Mathematics and numerics for data assimilation and state estimation – Lecture 14

Summer semester 2020

Overview

Summary of lecture 13

- In additive Gaussian noise setting (both for dynamics and observations), smoothing density for $V_{0:J}|Y_{1:J} = y_{1:J}$:

$$\pi(v_{0:J}|y_{1:J}) = \frac{1}{Z}\exp\Bigg(-\frac{1}{2}\sum_{j=1}^{J}|h(v_j) - y_j|_\Gamma^2 \\ -\frac{1}{2}|v_0 - m_0|_{C_0}^2 - \frac{1}{2}\sum_{j=0}^{J-1}|v_{j+1} - \Psi(v_j)|_\Sigma^2\Bigg)$$

- Stability of the density wrt perturbations (under some assumptions on the dynamics),

$$d_H(\pi_{V_{0:J}|Y_{1:J}}(\cdot \mid y_{1:J}), \pi_{V_{0:J}|Y_{1:J}}(\cdot \mid \tilde{y}_{1:J})) \leq c\sqrt{\sum_{j=1}^{J}|y_j - \tilde{y}_j|^2}$$

- For deterministic dynamics with uncertain initial condition, we derived a smoothing density for $V_0|Y_{1:J} = y_{1:J}$.

Overview

Dynamics and observation setting

Continuous state-space dynamics: A mapping $\Psi \in C(\mathbb{R}^d, \mathbb{R}^d)$ is associated to the dynamics

$$\begin{aligned} V_{j+1} &= \Psi(V_j) + \xi_j, \quad j = 0, 1, \ldots \\ V_0 &\sim N(m_0, C_0) \end{aligned} \tag{1}$$

with an iid sequence $\xi_j \sim N(0, \Sigma)$.

Observations:

$$Y_j = h(V_j) + \eta_j, \quad j = 1, 2, \ldots, \tag{2}$$

where $h \in C(\mathbb{R}^d, \mathbb{R}^k)$ and iid sequence $\eta_j \sim N(0, \Gamma)$.

Independence assumptions:

$$\{\eta_j\} \perp \{\xi_j\} \perp \{V_0\}$$

Objective: Derive iterative formulas for pdfs of $V_n | Y_{1:n} = y_{1:n}$ and $V_{n+1} | Y_{1:n} = y_{1:n}$ for $n \geq 1$.

Filtering – the prediction step

Setting: At time $n \geq 0$, we have observations $Y_{1:n} = y_{1:n}$ and we have computed $\pi_{V_n|Y_{1:n}}(v_n|y_{1:n}) =: \pi(v_n|y_{1:n})$ (for $n = 0$, we mean by this $\pi_{V_0}(v_0)$).

What is the distribution of $V_{n+1}|Y_{1:n} = y_{1:n}$?

Prediction: By the law of total probability

$$\pi(v_{n+1}|y_{1:n}) = \int_{\mathbb{R}^d} \pi(v_{n+1}, v_n \mid y_{1:n})\, dv_n$$

$$= \int_{\mathbb{R}^d} \pi(v_{n+1}|v_n, y_{1:n})\pi(v_n|y_{1:n})\, dv_n$$

$$= \int_{\mathbb{R}^d} \pi(v_{n+1}|v_n)\pi(v_n|y_{1:n})\, dv_n$$

The last step follows from $\xi_n \perp \{Y_{1:n}\}$ and

$$V_{n+1}|(V_n = v_n, Y_{1:n} = y_{1:n}) \quad = \quad \Psi(V_n) + \xi_n|(V_n = v_n, Y_{1:n} = y_{1:n})$$

$$= \Psi(v_n) + \xi_n|(V_n = v_n, Y_{1:n} = y_{1:n})$$

$$= \Psi(v_n) + \xi_n|(V_n = v_n).$$

Filtering – the analysis step

Setting: At time $n+1$, we have the old observations $Y_{1:n} = y_{1:n}$ and we have computed the prediction density $\pi(v_{n+1}|y_{1:n})$. Now we seek to assimilate the new observation $Y_{n+1} = y_{n+1}$ into our state estimate.

What is the distribution of $V_{n+1}|(Y_{1:n} = y_{1:n}, Y_{n+1} = y_{n+1})$?

