n | Y_n) \sim \sum_{\ell=1}^{\ell_n} \delta_\ell \varphi_\ell(x_n | Y_n) \Rightarrow \sum_{\ell=1}^m \delta'_\ell \varphi'_\ell(x_n | Y_n)$$

原理的にはKL-情報量最大化

$$I(p; q) \int \log \frac{p(x)}{q(x)} p(x) dx$$

現実的には分布数の逐次削減

$$D(\varphi_k, \varphi_m) = -2\delta_k \delta_m \{I(\varphi_k; \varphi_m) + I(\varphi_m; \varphi_k)\}$$

$$= \delta_k \delta_m \{V_k^{-1} V_m + V_m^{-1} V_k + (\mu_k - \mu_m)^T (V_k^{-1} + V_m^{-1}) (\mu_k - \mu_m)\}$$

$$\mu_{km} = \frac{\delta_k \mu_k + \delta_m \mu_m}{\delta_k + \delta_m}$$

$$V_{km} = \frac{\delta_k \{V_k + (\mu_{km} - \mu_k)(\mu_{km} - \mu_k)^T\} + \delta_m \{V_m + (\mu_{km} - \mu_m)(\mu_{km} - \mu_m)^T\}}{\delta_k + \delta_m}$$

ガウス成分数の削減

- 原理的には**KL-divergence**の最小化によって最適な近似が得られる

$$I(g(x); f(x)) = E \left[\log \left\{ \frac{g(x)}{f(x)} \right\} \right] = \int \log \left\{ \frac{g(x)}{f(x)} \right\} g(x) dx$$

$$g(x) = \sum_{i=1}^m \alpha_i \varphi(x | \mu_i, V_i) \quad f(x) = \sum_{j=1}^{\ell} \beta_j \varphi(x | \xi_j, W_j)$$

実際にはKL-divergence最小化は困難

1. 解析的表現がない： 数値積分が必要
2. 多数のパラメータが存在： 数値的最適化が必要

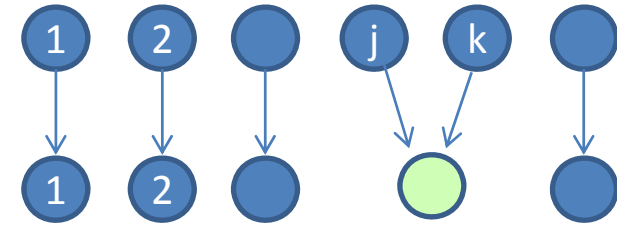
$$\beta_j, \xi_j, W_j \quad (j=1, \dots, \ell)$$

逐次併合法

$$g(x) = \sum_{i=1}^m w_i \varphi(x | \mu_i, \sigma_i^2)$$

$$f_\ell(x) = \sum_{i=1}^{\ell} \alpha_i \varphi(x | \xi_i, \tau_i^2)$$

$$h_{jk}(x) = \sum_{i \notin \{j,k\}} \alpha_i \varphi(x | \xi_i, \tau_i^2) + (\alpha_j + \alpha_k) \varphi(x | \xi_{jk}, \tau_{jk}^2)$$



Moment preserving pooling

$$\xi_{jk} = (\alpha_j + \alpha_k)^{-1} \{ \alpha_j \xi_j + \alpha_k \xi_k \}$$

$$\tau_{jk}^2 = (\alpha_j + \alpha_k)^{-1} \left[\alpha_j \{ \tau_j^2 + (\xi_j - \xi_{jk})^2 \} + \alpha_k \{ \tau_k^2 + (\xi_k - \xi_{jk})^2 \} \right]$$

$\min_{j,k} I(g(x); h_{jk}(x))$ を最小にする j と k を選択 ($I(g;h)$ は適当な評価基準)

$$g(x) \Rightarrow f_{m-1}(x) \Rightarrow f_{m-2}(x) \Rightarrow \cdots \Rightarrow f_\ell(x)$$

削減法の比較

- Ad hoc. Criteria

- Kitagawa(1989,1994): 併合する二つの分布の重み付き KL-divergence

$$D(\varphi_k, \varphi_j) = \beta_k \beta_j \left\{ V_k^{-1} V_j + V_j^{-1} V_k + (\mu_k - \mu_j)^T (V_k^{-1} + V_j^{-1}) (\mu_k - \mu_j) \right\}$$

- Runnalls(2006) : 併合によるKL-divergenceの増加の上限

$$B(\varphi_k, \varphi_j) = \frac{1}{2} \left\{ (\beta_k + \beta_j) \log \det(V_{kj}) - \beta_k \log \det(V_k) - \beta_j \log \det(V_j) \right\}$$

- 併合した分布のKL-divergence (数値積分による)

$$I(g(x); f_{jk}(x)) = \int \log g(x) g(x) dx - \int \log f_{jk}(x) g(x) dx$$

$f_{jk}(x)$ のパラメータはモーメント保存法による

- KL-divergence 最小化 (数値積分と数値的最適化による)

