Kalman Filter and Smoother

Dr. Jarad Niemi

STAT 6150 - Iowa State University

October 7, 2025

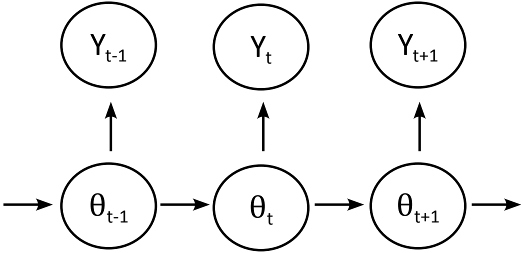
Structure

$$
\begin{aligned}
Y_t &= F_t\theta_t + v_t & v_t &\stackrel{ind}{\sim} N_m(0, V_t) \\
\theta_t &= G_t\theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_p(0, W_t) \\
& & \theta_0 &\sim N_p(m_0, C_0)
\end{aligned}
$$

Structure

$$\begin{aligned} Y_t &= F_t\theta_t + v_t & v_t &\stackrel{ind}{\sim} N_m(0, V_t) \\ \theta_t &= G_t\theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_p(0, W_t) \\ & & \theta_0 &\sim N_p(m_0, C_0) \end{aligned}$$

where v_t and w_t are independent across time and all are independent of θ_0.



Local level model

$$
\begin{aligned}
Y_t &= \theta_t + v_t & v_t &\stackrel{ind}{\sim} N_1(0, V) \\
\theta_t &= \theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_1(0, W) \\
& & \theta_0 &\sim N_1(m_0, C_0)
\end{aligned}
$$

Local level model

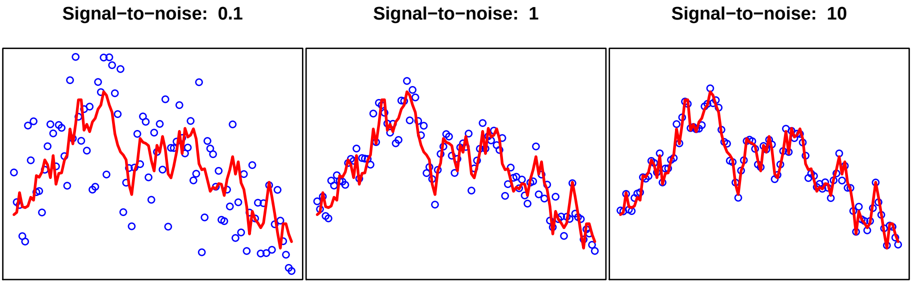
$$\begin{aligned} Y_t &= \theta_t + v_t & v_t &\stackrel{ind}{\sim} N_1(0, V) \\ \theta_t &= \theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_1(0, W) \\ & & \theta_0 &\sim N_1(m_0, C_0) \end{aligned}$$

Signal-to-noise, $r = W/V$.

Local level model

$$\begin{aligned} Y_t &= \theta_t + v_t & v_t &\stackrel{ind}{\sim} N_1(0,V) \\ \theta_t &= \theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_1(0,W) \\ & & \theta_0 &\sim N_1(m_0, C_0) \end{aligned}$$

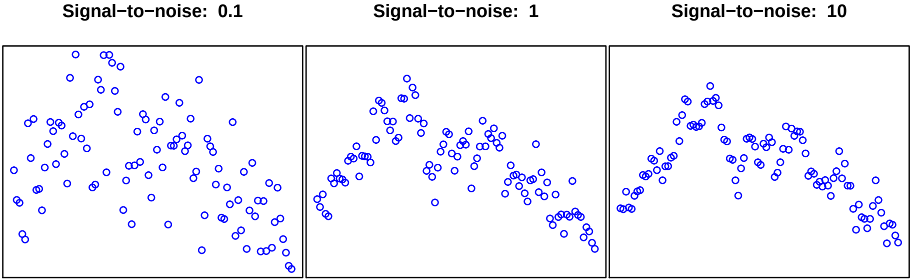
Signal-to-noise, $r = W/V$.



Local level model

$$\begin{aligned} Y_t &= \theta_t + v_t & v_t &\stackrel{ind}{\sim} N(0,V) \\ \theta_t &= \theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N(0,W) \\ & & \theta_0 &\sim N(m_0, C_0) \end{aligned}$$

Signal-to-noise, $r = W/V$.



Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$
- Prior for θ_t

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$
- Prior for θ_t

$$p(\theta_t|y_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}$$

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$
- Prior for θ_t

$$p(\theta_t|y_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}$$

- One-step ahead predictive distribution for y_t

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$
- Prior for θ_t

$$p(\theta_t|y_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}$$

- One-step ahead predictive distribution for y_t

$$p(y_t|y_{1:t-1}) = \int p(y_t|\theta_t)p(\theta_t|y_{1:t-1})d\theta_t$$

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$
- Prior for θ_t

$$p(\theta_t|y_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}$$

- One-step ahead predictive distribution for y_t

$$p(y_t|y_{1:t-1}) = \int p(y_t|\theta_t)p(\theta_t|y_{1:t-1})d\theta_t$$

- Filtered distribution for θ_t

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$
- Prior for θ_t

$$p(\theta_t|y_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}$$

- One-step ahead predictive distribution for y_t

$$p(y_t|y_{1:t-1}) = \int p(y_t|\theta_t)p(\theta_t|y_{1:t-1})d\theta_t$$

- Filtered distribution for θ_t

$$p(\theta_t|y_{1:t}) = \frac{p(y_t|\theta_t)p(\theta_t|y_{1:t-1})}{p(y_t|y_{1:t-1})}$$

Kalman filter idea

Goal: obtain $p(\theta_t|y_{1:t})$

Recursive procedure:

- Assume $p(\theta_{t-1}|y_{1:t-1})$ $\qquad p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$
- Prior for θ_t

$$p(\theta_t|y_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|y_{1:t-1})d\theta_{t-1}$$

- One-step ahead predictive distribution for y_t

$$p(y_t|y_{1:t-1}) = \int p(y_t|\theta_t)p(\theta_t|y_{1:t-1})d\theta_t$$

- Filtered distribution for θ_t

$$p(\theta_t|y_{1:t}) = \frac{p(y_t|\theta_t)p(\theta_t|y_{1:t-1})}{p(y_t|y_{1:t-1})}$$

- Prior for θ_t

- Prior for θ_t

$$
\begin{aligned}
p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t) N(\theta_{t-1}; m_{t-1}, C_{t-1}) d\theta_{t-1} \\
&= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1}
\end{aligned}
$$

- Prior for θ_t

$$\begin{aligned}
p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t) N(\theta_{t-1}; m_{t-1}, C_{t-1}) d\theta_{t-1} \\
&= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1} \\
&= N(a_t, R_t)
\end{aligned}$$

- Prior for θ_t

$$\begin{aligned} p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t) N(\theta_{t-1}; m_{t-1}, C_{t-1}) d\theta_{t-1} \\ &= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\ &\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1} \\ &= N(a_t, R_t) \end{aligned}$$

- One-step ahead predictive distribution for y_t

- Prior for θ_t

$$\begin{aligned}
p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t)N(\theta_{t-1}; m_{t-1}, C_{t-1})d\theta_{t-1} \\
&= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1} \\
&= N(a_t, R_t)
\end{aligned}$$

- One-step ahead predictive distribution for y_t

$$\begin{aligned}
p(y_t|y_{1:t-1}) &= \int p(y_t|\theta_t)p(\theta_t|y_{1:t-1})d\theta_t \\
&= \int N(y_t; F_t\theta_t, V_t)N(\theta_t; a_t, R_t)d\theta_t \\
&= \int \frac{1}{(2\pi)^{m/2}|V_t|^{1/2}} \exp\left(-\frac{1}{2}(y_t - F_t\theta_t)^\top V_t^{-1}(y_t - F_t\theta_t)\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|R_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - a_t)^\top R_t^{-1}(\theta_t - a_t)\right) d\theta_t
\end{aligned}$$

- Prior for θ_t

$$
\begin{aligned}
p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t) N(\theta_{t-1}; m_{t-1}, C_{t-1}) d\theta_{t-1} \\
&= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1} \\
&= N(a_t, R_t)
\end{aligned}
$$

- One-step ahead predictive distribution for y_t

$$
\begin{aligned}
p(y_t|y_{1:t-1}) &= \int p(y_t|\theta_t) p(\theta_t|y_{1:t-1}) d\theta_t \\
&= \int N(y_t; F_t\theta_t, V_t) N(\theta_t; a_t, R_t) d\theta_t \\
&= \int \frac{1}{(2\pi)^{m/2}|V_t|^{1/2}} \exp\left(-\frac{1}{2}(y_t - F_t\theta_t)^\top V_t^{-1}(y_t - F_t\theta_t)\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|R_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - a_t)^\top R_t^{-1}(\theta_t - a_t)\right) d\theta_t \\
&= N(f_t, Q_t)
\end{aligned}
$$