Analysis step

$$\begin{aligned}
\pi(v_{n+1}|y_{1:n}, y_{n+1}) &= \frac{\pi(v_{n+1}, y_{n+1}|y_{1:n})}{\pi(y_{n+1}|y_{1:n})} \\
&= \frac{\pi(y_{n+1}|v_{n+1}, y_{1:n})\pi(v_{n+1}|y_{1:n})}{\pi(y_{n+1}|y_{1:n})} \\
&= \frac{\pi(y_{n+1}|v_{n+1})\pi(v_{n+1}|y_{1:n})}{\pi(y_{n+1}|y_{1:n})}
\end{aligned}$$

Here we used that $\eta_{n+1} \perp \{Y_{1:n}\}$:

$$\begin{aligned}
Y_{n+1}|(V_{n+1} = v_{n+1}, Y_{1:n} = y_{1:n}) &= h(v_{n+1}) + \eta_{n+1}|(V_{n+1} = v_{n+1}, Y_{1:n} = y_{1:n}) \\
&= h(v_{n+1}) + \eta_{n+1}|(V_{n+1} = v_{n+1})
\end{aligned}$$

Summary filtering steps

Prediction step:

$$\pi(v_{n+1}|y_{1:n}) = \int_{\mathbb{R}^d} \pi(v_{n+1}|v_n)\pi(v_n|y_{1:n})\, dv_n$$

Analysis step:

$$\pi(v_{n+1}|y_{1:n+1}) = \frac{\pi(y_{n+1}|v_{n+1})\pi(v_{n+1}|y_{1:n})}{\pi(y_{n+1}|y_{1:n})}$$

Remarks:

- $\pi(v_{n+1}|v_n)$ is the transition kernel density :

$$\pi(v_{n+1}|v_n) \text{ "} = \text{"prob density of going from } v_n \text{ to } v_{n+1}$$

- Generally, it is not easy to derive usable closed-form filtering densities from the above steps, and they are rather a starting point for approximation filtering algorithms.

Relationship between smoothing and filtering pdfs

The derived equations are **exact** both for

- the updated filtering pdf $\pi_{V_n|Y_{1:n}}(v_n|y_{1:n})$
- and for the smoothing pdf $\pi_{V_{0:n}|Y_{1:n}}(v_{0:n}|y_{1:n})$.

Consequently,

$$\pi_{V_n|Y_{1:n}}(v_n|y_{1:n}) = \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \pi_{V_{0:n}|Y_{1:n}}(v_{0:n}|y_{1:n})\, dv_0 \dots dv_{n-1}. \tag{3}$$

Explicit computations of either of them is often complicated when Ψ and/or h are nonlinear, following from their effects on the pdf

$$\pi(v_{0:n}|y_{1:n}) = \frac{1}{Z} \exp\Bigg(-\frac{1}{2}\sum_{j=1}^{n} |h(v_j) - y_j|_\Gamma^2 \\ - \frac{1}{2}|v_0 - m_0|_{C_0}^2 - \frac{1}{2}\sum_{j=0}^{n-1} |v_{j+1} - \Psi(v_j)|_\Sigma^2 \Bigg)$$

Well-posedness of the filter pdf

Corollary 1 (SST 7.7)

Fix $n \in \mathbb{N}$, a pair of observation sequences $y_{1:n}, \tilde{y}_{1:n} \in \mathbb{R}^{k \times n}$, and assume that the dynamics V_j satisfies

$$\mathbb{E}\left[\sum_{j=0}^{n} (1 + |h(V_j)|^2)\right] < \infty.$$

Then there exists a constant $c > 0$ that depends on $y_{1:n}$ and $\tilde{y}_{1:n}$ such that

$$d_{TV}(\pi_{V_n|Y_{1:n}}(\cdot|y_{1:n}), \pi_{V_n|Y_{1:n}}(\cdot|\tilde{y}_{1:n})) \leq c\sqrt{\sum_{j=1}^{n} |y_j - \tilde{y}_j|^2}$$

Proof:

We will use

$$d_H(\pi_{V_{0:n}|Y_{1:n}}(\cdot|y_{1:n}), \pi_{V_{0:n}|Y_{1:n}}(\cdot|\tilde{y}_{1:n})) \leq c\sqrt{\sum_{j=1}^{n} |y_j - \tilde{y}_j|^2} \quad \text{(LSZ 2.15)}$$

$$\text{and that} \quad d_{TV}(\hat{\pi}, \check{\pi}) \leq \sqrt{2} d_H(\hat{\pi}, \check{\pi}) \quad \text{for any } \hat{\pi}, \check{\pi} \in \mathcal{M}.$$