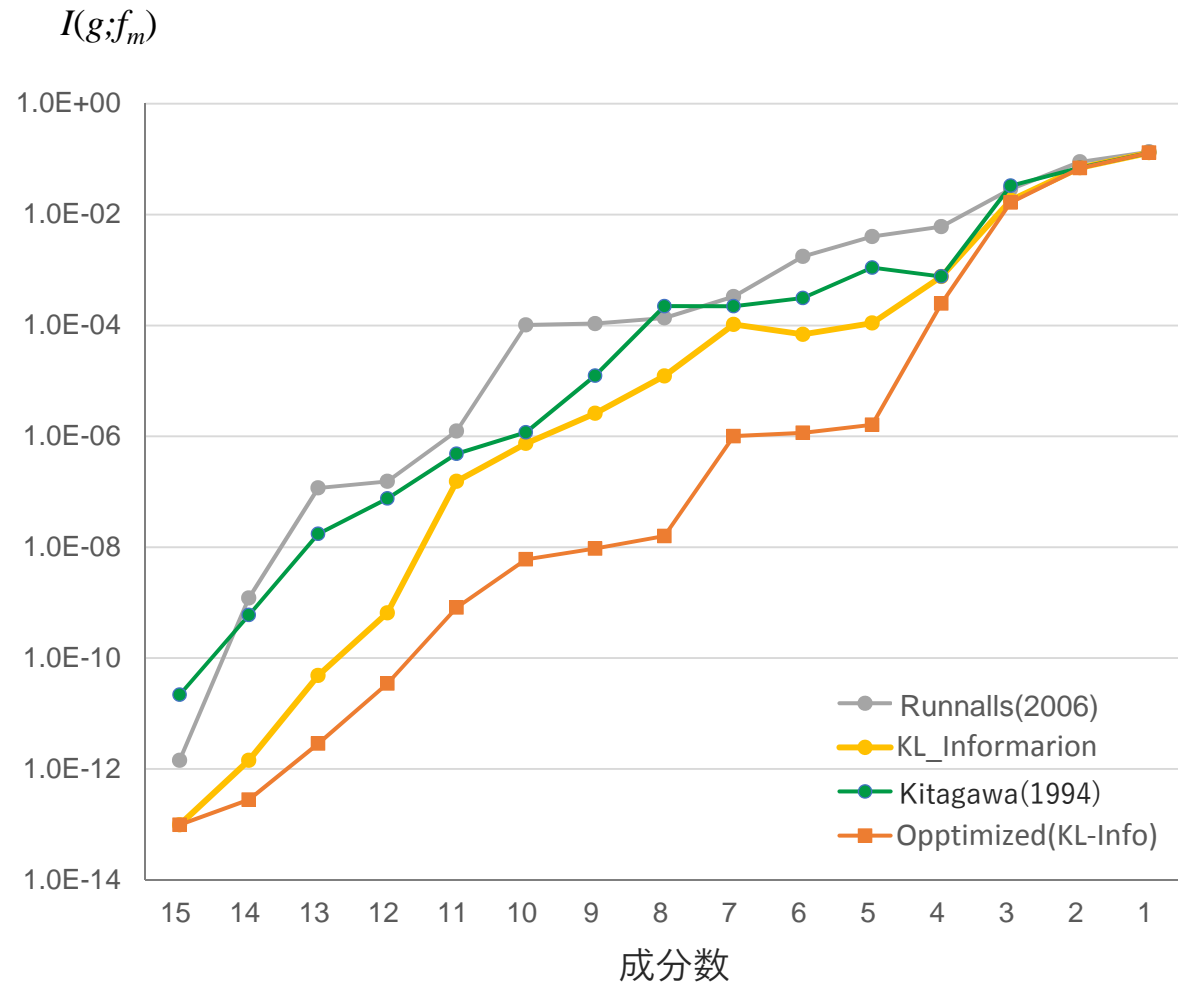
$$I(g(x); \hat{f}_{jk}(x)) = \int \log g(x) g(x) dx - \int \log \hat{f}_{jk}(x) g(x) dx$$