- Prior for θ_t

$$\begin{aligned}
p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t) N(\theta_{t-1}; m_{t-1}, C_{t-1}) d\theta_{t-1} \\
&= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1} \\
&= N(a_t, R_t)
\end{aligned}$$

- One-step ahead predictive distribution for y_t

$$\begin{aligned}
p(y_t|y_{1:t-1}) &= \int p(y_t|\theta_t) p(\theta_t|y_{1:t-1}) d\theta_t \\
&= \int N(y_t; F_t\theta_t, V_t) N(\theta_t; a_t, R_t) d\theta_t \\
&= \int \frac{1}{(2\pi)^{m/2}|V_t|^{1/2}} \exp\left(-\frac{1}{2}(y_t - F_t\theta_t)^\top V_t^{-1}(y_t - F_t\theta_t)\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|R_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - a_t)^\top R_t^{-1}(\theta_t - a_t)\right) d\theta_t \\
&= N(f_t, Q_t)
\end{aligned}$$

- Filtered distribution for θ_t

- Prior for θ_t

$$
\begin{aligned}
p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t) N(\theta_{t-1}; m_{t-1}, C_{t-1}) d\theta_{t-1} \\
&= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1} \\
&= N(a_t, R_t)
\end{aligned}
$$

- One-step ahead predictive distribution for y_t

$$
\begin{aligned}
p(y_t|y_{1:t-1}) &= \int p(y_t|\theta_t)p(\theta_t|y_{1:t-1}) d\theta_t \\
&= \int N(y_t; F_t\theta_t, V_t) N(\theta_t; a_t, R_t) d\theta_t \\
&= \int \frac{1}{(2\pi)^{m/2}|V_t|^{1/2}} \exp\left(-\frac{1}{2}(y_t - F_t\theta_t)^\top V_t^{-1}(y_t - F_t\theta_t)\right) \\
&\quad \frac{1}{(2\pi)^{p/2}|R_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - a_t)^\top R_t^{-1}(\theta_t - a_t)\right) d\theta_t \\
&= N(f_t, Q_t)
\end{aligned}
$$

- Filtered distribution for θ_t

$$
p(\theta_t|y_{1:t}) = \frac{p(y_t|\theta_t)p(\theta_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} = \frac{N(y_t; F_t\theta_t, V_t) N(\theta_t; a_t, R_t)}{N(y_t; f_t, Q_t)}
$$

- Prior for θ_t

$$\begin{aligned} p(\theta_t|y_{1:t-1}) &= \int N(\theta_t; G_t\theta_{t-1}, W_t) N(\theta_{t-1}; m_{t-1}, C_{t-1}) d\theta_{t-1} \\ &= \int \frac{1}{(2\pi)^{p/2}|W_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})^\top W_t^{-1}(\theta_t - G_t\theta_{t-1})\right) \\ &\quad \frac{1}{(2\pi)^{p/2}|C_{t-1}|^{1/2}} \exp\left(-\frac{1}{2}(\theta_{t-1} - m_{t-1})^\top C_{t-1}^{-1}(\theta_{t-1} - m_{t-1})\right) d\theta_{t-1} \\ &= N(a_t, R_t) \end{aligned}$$

- One-step ahead predictive distribution for y_t

$$\begin{aligned} p(y_t|y_{1:t-1}) &= \int p(y_t|\theta_t) p(\theta_t|y_{1:t-1}) d\theta_t \\ &= \int N(y_t; F_t\theta_t, V_t) N(\theta_t; a_t, R_t) d\theta_t \\ &= \int \frac{1}{(2\pi)^{m/2}|V_t|^{1/2}} \exp\left(-\frac{1}{2}(y_t - F_t\theta_t)^\top V_t^{-1}(y_t - F_t\theta_t)\right) \\ &\quad \frac{1}{(2\pi)^{p/2}|R_t|^{1/2}} \exp\left(-\frac{1}{2}(\theta_t - a_t)^\top R_t^{-1}(\theta_t - a_t)\right) d\theta_t \\ &= N(f_t, Q_t) \end{aligned}$$

- Filtered distribution for θ_t

$$\begin{aligned} p(\theta_t|y_{1:t}) &= \frac{p(y_t|\theta_t) p(\theta_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} = \frac{N(y_t; F_t\theta_t, V_t) N(\theta_t; a_t, R_t)}{N(y_t; f_t, Q_t)} \\ &= N(m_t, C_t) \end{aligned}$$

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal
- $E[\theta_t|y_{1:t-1}]$

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal
- $E[\theta_t|y_{1:t-1}] = G_t m_{t-1}$