By definition

$$d_{TV}\left(\pi_{V_n|Y_{1:n}}(\cdot|y_{1:n}), \pi_{V_n|Y_{1:n}}(\cdot|\tilde{y}_{1:n})\right) = \frac{1}{2}\int_{\mathbb{R}^d} \left|\pi(v_n|y_{1:n}) - \pi(v_n|\tilde{y}_{1:n})\right| dv_n$$

$$\stackrel{(3)}{=} \frac{1}{2}\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} \pi(v_{0:n}|y_{1:n}) - \pi(v_{0:n}|\tilde{y}_{1:n})\, dv_0 \dots dv_{n-1} \right| dv_n$$

$$\leq \frac{1}{2}\int_{\mathbb{R}^d} \dots \int_{\mathbb{R}^d} \left|\pi(v_{0:n}|y_{1:n}) - \pi(v_{0:n}|\tilde{y}_{1:n})\right| dv_0 \dots dv_n$$

$$= d_{TV}\left(\pi_{V_{0:n}|Y_{1:n}}(\cdot|y_{1:n}), \pi_{V_{0:n}|Y_{1:n}}(\cdot|\tilde{y}_{1:n})\right)$$

$$\leq \sqrt{2} d_H\left(\pi_{V_{0:n}|Y_{1:n}}(\cdot|y_{1:n}), \pi_{V_{0:n}|Y_{1:n}}(\cdot|\tilde{y}_{1:n})\right) \leq \sqrt{2}c|y_{1:n} - \tilde{y}_{1:n}|$$

Overview

Kalman filter

- Is the filtering problem with additive Gaussian noise (all independent) and both linear dynamics $\Psi(v) = Av$ and linear observations $h(v) = Hv$.

- In this setting the filtering pdfs will remain Gaussian for all times, and we obtain surprisingly simple recursive formulas the pdfs.

- Groundbreaking paper by Richard Kalman, "A new approach to linear filtering and prediction problems" J. Basic Engineering 1960, has, according to Google Scholar, been cited more than 33000 times.

Applications in control theory

In many real application, the state estimation and state prediction of filtering is often combined with control

$$\begin{aligned} V_{j+1} &= AV_j + B\,u_j + \xi_j \qquad && \textbf{dynamics} \\ Y_j &= HV_j + \eta_j && \textbf{observations}, \end{aligned}$$

where u_j belongs to set of admissible controls, e.g., $u_j \in \sigma(Y_{1:j})$.

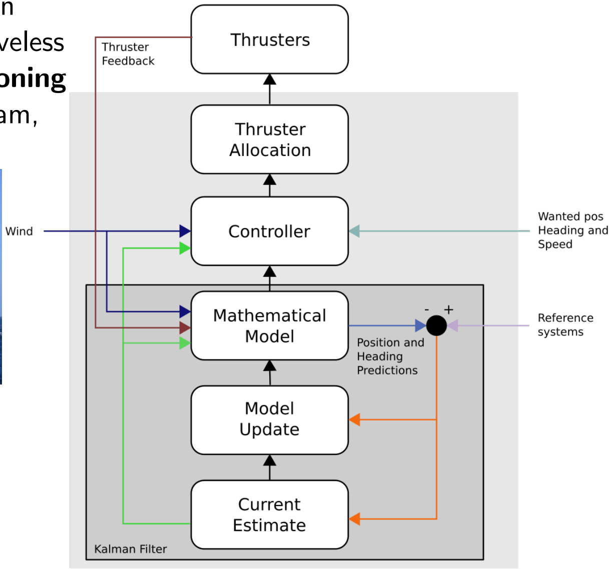
For example, the linear quadratic Gaussian control problem

$$\min_{u_n, u_{n+1}, \ldots, u_N} \mathbb{E}\left[V_N^T Q_0 V_N + \sum_{j=n}^{N} (V_j^T Q_1 V_j + u_j^T Q_2 u_j) \,\middle|\, Y_{1:n} = y_{1:n} \right]$$

Applications:

- Guidance and navigation systems [autopilots, driveless cars, **dynamical positioning in ships** , Apollo program, missiles, . . .]

- econometric time-series analysis and signal processing
- and seed of many approximate Gaussian filtering methods.