$\hat{f}_{jk}(x)$ のパラメータは $I(g; f_{jk})$ の最小化により推定

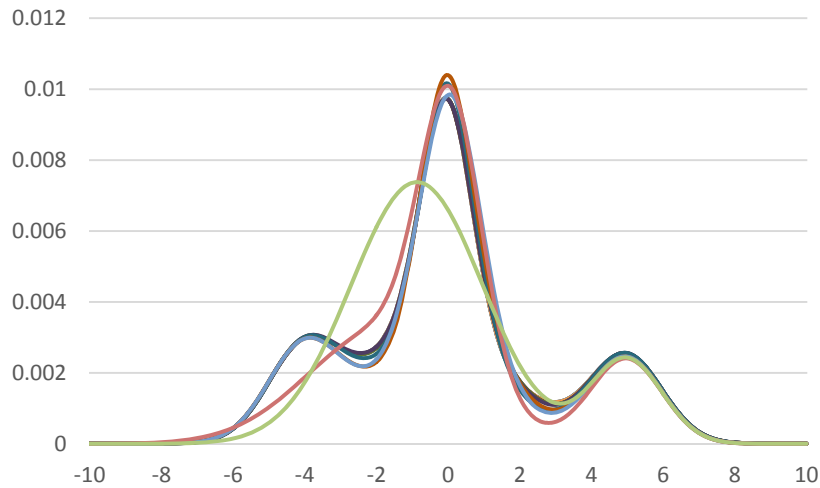
併合アルゴリズムの比較

逐次併合によるKL-divergence
の増加

| m | j | k | $I(g;f_m)$ |
|-----|-----|-----|------------|
| 15 | 10 | 15 | 9.800E-14 |
| 14 | 9 | 12 | 1.431E-12 |
| 13 | 9 | 13 | 4.178E-11 |
| 12 | 9 | 13 | 5.638E-10 |
| 11 | 10 | 11 | 1.376E-07 |
| 10 | 9 | 10 | 5.064E-07 |
| 9 | 7 | 8 | 1.535E-06 |
| 8 | 5 | 7 | 6.298E-06 |
| 7 | 4 | 6 | 8.464E-05 |
| 6 | 5 | 7 | 1.206E-05 |
| 5 | 4 | 6 | 8.587E-05 |
| 4 | 4 | 5 | 0.0005982 |
| 3 | 3 | 4 | 0.0185856 |
| 2 | 1 | 3 | 0.0720529 |
| 1 | 1 | 2 | 0.1304686 |

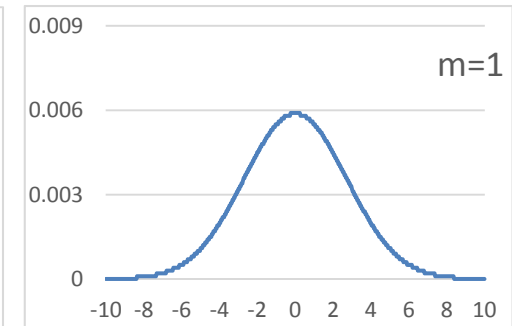
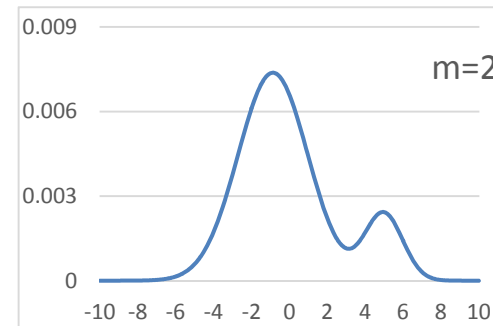
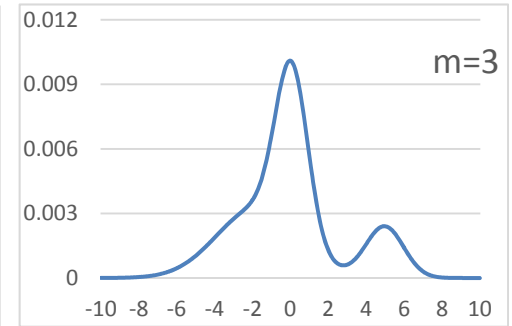
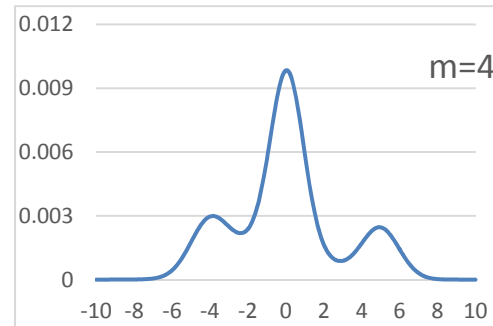
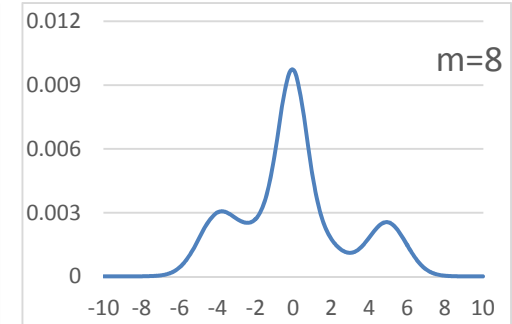
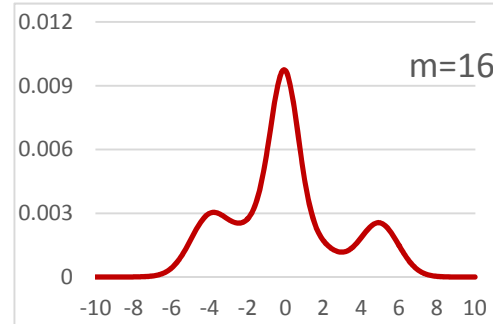


