Applied Math Reading Seminar

Data Assimilation I

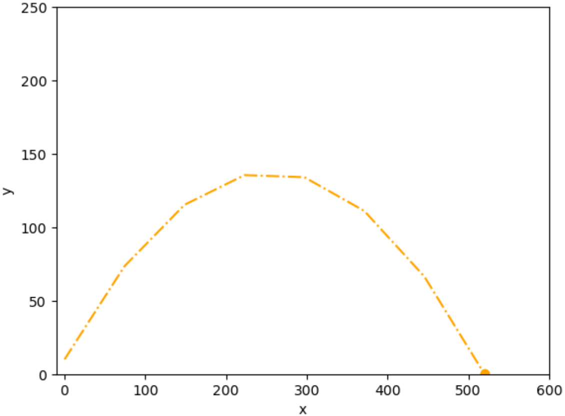
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

March 20, 2023

Motivating Toy Example

$$x(t) = x_0 + v_0^x t$$

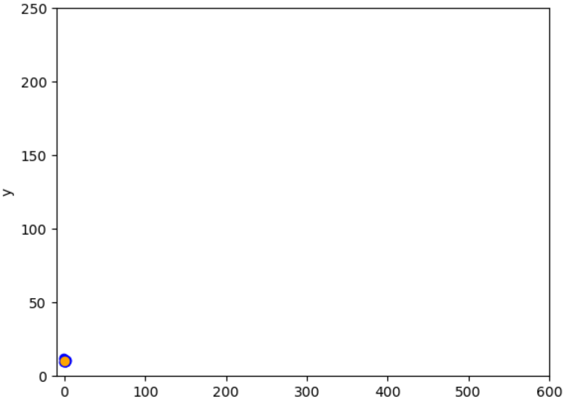
$$y(t) = y_0 + v_0^y t - \frac{g}{2} t^2$$



Motivating Toy Example

$$x(t) = x_0 + v_0^x t$$

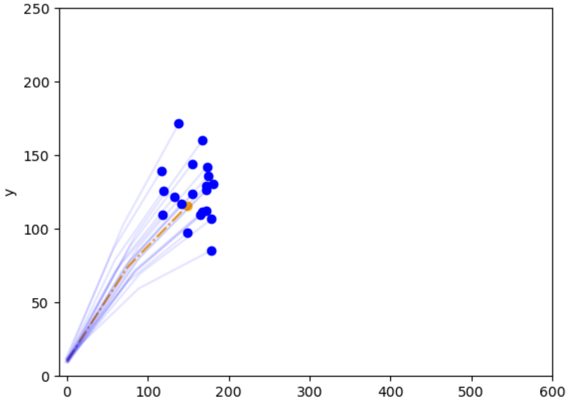
$$y(t) = y_0 + v_0^y t - \frac{g}{2} t^2$$



Motivating Toy Example

$$x(t) = x_0 + v_0^x t$$

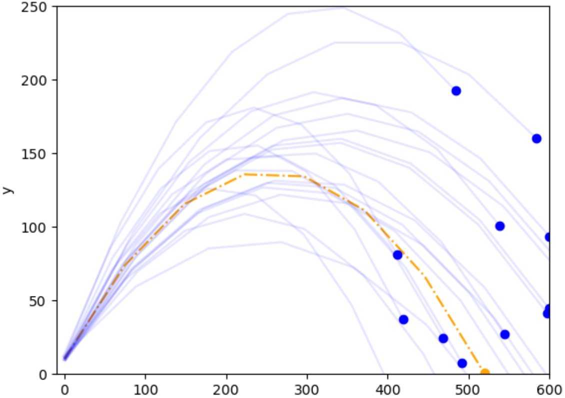
$$y(t) = y_0 + v_0^y t - \frac{g}{2} t^2$$