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal
- $E[\theta_t|y_{1:t-1}] = G_t m_{t-1} = a_t$

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal
- $E[\theta_t|y_{1:t-1}] = G_t m_{t-1} = a_t$
- $Var[\theta_t|y_{1:t-1}]$

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal
- $E[\theta_t|y_{1:t-1}] = G_t m_{t-1} = a_t$
- $Var[\theta_t|y_{1:t-1}] = G_t C_{t-1} G_t^\top + W_t$

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal
- $E[\theta_t|y_{1:t-1}] = G_t m_{t-1} = a_t$
- $Var[\theta_t|y_{1:t-1}] = G_t C_{t-1} G_t^\top + W_t = R_t$

Latent state prior

Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$. Find $p(\theta_t|y_{1:t-1})$.

Evolution equation: $\theta_t = G_t\theta_{t-1} + w_t$ where $w_t \stackrel{ind}{\sim} N_p(0, W_t)$ (independent of $p(\theta_{t-1}|y_{1:t-1})$.

- CONAN $\implies$ $\theta_t|y_{1:t-1}$ is normal
- $E[\theta_t|y_{1:t-1}] = G_t m_{t-1} = a_t$
- $Var[\theta_t|y_{1:t-1}] = G_t C_{t-1} G_t^\top + W_t = R_t$
- $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$.

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies$ $y_t|y_{1:t-1}$ is normal

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies$ $y_t|y_{1:t-1}$ is normal
- $E[y_t|y_{1:t-1}]$

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies y_t|y_{1:t-1}$ is normal
- $E[y_t|y_{1:t-1}] = F_t a_t$

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies$ $y_t|y_{1:t-1}$ is normal
- $E[y_t|y_{1:t-1}] = F_t a_t = f_t$

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies y_t|y_{1:t-1}$ is normal
- $E[y_t|y_{1:t-1}] = F_t a_t = f_t$
- $Var[y_t|y_{1:t-1}]$

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies$ $y_t|y_{1:t-1}$ is normal
- $E[y_t|y_{1:t-1}] = F_t a_t = f_t$
- $Var[y_t|y_{1:t-1}] = F_t R_t F_t^\top + V_t$

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies$ $y_t|y_{1:t-1}$ is normal
- $E[y_t|y_{1:t-1}] = F_t a_t = f_t$
- $Var[y_t|y_{1:t-1}] = F_t R_t F_t^\top + V_t = Q_t$

One-step ahead prediction

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(y_t|y_{1:t-1})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- CONAN $\implies$ $y_t|y_{1:t-1}$ is normal
- $E[y_t|y_{1:t-1}] = F_t a_t = f_t$
- $Var[y_t|y_{1:t-1}] = F_t R_t F_t^\top + V_t = Q_t$
- $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$.

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal
- $Var[\theta_t|y_{1:t}]$

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal
- $Var[\theta_t|y_{1:t}] = (R_t^{-1} + F_t^\top V_t^{-1} F_t)^{-1}$

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal
- $Var[\theta_t|y_{1:t}] = (R_t^{-1} + F_t^\top V_t^{-1} F_t)^{-1} = C_t$

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal
- $Var[\theta_t|y_{1:t}] = (R_t^{-1} + F_t^\top V_t^{-1} F_t)^{-1} = C_t$
- $E[\theta_t|y_{1:t}]$

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal
- $Var[\theta_t|y_{1:t}] = (R_t^{-1} + F_t^\top V_t^{-1} F_t)^{-1} = C_t$
- $E[\theta_t|y_{1:t}] = C_t(R_t^{-1}a_t + F_t^\top V_t^{-1} y_t)$

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal
- $Var[\theta_t|y_{1:t}] = (R_t^{-1} + F_t^\top V_t^{-1} F_t)^{-1} = C_t$
- $E[\theta_t|y_{1:t}] = C_t(R_t^{-1}a_t + F_t^\top V_t^{-1} y_t) = m_t$

Latent state posterior - linear regression approach

We have $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$. Find $p(\theta_t|y_{1:t})$.

Observation equation: $y_t = F_t\theta_t + v_t$ where $v_t \stackrel{ind}{\sim} N_m(0, V_t)$ (independent of $p(\theta_t|y_{1:t-1})$.

- Linear regression $\implies$ $\theta_t|y_{1:t}$ is normal
- $Var[\theta_t|y_{1:t}] = (R_t^{-1} + F_t^\top V_t^{-1} F_t)^{-1} = C_t$
- $E[\theta_t|y_{1:t}] = C_t(R_t^{-1}a_t + F_t^\top V_t^{-1} y_t) = m_t$
- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$.