ガウス成分の逐次併合

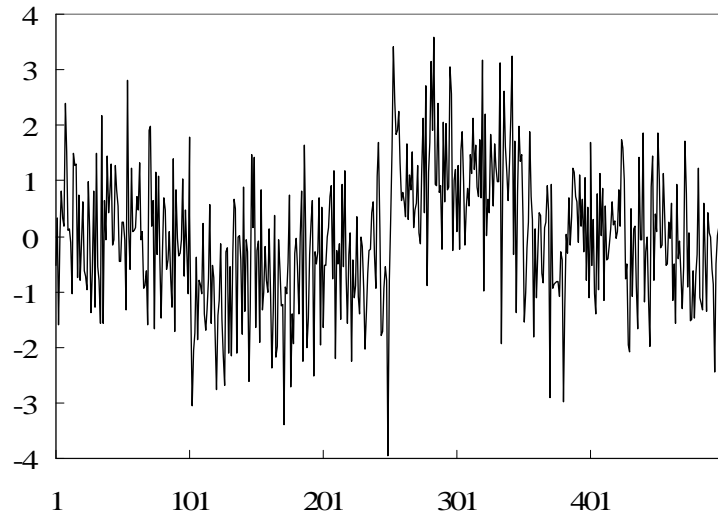


真のモデル(m=16)

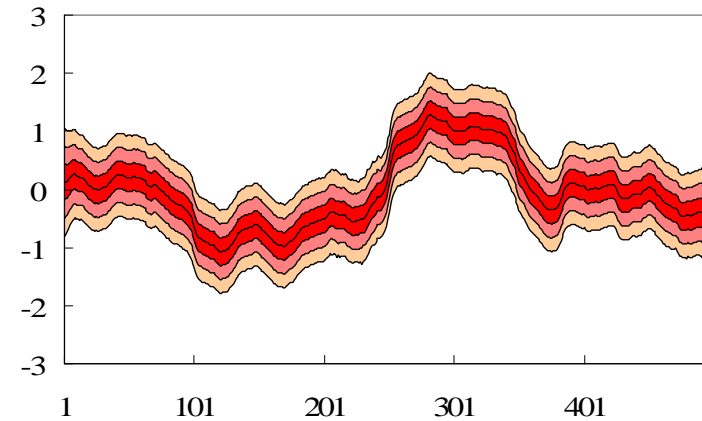
$$g(x) = \sum_{i=1}^m w_i \varphi(x | \mu_i, \sigma_i^2)$$