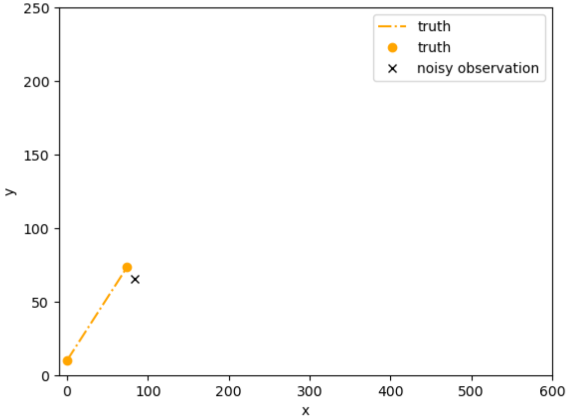
Motivating Toy Example

$$x(t) = x_0 + v_0^x t$$

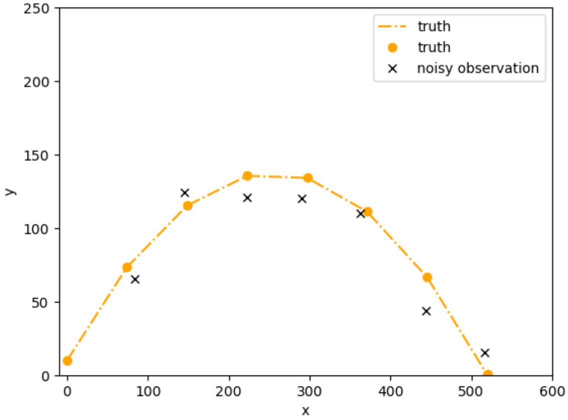
$$y(t) = y_0 + v_0^y t - \frac{g}{2} t^2$$



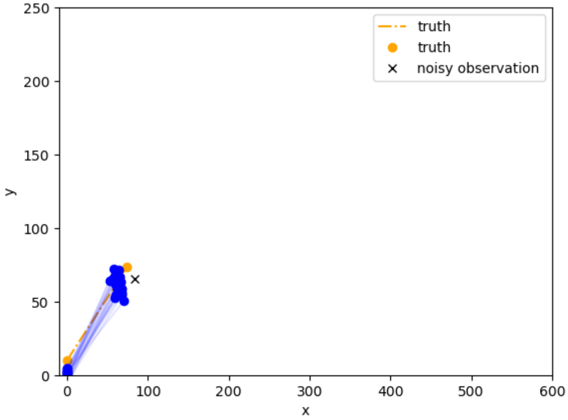
Motivating Toy Example



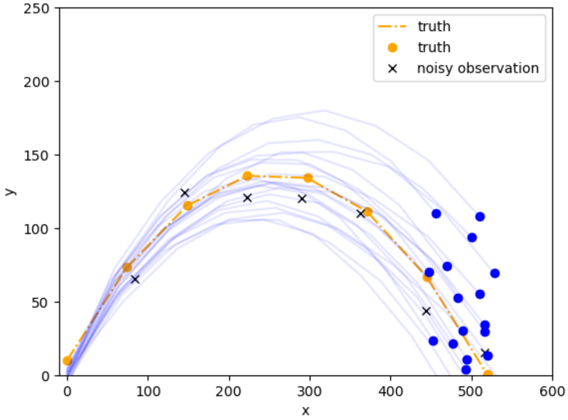
Motivating Toy Example



Motivating Toy Example



Motivating Toy Example



Notation

- $\mathbb{Z}^+ = \mathbb{N}_0 = \{0, 1, 2, \dots\}$
- $|\cdot|$ **Euclidean**
- For symmetric, positive definite A
  - $\langle \cdot, \cdot \rangle_A := \langle \cdot, A^{-1} \cdot \rangle$
  - $|\cdot|_A := |A^{-1/2} \cdot|$

Data Assimilation Setting

State Space Model

$$
\begin{aligned}
\text{Dynamics models:}\quad & v_{j+1} = \Psi\left(v_j\right) + \xi_j,\ j \in \mathbb{Z}^+ \\
\text{Data Model:}\quad & y_{j+1} = h(v_{j+1}) + \eta_{j+1},\ j \in \mathbb{Z}^+ \\
\text{Probabilistic Structure:}\quad & v_0 \sim \mathcal{N}\left(m_0, C_0\right),\ \xi_j \sim \mathcal{N}(0, \Sigma),\ \eta_j \sim \mathcal{N}(0, \Gamma) \text{ i.i.d.} \\
\text{Probabilistic Structure:}\quad & v_0 \perp \{\xi_j\} \perp \{\eta_j\}
\end{aligned}
$$

Assume C_0, Σ, Γ are positive definite, $\Psi \in C(\mathbb{R}^d, \mathbb{R}^d)$,and $h \in C(\mathbb{R}^d, \mathbb{R}^k)$.