The linear-Gaussian setting

We consider the **dynamics** on $\mathbb{R}^d$:

$$\begin{aligned} V_{j+1} &= AV_j + \xi_j, \qquad j = 0, 1, \dots \\ V_0 &\sim N(m_0, C_0) \end{aligned}$$

with $\xi_j \overset{iid}{\sim} N(0, \Sigma)$, and the **observations** on $\mathbb{R}^k$:

$$Y_j = HV_j + \eta_j, \quad j = 1, 2, \dots$$

with $\eta_j \overset{iid}{\sim} N(0, \Gamma)$.

Independence assuptions: $V_0 \perp \{\xi_j\} \perp \{\eta_j\}$.

Objective: Show that, under assumption $C_0, \Sigma, \Gamma > 0$,

$$V_n | Y_{1:n} = y_{1:n} \sim N(m_n, C_n), \quad V_{n+1} | Y_{1:n} = y_{1:n} \sim N(\hat{m}_{n+1}, \hat{C}_{n+1})$$

for all $n > 0$, and describe recursive formulas for evolution of pdfs

$$(m_n, C_n) \mapsto (\hat{m}_{n+1}, \hat{C}_{n+1}) \quad \text{and} \quad (\hat{m}_{n+1}, \hat{C}_{n+1}, y_{n+1}) \mapsto (m_{n+1}, C_{n+1}).$$

Gaussianity of the filtering pdfs

Property 1: The dynamics $V_{j+1} = AV_j + \xi_j$ is Gaussian for any $j \geq 0$.

Motivation: Assuming V_j is Gaussian, $AV_j + \xi_j$ is a linear combination of independent Gaussians, which again is a Gaussian (cf. LSZ 1.5 and Ubung 6). Holds by induction, since V_0 is Gaussian.

Property 2: If $V_j|Y_{1:j} = y_{1:j} \sim N(m_j, C_j)$, then $V_{j+1}|Y_{1:j} = y_{1:j} \sim N(\hat{m}_{j+1}, \hat{C}_{j+1})$ for computable moments with $\hat{C}_{j+1} > 0$.

Motivation: Writing $Z_j := V_j|(Y_{1:j} = y_{1:j})$, observe that

$$
\begin{aligned}
V_{j+1}|(Y_{1:j} = y_{1:j}) &= AV_j + \xi_j|(Y_{1:j} = y_{1:j}) \\
&= A\Big(V_j|(Y_{1:j} = y_{1:j})\Big) + \xi_j \\
&= AZ_j + \xi_j
\end{aligned}
$$

Hence, $V_{j+1}|(Y_{1:j} = y_{1:j})$ is linear combination of independent Gaussians and thus itself Gaussian. Moreover,

$$
\begin{aligned}
\hat{m}_{j+1} &= \mathbb{E}\left[V_{j+1}|Y_{1:j} = y_{1:j}\right] \\
&= \mathbb{E}\left[AV_j + \xi_j|Y_{1:j} = y_{1:j}\right] \\
&= A\mathbb{E}\left[V_j|Y_{1:j} = y_{1:j}\right] + \mathbb{E}\left[\xi_j\right] \\
&= Am_j,
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{C}_{j+1} &= \mathbb{E}\left[(V_{j+1} - \hat{m}_{j+1})(V_{j+1} - \hat{m}_{j+1})^T | Y_{1:j} = y_{1:j}\right] \\
&= \mathbb{E}\left[(AV_j + \xi_j - Am_j)(AV_j + \xi_j - Am_j)^T | Y_{1:j} = y_{1:j}\right] \\
&= \mathbb{E}\left[A(V_j - m_j)(V_j - m_j)^T A^T | Y_{1:j} = y_{1:j}\right] \\
&+ \mathbb{E}\left[\xi_j\right] \mathbb{E}\left[(V_j - m_j)^T A^T | Y_{1:j} = y_{1:j}\right] \\
&+ \mathbb{E}\left[A(V_j - m_j) | Y_{1:j} = y_{1:j}\right] \mathbb{E}\left[\xi_j^T\right] + \mathbb{E}\left[\xi_j \xi_j^T\right] \\
&= A\mathbb{E}\left[(V_j - m_j)(V_j - m_j)^T | Y_{1:j} = y_{1:j}\right] A^T + \Sigma \\
&= \underbrace{AC_jA^T}_{\geq 0} + \underbrace{\Sigma}_{> 0}. \quad > 0
\end{aligned}
$$

Property 3: If $V_{j+1}|Y_{1:j} = y_{1:j} \sim N(\hat{m}_{j+1}, \hat{C}_{j+1})$ with $\hat{C}_{j+1} > 0$, then for any $y_{j+1} \in \mathbb{R}^k$ we have that $V_{j+1}|Y_{1:j+1} = y_{1:j+1} \sim N(m_{j+1}, C_{j+1})$ and the moments are computable.