非ガウス型平滑化



カルマンフィルタ・平滑化



トレンドモデル

$$t_n = t_{n-1} + v_n$$

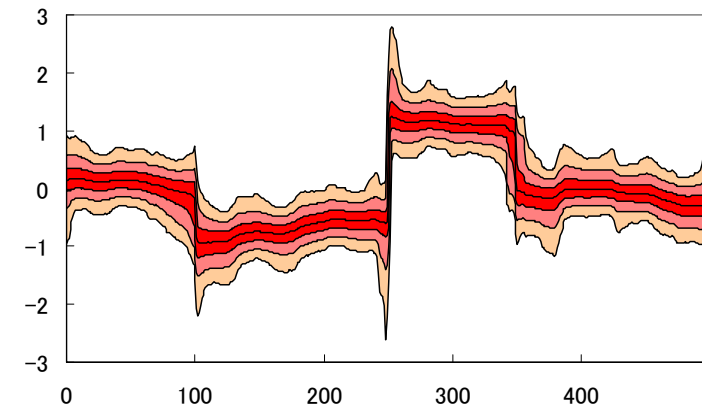
$$y_n = t_n + w_n$$

ノイズ分布

$$v_n \sim N(0, \tau^2) \text{ or } C(0, \tau^2)$$

$$w_n \sim N(0, \sigma^2)$$

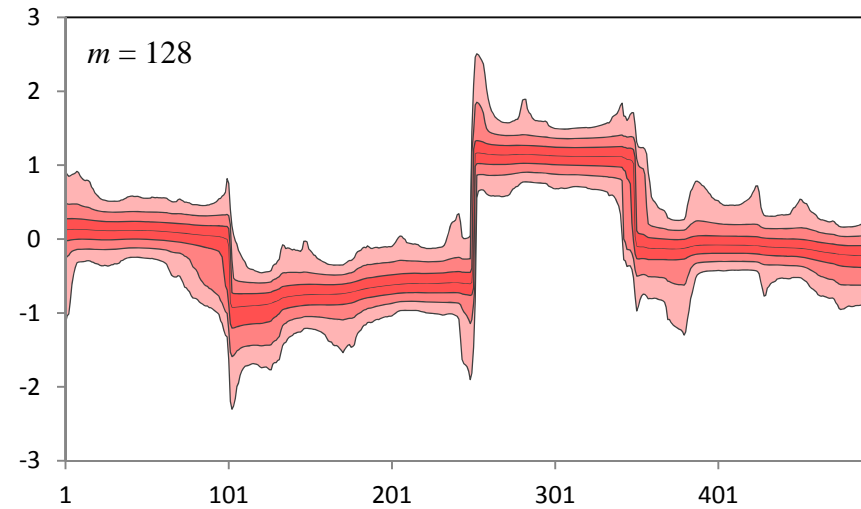
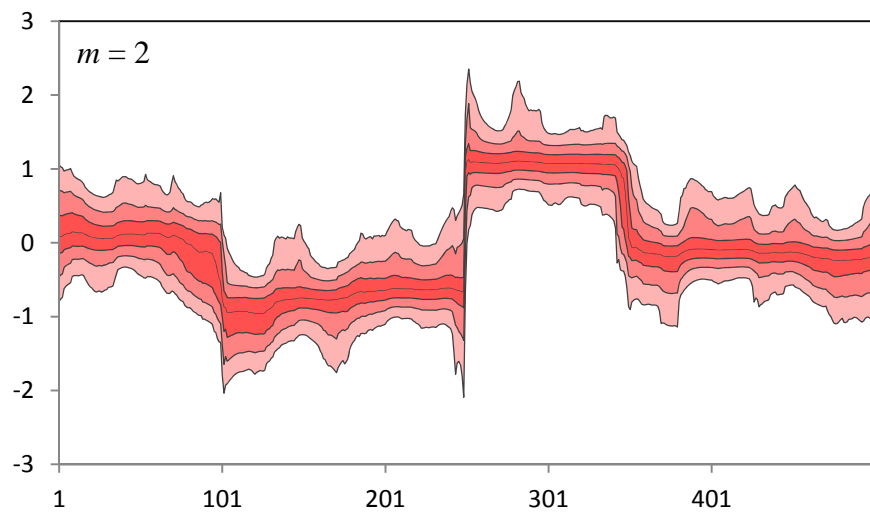
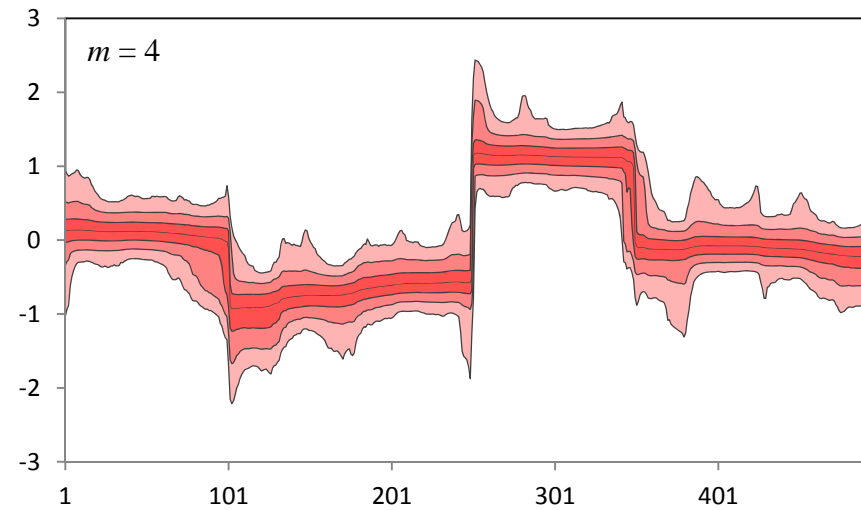
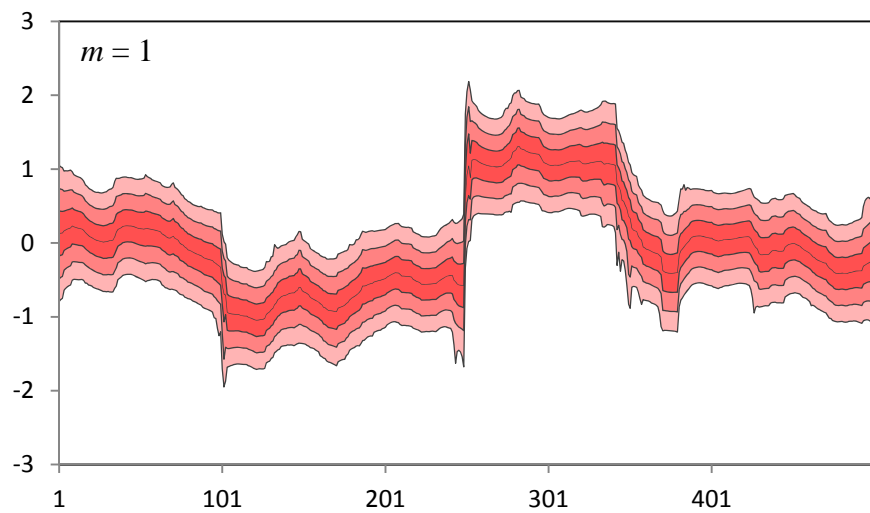
非ガウス型フィルタ・平滑化



ガウス和フィルタ・平滑化

| m | Log-likelihood | cpu time (秒) | |
|-----|----------------|--------------|---------|
| | | フィルタ | 平滑化 |
| 1 | -741.930 | 0.00 | 0.08 |
| 2 | -741.047 | 0.02 | 0.23 |
| 4 | -740.816 | 0.02 | 0.94 |
| 8 | -740.748 | 0.05 | 3.70 |
| 16 | -740.702 | 0.27 | 14.85 |
| 32 | -740.704 | 1.86 | 59.53 |
| 64 | -740.704 | 14.26 | 243.47 |
| 128 | -740.704 | 112.51 | 1018.20 |