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array} \right] \middle| y_{1:t-1} \right)$$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array} \right] \middle| y_{1:t-1} \right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array} \right], \left[\begin{array}{cc} Q_t & \\ & R_t \end{array} \right] \right)$$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array} \right] \middle| y_{1:t-1} \right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array} \right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array} \right] \right)$$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array} \right] \middle| y_{1:t-1} \right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array} \right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array} \right] \right)$$

- MVN theory $\implies \theta_t|y_{1:t-1}, y_t$ is normal

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array} \right] \middle| y_{1:t-1} \right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array} \right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array} \right] \right)$$

- MVN theory $\implies \theta_t|y_{1:t-1}, y_t$ is normal
- $E[\theta_t|y_{1:t}]$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array}\right] \middle| y_{1:t-1}\right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array}\right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array}\right]\right)$$

- MVN theory $\implies$ $\theta_t|y_{1:t-1}, y_t$ is normal
- $E[\theta_t|y_{1:t}] = a_t + R_t F_t^\top Q_t^{-1}(y_t - f_t)$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array}\right]\middle| y_{1:t-1}\right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array}\right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array}\right]\right)$$

- MVN theory $\implies \theta_t|y_{1:t-1}, y_t$ is normal
- $E[\theta_t|y_{1:t}] = a_t + R_t F_t^\top Q_t^{-1}(y_t - f_t) = m_t$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array}\right] \middle| y_{1:t-1}\right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array}\right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array}\right]\right)$$

- MVN theory $\implies \theta_t|y_{1:t-1}, y_t$ is normal
- $E[\theta_t|y_{1:t}] = a_t + R_t F_t^\top Q_t^{-1}(y_t - f_t) = m_t$
- $Var[\theta_t|y_{1:t}]$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array}\right]\middle| y_{1:t-1}\right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array}\right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array}\right]\right)$$

- MVN theory $\implies$ $\theta_t|y_{1:t-1}, y_t$ is normal
- $E[\theta_t|y_{1:t}] = a_t + R_t F_t^\top Q_t^{-1}(y_t - f_t) = m_t$
- $Var[\theta_t|y_{1:t}] = R_t - R_t F_t^\top Q_t^{-1} F_t R_t$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array} \right] \middle| y_{1:t-1} \right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array} \right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array} \right] \right)$$

- MVN theory $\implies \theta_t|y_{1:t-1}, y_t$ is normal
- $E[\theta_t|y_{1:t}] = a_t + R_t F_t^\top Q_t^{-1}(y_t - f_t) = m_t$
- $Var[\theta_t|y_{1:t}] = R_t - R_t F_t^\top Q_t^{-1} F_t R_t = C_t$

Latent state posterior - multivariate normal approach

Prior is $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ and $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$. Find $p(\theta_t|y_{1:t})$.

Consider

$$p\left(\left[\begin{array}{c} y_t \\ \theta_t \end{array}\right]\middle| y_{1:t-1}\right) = N\left(\left[\begin{array}{c} f_t \\ a_t \end{array}\right], \left[\begin{array}{cc} Q_t & F_t R_t \\ R_t F_t^\top & R_t \end{array}\right]\right)$$

- MVN theory $\implies \theta_t|y_{1:t-1}, y_t$ is normal
- $E[\theta_t|y_{1:t}] = a_t + R_t F_t^\top Q_t^{-1}(y_t - f_t) = m_t$
- $Var[\theta_t|y_{1:t}] = R_t - R_t F_t^\top Q_t^{-1} F_t R_t = C_t$
- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$.

Kalman filter

1. Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$.
2. Obtain prior $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ where

$$a_t = G_t m_{t-1} \quad \text{and} \quad R_t = G_t C_{t-1} G_t^\top + W_t.$$

3. Obtain one step ahead predictive $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$ where

$$f_t = F_t a_t \quad \text{and} \quad Q_t = F_t R_t F_t^\top + V_t.$$

4. Obtain posterior $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ where

$$\begin{array}{rclcrcl} m_t &=& a_t + K_t e_t & \text{and} & C_t &=& R_t - K_t Q_t K_t^\top \\ e_t &=& y_t - f_t & \text{and} & K_t &=& R_t F_t^\top Q_t^{-1} \end{array}$$