Motivation: By the previous derivations and using that

$$\begin{aligned} \pi(v_{j+1}|y_{1:j+1}) &\propto \pi(y_{j+1}|v_{j+1})\pi(v_{j+1}|y_{1:j}) \\ &\propto \exp\Big(-\frac{1}{2}|y_{j+1} - Hv_{j+1}|_{\Gamma}^2 - \frac{1}{2}|v_{j+1} - \hat{m}_{j+1}|_{\hat{C}_{j+1}}^2 \Big) \end{aligned} \tag{4}$$

where we used that

$$Y_{j+1}|V_{j+1} = v_{j+1} \;\; = \;\; Hv_{j+1} + \eta_j \sim N(Hv_{j+1}, \Gamma)$$

Making the ansatz $V_{j+1}|Y_{1:j+1} = y_{1:j+1} \sim N(m_{j+1}, C_{j+1})$ and equating same-order-term coefficients in the exponent of (4) the exponent of our ansatz pdf

$$\pi(v_{j+1}|y_{1:j+1}) \propto \exp\Big(-\frac{1}{2}|v_{j+1} - m_{j+1}|_{C_{j+1}}^2 \Big)$$

verifies the claim.

$$\hat{C} > 0 = U\begin{bmatrix}\lambda_1 & 0 \\ 0 & \ddots & \lambda_d\end{bmatrix}U^T \qquad \lambda_1 \geq \cdots \geq \lambda_d > 0$$

$$\Rightarrow \hat{C}^{-1} = U\begin{bmatrix}\lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_d^{-1}\end{bmatrix}U^T$$

$$\Rightarrow \hat{C}^{-1} > 0$$

Moreover, equating quadratic terms yields

$$C_{j+1}^{-1} = \underbrace{\hat{C}_{j+1}^{-1}}_{>0} + \underbrace{H^T \Gamma^{-1} H}_{\geq 0} > 0 \tag{5}$$

and equating linear terms yields

$$C_{j+1}^{-1} m_{j+1} = \hat{C}_{j+1}^{-1} \hat{m}_{j+1} + H^T \Gamma^{-1} y_{j+1}$$

(For more details on equating terms, see similar argument in Lecture 10.)

Consequence of these properties:

$V_1 = A V_0 + \xi_0$

Given a sequence $y_1, y_2, \ldots$,

- Starting from $V_0 \sim N(m_0, C_0)$ it follows by **Property 2** that $V_1 \sim N(\hat{m}_1, \hat{C}_1)$ with

$$\hat{m}_1 = A m_0 \quad \text{and} \quad \hat{C}_1 = A C_0 A^T + \Sigma > 0, \quad \text{since} \Sigma > 0$$

- Property 3 then implies that $V_1 | Y_1 = y_1 \sim N(m_1, C_1)$ with computable moments, where

$$C_1^{-1} = \hat{C}_1^{-1} + H^T \Gamma^{-1} H$$

is positive definite since $\hat{C}_1, \Gamma > 0$, and thus invertible.

$V_n | Y_{1:n} = y_{1:n} \sim N(m_n, C_n)$, $C_n > 0$

- By induction, $V_{n+1} | Y_{1:n} = y_{1:n} \sim N(\hat{m}_{n+1}, \hat{C}_{n+1})$ with $\hat{C}_{n+1} > 0$
- and $V_{n+1} | Y_{1:n+1} = y_{n+1} \sim N(m_{n+1}, C_{n+1})$ for computable moments with

$$C_{n+1}^{-1} = \hat{C}_{n+1}^{-1} + H^T \Gamma^{-1} H$$

which is positive definite since $\hat{C}_{n+1}, \Gamma > 0$, and thus invertible.