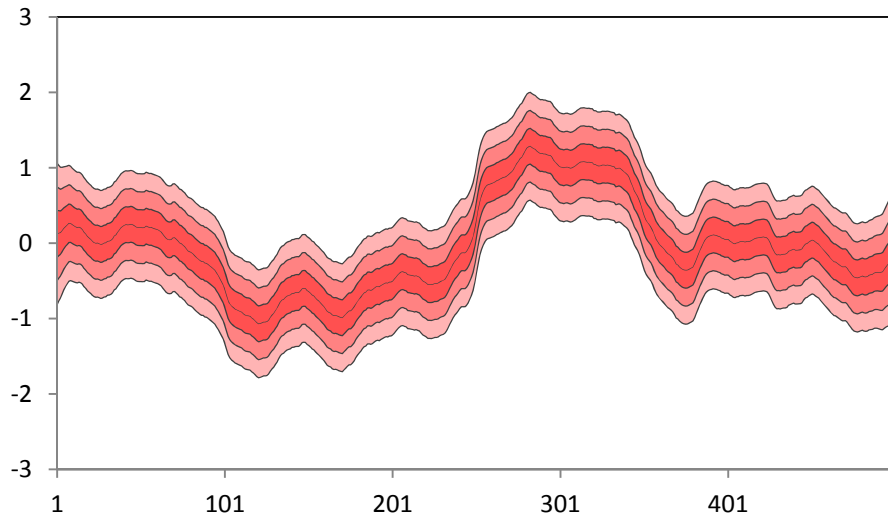
ガウス和平滑化 (m は成分数)



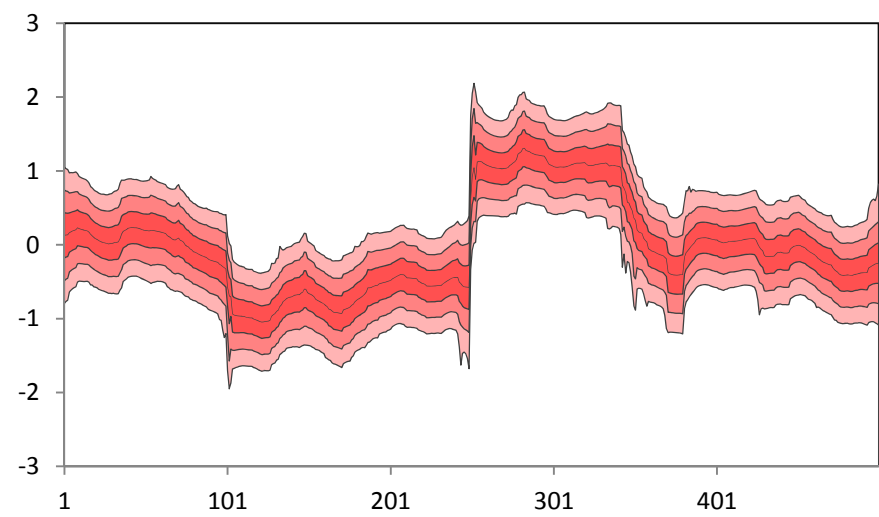
注意： $m=1$ としたガウス和平滑化

$m=1$ の場合でもガウス和平滑化はカルマン平滑化と異なる。

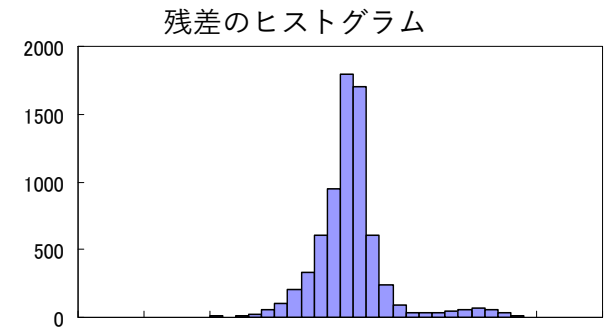
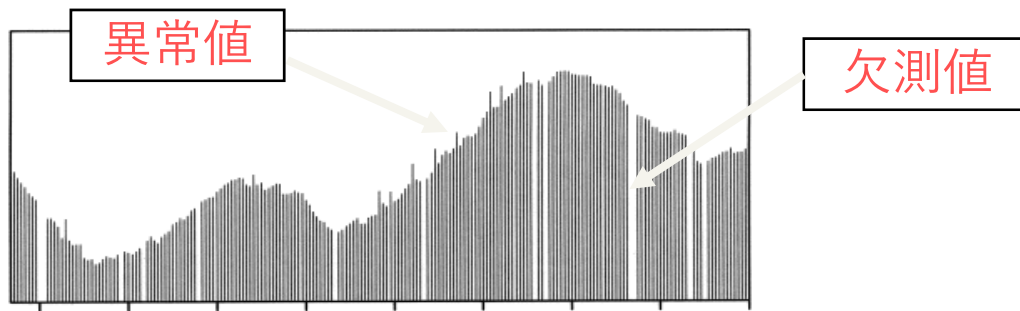
Kalman Smoother