Kalman filter

1. Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$.
2. Obtain prior $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ where

$$a_t = G_t m_{t-1} \quad \text{and} \quad R_t = G_t C_{t-1} G_t^\top + W_t.$$

3. Obtain one step ahead predictive $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$ where

$$f_t = F_t a_t \quad \text{and} \quad Q_t = F_t R_t F_t^\top + V_t.$$

4. Obtain posterior $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ where

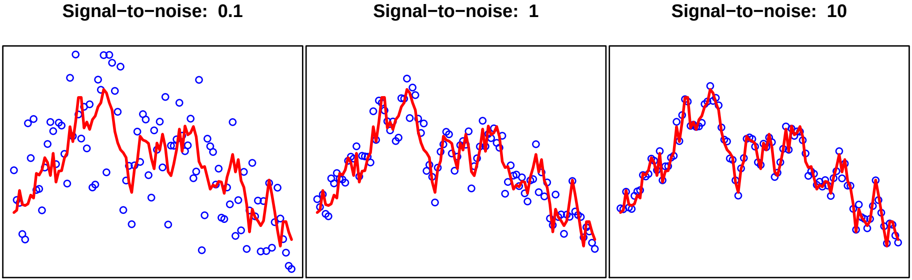
$$\begin{array}{rclcrcl} m_t &=& a_t + K_t e_t & \text{and} & C_t &=& R_t - K_t Q_t K_t^\top \\ e_t &=& y_t - f_t & \text{and} & K_t &=& R_t F_t^\top Q_t^{-1} \end{array}$$

K_t is the Kalman gain or adaptive coefficient (high values have more weight on the current observation while low values have more weight on the prior information).

Local level model

$$
\begin{aligned}
Y_t &= \theta_t + v_t & v_t &\stackrel{ind}{\sim} N(0,V) \\
\theta_t &= \theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N(0,W) \\
p(\theta_0) &= N(m_0, C_0)
\end{aligned}
$$

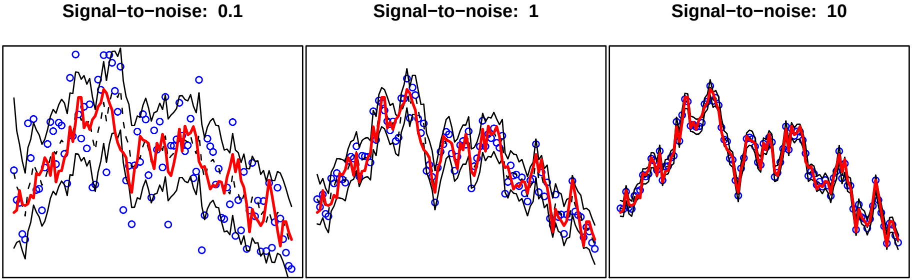
Signal-to-noise, $r = W/V$.



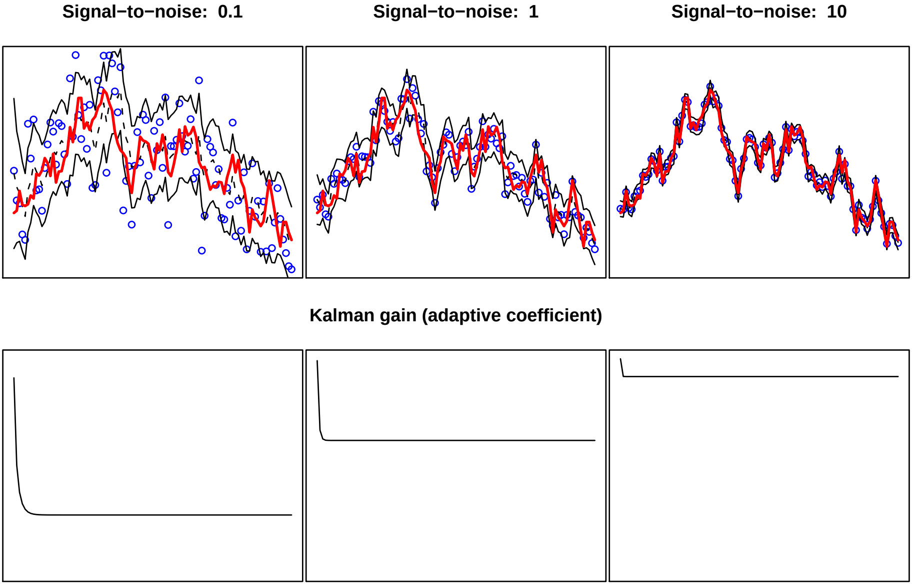
Local level model

$$
\begin{aligned}
Y_t &= \theta_t + v_t & v_t &\stackrel{ind}{\sim} N(0, V) \\
\theta_t &= \theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N(0, W) \\
p(\theta_0) &= N(m_0, C_0)
\end{aligned}
$$

Signal-to-noise, $r = W/V$.