Theorem 2 (LSZ 4.1)

For the linear-Gaussian filtering problem with $\underbrace{C_0, \Sigma, \Gamma}_{>0}$*, it holds for any observation sequence* $y_1, y_2, \dots$ *and* $n \geq 1$ *that*
$V_n | Y_{1:n} = y_{1:n} \sim N(m_n, C_n)$ *where*

$$C_n^{-1} = \hat{C}_n^{-1} + H^T \Gamma^{-1} H$$

is positive definite and thus invertible, and

$$C_n^{-1} m_n = \hat{C}_n^{-1} \hat{m}_n + H^T \Gamma^{-1} y_n.$$

To avoid dealing with the inverse of C_n, we apply the Woodbury matrix identity (LSZ 4.4) to obtain

$$C_n = (\hat{C}_n^{-1} + H^T \Gamma^{-1} H)^{-1} = \hat{C}_n - \underbrace{\hat{C}_n H^T (H \hat{C}_n H^T + \Gamma)^{-1}}_{=:K_n} H \hat{C}_n$$

$$= (I - K_n H) \hat{C}_n$$

and

$$m_n = (I - K_n H) \hat{m}_n + K_n y_n \qquad \text{(ubung 7)}$$

Kalman filtering iteration algorithm

Given any sequence $y_1, y_2, \ldots$ and $V_n | Y_{1:n} = y_{1:n} \sim N(m_n, C_n)$ the next-time filtering distributions are iteratively determined by

Prediction

$V_{n+1} | Y_{1:n} = y_{1:n} \sim N(\hat{m}_{n+1}, \hat{C}_{n+1})$

$$\hat{m}_{n+1} = A m_n$$
$$\hat{C}_{n+1} = A C_n A^T + \Sigma$$

and

Analysis

$$d_{n+1} = y_{n+1} - H\hat{m}_{n+1} \qquad \textbf{innovation}$$
$$K_{n+1} = \hat{C}_{n+1} H^T (H \hat{C}_{n+1} H^T + \Gamma)^{-1} \qquad \textbf{Kalman gain}$$
$$m_{n+1} = \hat{m}_{n+1} + K_{n+1} d_{n+1} = (I - K_{n+1} H)\hat{m}_{n+1} + K_{n+1} y_{n+1}$$
$$C_{n+1} = (I - K_{n+1} H)\hat{C}_{n+1}$$

Example

Dynamics on $\mathbb{R}^2$

$$V_{j+1} = \overbrace{\begin{bmatrix}1 & 0.1\\ 0 & 1\end{bmatrix}}^{A} V_j + \xi_j,$$

$$V_0 \sim N\left(\underbrace{\begin{bmatrix}0\\ 1\end{bmatrix}}_{m_0}, \underbrace{\begin{bmatrix}1/4 & 0\\ 0 & 1/4\end{bmatrix}}_{C_0} \right)$$

where $\xi_j \overset{iid}{\sim} N(0, \Sigma)$ with $\Sigma = \begin{bmatrix}0.01 & 0\\ 0 & 0.1\end{bmatrix}$.

And **observations** on $\mathbb{R}$:

$$Y_j = \underbrace{\begin{bmatrix}0 & 1\end{bmatrix}}_{H} V_j + \eta_j, \qquad \eta_j \overset{iid}{\sim} N(0, \underbrace{1/4}_{\Gamma}).$$

An observation sequence is generated from synthetic data:
$y_j = V_{j,2}^{\dagger}(\omega) + \eta_j(\omega)$.

```
% Dynamics parameters
A = [1  0.1; 0 1];
Sigma = [0.01 0; 0 0.1];
m0 = [0; 1]; C0 = [1/4 0; 0 1/4];

%Observation parameters
H = [0 1];Gamma = 1/4;

n =40;

%generate observation sequence
rng(12009) %set seed for reproducibility
v = zeros(2, n+1); y = zeros(1,n);
v(:,1) = m0+ sqrt(C0)*randn(2,1);
for j=1:n
    v(:,j+1) = A*v(:,j) + sqrt(Sigma)*randn(2,1);
    y(j) = H*v(:,j+1) + sqrt(Gamma)*randn();
end
```

Continuation of Matlab program

```
% Filtering distributions
m = zeros(2, n+1);
C = zeros(2,2,n+1);

m(:,1) = m0;
C(:,:,1) = C0;
for j=1:n
    %prediction step
    m(:,j+1)    = A*m(:, j);
    C(:,:,j+1)  = A*C(:,:, j)*A' + Sigma;

    %Analysis
    K          = C(:,:,j+1)*H'/(H*C(:,:,j+1)*H' + Gamma);
    m(:,j+1)   = m(:,j+1) + K*(y(j) - H*m(:,j+1));
    C(:,:,j+1) = (eye(2)-K*H)* C(:,:,j+1);
end
```