Gaussian-sum Smoother with $m=1$



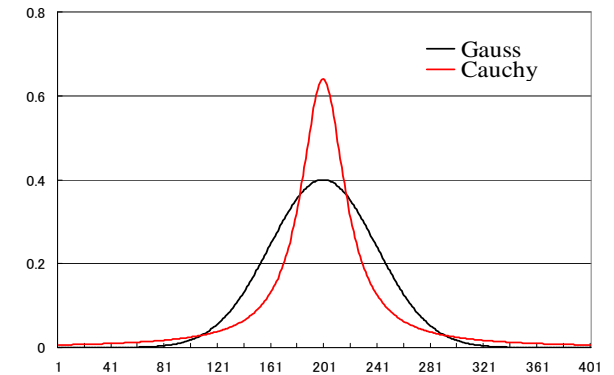
欠測値と異常値の処理



状態空間モデル

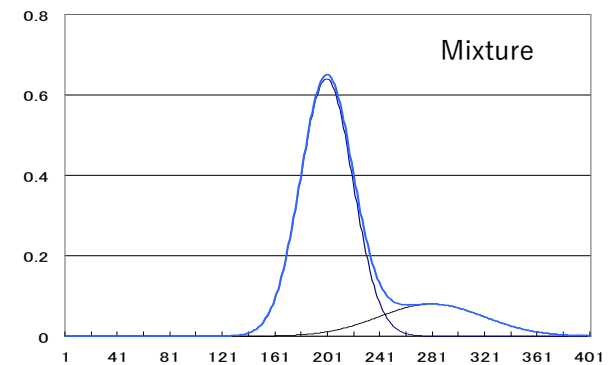
$$\begin{aligned}
 t_n &= t_{n-1} + v_n \\
 y_n &= t_n + w_n
 \end{aligned}$$

y_n : 時系列
 t_n : 信号
 v_n : システムノイズ
 w_n : 観測ノイズ

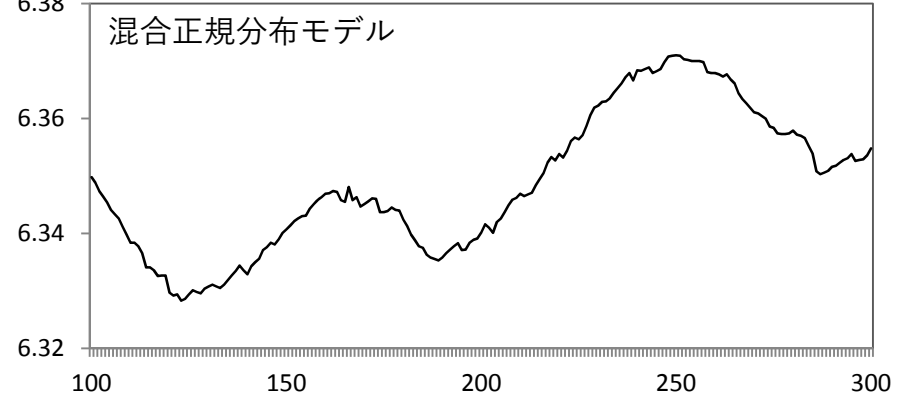
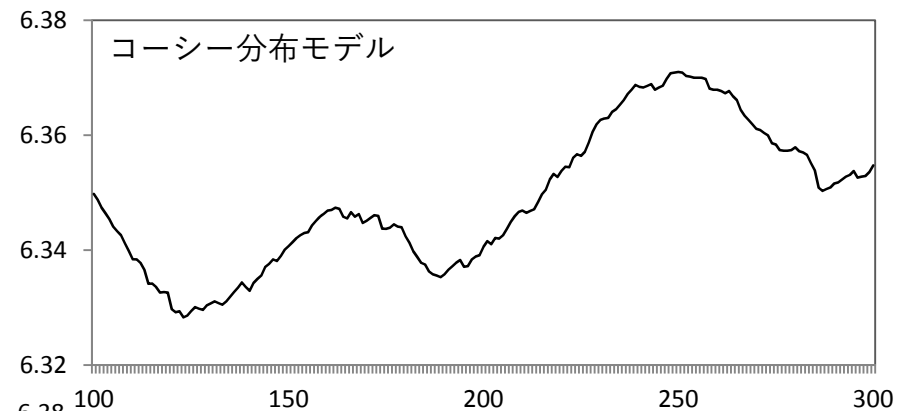
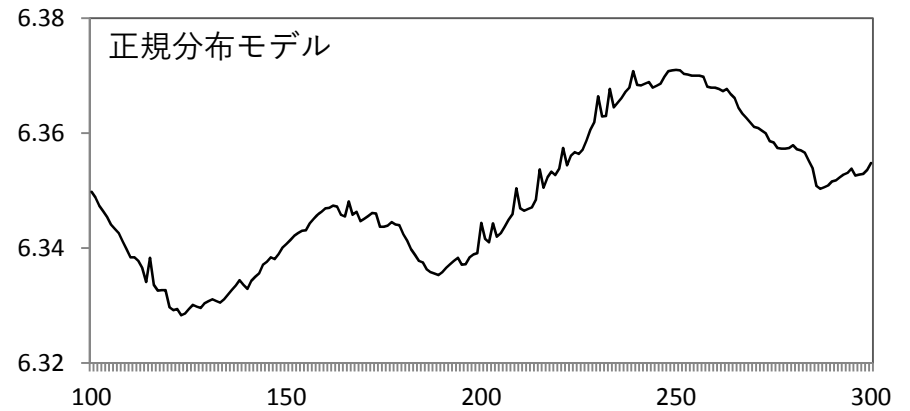
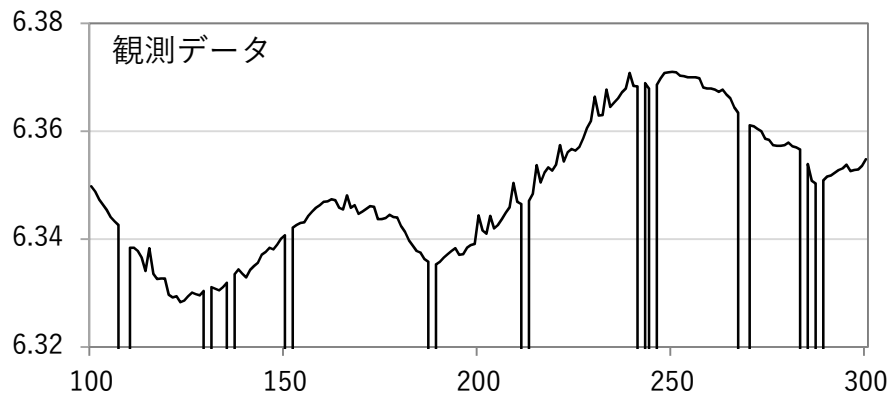