Kalman gain (adaptive coefficient)



Kalman filter with missing data

- Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$.
- Obtain prior $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ where

$$a_t = G_t m_{t-1} \quad \text{and} \quad R_t = G_t C_{t-1} G_t^\top + W_t.$$

- Obtain one step ahead predictive $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$ where

$$f_t = F_t a_t \quad \text{and} \quad Q_t = F_t R_t F_t^\top + V_t.$$

- Obtain posterior $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ where
  - If y_t is observed,

$$\begin{array}{rclcrcl} m_t &=& a_t + K_t e_t & \text{and} & C_t &=& R_t - K_t Q_t K_t^\top \\ e_t &=& y_t - f_t & \text{and} & K_t &=& R_t F_t^\top Q_t^{-1} \end{array}$$

Kalman filter with missing data

- Assume $p(\theta_{t-1}|y_{1:t-1}) = N(m_{t-1}, C_{t-1})$.
- Obtain prior $p(\theta_t|y_{1:t-1}) = N(a_t, R_t)$ where

$$a_t = G_t m_{t-1} \quad \text{and} \quad R_t = G_t C_{t-1} G_t^\top + W_t.$$

- Obtain one step ahead predictive $p(y_t|y_{1:t-1}) = N(f_t, Q_t)$ where

$$f_t = F_t a_t \quad \text{and} \quad Q_t = F_t R_t F_t^\top + V_t.$$

- Obtain posterior $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ where
  - If y_t is observed,

$$\begin{array}{rclcrcl} m_t &=& a_t + K_t e_t & \text{and} & C_t &=& R_t - K_t Q_t K_t^\top \\ e_t &=& y_t - f_t & \text{and} & K_t &=& R_t F_t^\top Q_t^{-1} \end{array}$$

  - If y_t is not observed, $m_t = a_t$ and $C_t = R_t$.

Forecasting

Forecasting is simply the Kalman filter with missing observations.

Forecasting

Forecasting is simply the Kalman filter with missing observations. So,

$$\left(\begin{array}{c}\theta_{t+k}\\ Y_{t+k}\end{array}\right) \sim N\left(\left[\begin{array}{c}a_t(k)\\ f_t(k)\end{array}\right], \left[\begin{array}{cc}R_t(k) & F_{t+k}R_t(k)\\ R_t(k)F_{t+k}^\top & Q_t(k)\end{array}\right]\right)$$

Forecasting

Forecasting is simply the Kalman filter with missing observations. So,

$$\left(\begin{array}{c} \theta_{t+k} \\ Y_{t+k} \end{array} \right) \sim N\left(\left[\begin{array}{c} a_t(k) \\ f_t(k) \end{array} \right], \left[\begin{array}{cc} R_t(k) & F_{t+k}R_t(k) \\ R_t(k)F_{t+k}^\top & Q_t(k) \end{array} \right] \right)$$

where

$$\begin{array}{rl} a_t(k) & = G_{t+k}a_t(k-1) \\ R_t(k) & = G_{t+k}R_t(k-1)G_{t+k}^\top + W_{t+k} \\ f_t(k) & = F_{t+k}a_t(k) \\ Q_t(k) & = F_{t+k}R_t(k)R_{t+k}^\top + V_t \end{array}$$

Forecasting

Forecasting is simply the Kalman filter with missing observations. So,

$$\begin{pmatrix} \theta_{t+k} \\ Y_{t+k} \end{pmatrix} \sim N\left(\begin{bmatrix} a_t(k) \\ f_t(k) \end{bmatrix}, \begin{bmatrix} R_t(k) & F_{t+k}R_t(k) \\ R_t(k)F_{t+k}^\top & Q_t(k) \end{bmatrix} \right)$$

where

$$\begin{array}{rl} a_t(k) & = G_{t+k}a_t(k-1) \\ R_t(k) & = G_{t+k}R_t(k-1)G_{t+k}^\top + W_{t+k} \\ f_t(k) & = F_{t+k}a_t(k) \\ Q_t(k) & = F_{t+k}R_t(k)R_{t+k}^\top + V_t \end{array}$$

with $a_t(0) = m_t$ and $R_t(0) = C_t$.

Kalman Smoother

Smoothing can be accomplished in a manner similar to the Kalman filter via the Kalman smoother.