Numerical results - noisy case

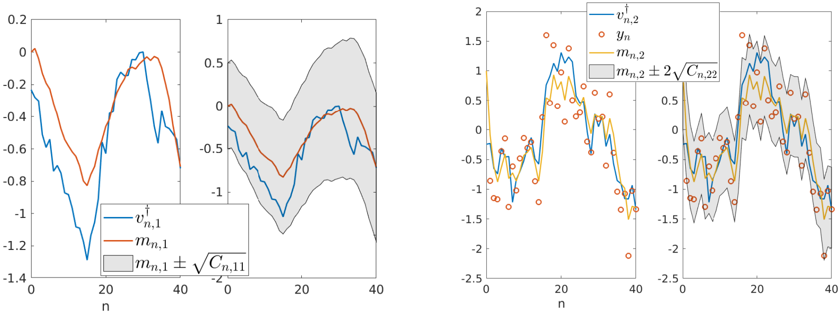


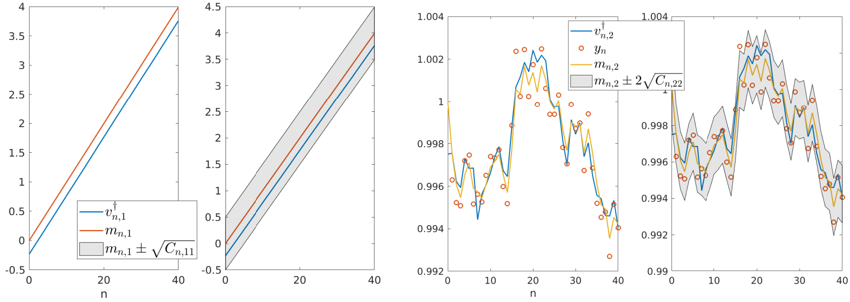
Figure: Left pair of figures: Evolution of first component, mean and "one standard deviation" grey uncertainty region in the right plot. Right pair of figures: Same for the second component, but here also including measurements.

What is a good error measure? Is it $\|m - v^{\dagger}\|$ or $\|m_{n,2} - y_n\|$, or should we also rely on uncertainty regions?

Numerical results - “noiseless case”

We consider the same problem, but now with almost no noise, except for in $V_{0,1}$:

$$C_0 = \begin{bmatrix} 1/4 & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Gamma = 10^{-6}.$$

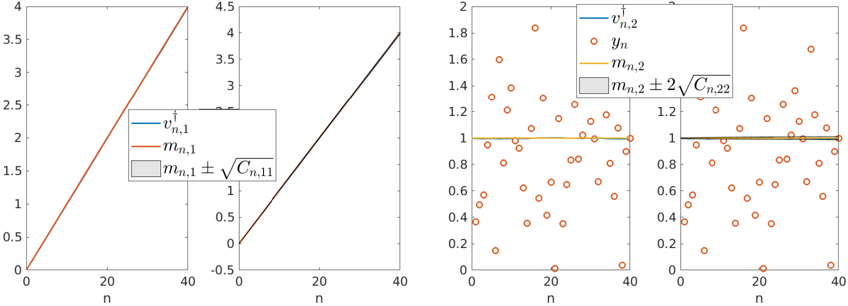


Note that uncertainty in first component remains for all times!

Numerical results - "noiseless case 2"

We consider the same problem, but now with almost no noise, except for in Γ:

$$C_0 = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Gamma = 1/4.$$

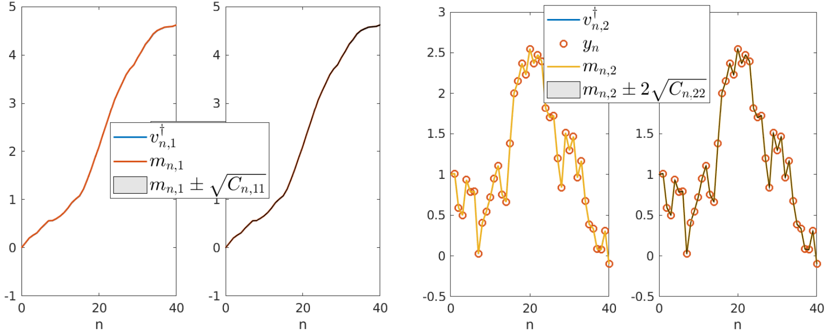


$|\Gamma| \gg |C_n| \implies |K_n| \ll 1 \implies$ we do almost not take observations into account, and then the problem is almost deterministic.