観測ノイズのモデル (非ガウス型モデル)

$$\begin{aligned}
 \text{Gauss} & \quad r(w) \sim N(0, \sigma^2) \\
 \text{Cauchy} & \quad r(w) = \frac{\sigma}{\pi(w^2 + \sigma^2)} \\
 \text{Mixture} & \quad r(w) \sim (1-\alpha)N(0, \sigma^2) + \alpha N(\mu, \zeta^2)
 \end{aligned}$$



観測ノイズモデルによる違い

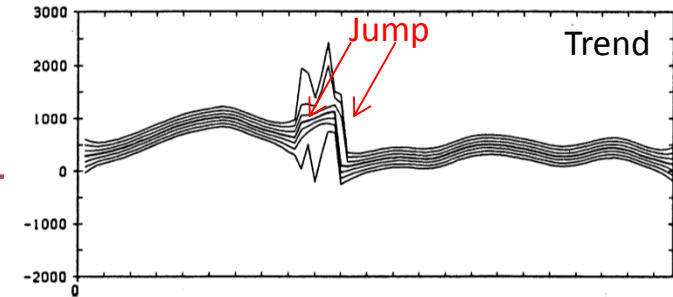
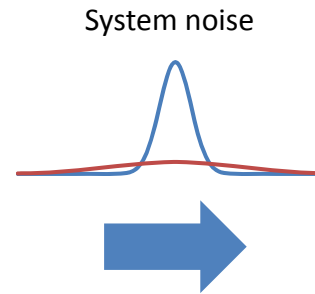
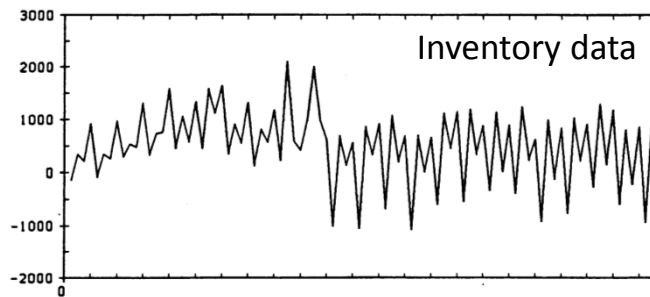


| Model | AIC |
|---------|---------|
| Gauss | -1827.9 |
| Cauchy | -1881.7 |
| Mixture | -1917.4 |

```
# 榛原 地下水位データ
haibara <- as.ts(read.csv("haibara_new.csv"))
haibara_water <- haibara[,2]
# 一部分の詳細表示
plot(haibara_water[211:280],type="h",ylim=c(6.35,6.38))
```

季節調整モデル

● 突然の構造変化



$$x_n = Fx_{n-1} + Gv_n$$

$$y_n = Hx_n + w_n$$

$$t_n = 2t_{n-1} - t_{n-2} + \varepsilon_n$$

$$s_n = -(s_{n-1} + \dots + s_{n-p+1}) + \delta_n$$

$$x_n = \begin{bmatrix} t_n \\ t_{n-1} \\ s_n \\ s_{n-1} \\ \vdots \\ s_{n-p+1} \end{bmatrix}$$

$$\varepsilon_n \sim \sum_{i=1}^m \alpha_i N(\mu_i, \tau_i^2)$$

$$\delta_n = \sum_{i=1}^{\ell} \beta_i N(\xi_i, \lambda_i^2)$$