Kalman Smoother

Smoothing can be accomplished in a manner similar to the Kalman filter via the Kalman smoother. If we have $\theta_{t+1}|y_{1:T} \sim N(s_{t+1}, S_{t+1})$, then $\theta_t|y_{1:T} \sim N(s_t, S_t)$ where

$$\begin{aligned} s_t &= m_t + C_t G_{t+1}^\top R_{t+1}^{-1}(s_{t+1} - a_{t+1}) \\ S_t &= C_t - C_t G_{t+1}^\top R_{t+1}^{-1}(R_{t+1} - S_{t+1}) R_{t+1}^{-1} G_{t+1} C_t. \end{aligned}$$

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$ with

$$\begin{aligned} H_t &= (C_t^{-1} + G_{t+1}^\top W_{t+1}^{-1} G_{t+1})^{-1} \\ h_t &= H_t(C_t^{-1} m_t + G_{t+1}^\top W_{t+1}^{-1} \theta_{t+1}) \end{aligned}$$

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$ with

$$\begin{aligned} H_t &= (C_t^{-1} + G_{t+1}^\top W_{t+1}^{-1} G_{t+1})^{-1} \\ h_t &= H_t(C_t^{-1} m_t + G_{t+1}^\top W_{t+1}^{-1} \theta_{t+1}) \end{aligned}$$

The algorithm is then

- Forward filter to obtain $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ for all t.

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$ with

$$\begin{aligned} H_t &= (C_t^{-1} + G_{t+1}^\top W_{t+1}^{-1} G_{t+1})^{-1} \\ h_t &= H_t(C_t^{-1} m_t + G_{t+1}^\top W_{t+1}^{-1} \theta_{t+1}) \end{aligned}$$

The algorithm is then

- Forward filter to obtain $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ for all t.
- Sample $\theta_T \sim N(m_T, C_T)$.

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$ with

$$\begin{aligned} H_t &= (C_t^{-1} + G_{t+1}^\top W_{t+1}^{-1} G_{t+1})^{-1} \\ h_t &= H_t(C_t^{-1} m_t + G_{t+1}^\top W_{t+1}^{-1} \theta_{t+1}) \end{aligned}$$

The algorithm is then

- Forward filter to obtain $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ for all t.
- Sample $\theta_T \sim N(m_T, C_T)$.
- For $t = T-1, T-2, \ldots, 1, 0$,

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$ with

$$\begin{aligned} H_t &= (C_t^{-1} + G_{t+1}^\top W_{t+1}^{-1} G_{t+1})^{-1} \\ h_t &= H_t(C_t^{-1} m_t + G_{t+1}^\top W_{t+1}^{-1} \theta_{t+1}) \end{aligned}$$

The algorithm is then

- Forward filter to obtain $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ for all t.
- Sample $\theta_T \sim N(m_T, C_T)$.
- For $t = T-1, T-2, \ldots, 1, 0$,
  - Calculate h_t and H_t based on θ_{t+1}.

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$ with

$$
\begin{aligned}
H_t &= (C_t^{-1} + G_{t+1}^\top W_{t+1}^{-1} G_{t+1})^{-1} \\
h_t &= H_t(C_t^{-1} m_t + G_{t+1}^\top W_{t+1}^{-1} \theta_{t+1})
\end{aligned}
$$

The algorithm is then

- Forward filter to obtain $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ for all t.
- Sample $\theta_T \sim N(m_T, C_T)$.
- For $t = T-1, T-2, \ldots, 1, 0$,
  - Calculate h_t and H_t based on θ_{t+1}.
  - Draw $\theta_t \sim N(h_t, H_t)$.

Backward sampling

Recall

- $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ is available for all t from filtering and
- $p(\theta_t|\theta_{t+1}, y_{1:T}) = N(h_t, H_T)$ with

$$\begin{aligned} H_t &= (C_t^{-1} + G_{t+1}^\top W_{t+1}^{-1} G_{t+1})^{-1} \\ h_t &= H_t(C_t^{-1} m_t + G_{t+1}^\top W_{t+1}^{-1} \theta_{t+1}) \end{aligned}$$

The algorithm is then

- Forward filter to obtain $p(\theta_t|y_{1:t}) = N(m_t, C_t)$ for all t.
- Sample $\theta_T \sim N(m_T, C_T)$.
- For $t = T-1, T-2, \ldots, 1, 0$,
  - Calculate h_t and H_t based on θ_{t+1}.
  - Draw $\theta_t \sim N(h_t, H_t)$.

This is then a joint draw of $\theta_{0:T} \sim p(\theta_0, \ldots, \theta_T|y_{1:T})$.

Inference questions?

Any questions on performing inference on the latent states in a DLM?