Numerical results - "noiseless case 3"

We consider the same problem, but now with almost no noise, except for in Σ_{22}:

$$C_0 = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 0.1 \end{bmatrix}, \qquad \Gamma = 10^{-6}.$$

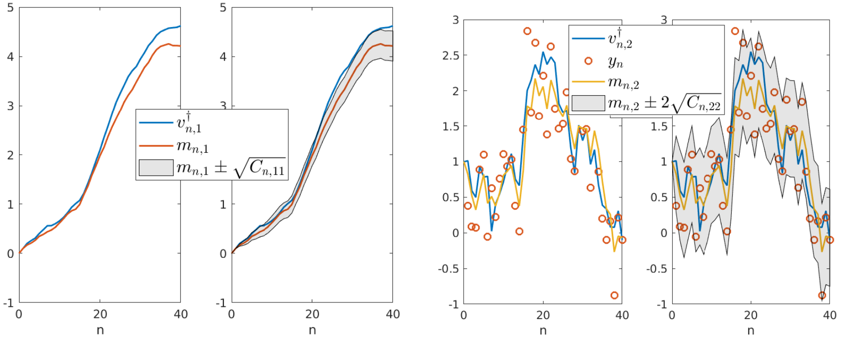


Very accurate observations $y_n \approx V_{n,2}$ means that by relying on the observations (and not the model) in $V_{n,2}$, we can track it very accurately.

Numerical results - “noiseless case 4”

We consider the same problem, but now with almost no noise, except for in Σ_{22} and Γ:

$$C_0 = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 0.1 \end{bmatrix}, \qquad \Gamma = 1/4$$

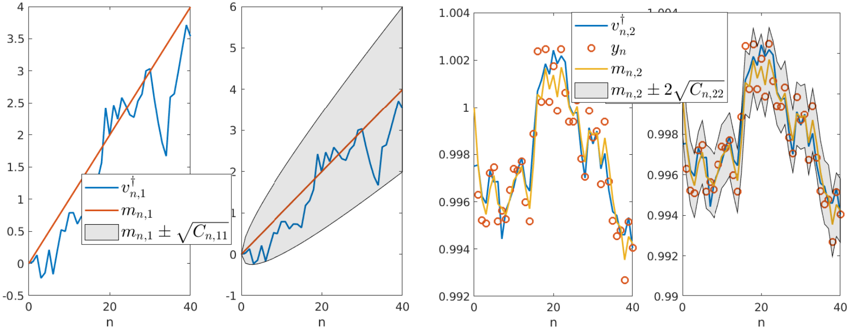


Both noisy dynamics and uncertain observations in the second component introduces uncertainty in both components.

Numerical results - “noiseless case 5”

We consider the same problem, but now with almost no noise, except for in Σ_{11}:

$$C_0 = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 0.1 & 0 \\ 0 & 10^{-6} \end{bmatrix}, \qquad \Gamma = 10^{-6}$$



Noisy dynamics in the first component, an unobserved component and does not influence the dynamics of the second component, will (almost) only introduce uncertainty in the first component.

Summary

For a filtering problem

$$V_{j+1} = \Psi(V_j) + \xi_j$$
$$Y_j = h(V_j) + \eta_j, \quad j = 1, 2, \dots,$$

with Gaussian noise and initial condition and Ψ and h linear mappings, we have derived iterative formulas for the distribution $V_n | Y_{1:n} = y_{1:n}$.

- Theory extends straightforwardly to settings with time-dependence: $\Psi_n(v) = A_n v$, $h(v) = H_n v$, Σ_n, Γ_n.
- Also possible to derive the moments for the Kalman smoother distribution $V_{0:n} | Y_{1:n} = y_{1:n}$, which also is a Gaussian, cf. LSZ 3.1
- Next time, we will look at Approximate Gaussian filters, which are extensions of Kalman filtering to nonlinear settings.
- No lectures or ubung during Pentecost week. Next lecture on Monday, June 8.