Filtering and Smoothing

$$V := \{v_0, \dots, v_J\}, Y := \{y_1, \dots, y_J\}, Y_j := \{y_1, \dots, y_j\}$$

$$\textbf{Smoothing Problem}: \quad \Pi(V) := \mathbb{P}(V \mid Y)$$

Filtering and Smoothing

$$V := \{v_0, \dots, v_J\}, Y := \{y_1, \dots, y_J\}, Y_j := \{y_1, \dots, y_j\}$$

$$\textbf{Smoothing Problem}: \quad \Pi(V) := \mathbb{P}(V \mid Y)$$

$$\textbf{Filtering Problem}: \quad \pi_j(v_j) := \mathbb{P}(v_j \mid Y_j),\ j \in [J]$$

Smoothing Problem Posterior Derivation

Baye's Theorem

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A)\mathbb{P}(B \mid A)}{\mathbb{P}(B)}$$

Smoothing Problem Posterior Derivation

Baye's Theorem

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A)\mathbb{P}(B \mid A)}{\mathbb{P}(B)}$$

$$\Pi(V) = \mathbb{P}(V \mid Y) \propto \mathbb{P}(V)\mathbb{P}(Y \mid V)$$

Smoothing Problem Posterior Derivation

$$\begin{aligned}\mathbb{P}(V) &= \mathbb{P}(v_J, v_{J-1}, \dots, v_0)\\ &= \mathbb{P}(v_J \mid v_{J-1}, \dots, v_0)\mathbb{P}(v_{J-1}, \dots, v_0)\\ &= \mathbb{P}(v_J \mid v_{J-1})\mathbb{P}(v_{J-1}, \dots, v_0)\\ &= \dots\\ &= \mathbb{P}(v_0)\prod_{j=0}^{J-1}\mathbb{P}(v_{j+1} \mid v_j)\end{aligned}$$

$$v_0 \sim \mathcal{N}(m_0, C_0) \qquad v_{j+1} \mid v_j \sim \mathcal{N}(\Psi(v_j), \Sigma)$$

$$\mathbb{P}(V) \propto \exp\left\{-\frac{1}{2}|v_0 - m_0|_{C_0}^2 - \frac{1}{2}\sum_{j=0}^{J-1}|v_{j+1} - \Psi(v_j)|_{\Sigma}^2\right\} =: \exp\{-R(V)\}$$

Smoothing Problem Posterior Derivation

$$\begin{aligned}\mathbb{P}(Y \mid V) &= \mathbb{P}(y_1, y_2, \ldots y_J \mid V) \\ &= \prod_{j=0}^{J-1} \mathbb{P}(y_{j+1} \mid V) \\ &= \prod_{j=0}^{J-1} \mathbb{P}(y_{j+1} \mid v_{j+1})\end{aligned}$$

$$y_{j+1} \mid v_{j+1} \sim \mathcal{N}(h(v_{j+1}, \Gamma)$$

$$\mathbb{P}(Y \mid V) \propto \exp\left\{-\frac{1}{2}\sum_{j=0}^{J-1} |y_{j+1} - h(v_{j+1})|_{\Sigma}^{2}\right\} =: \exp\{-L(V; Y)\}$$

Filtering Problem Posterior Distribution

$$\pi_j(v_j) := \mathbb{P}(v_j \mid Y_j) \quad j \in [J]$$

$$\textbf{Prediction Step}: \quad \hat{\pi}_{j+1} = \mathcal{P}\pi_j = \mathbb{P}(v_{j+1} \mid Y_j)$$

$$\textbf{Analysis Step:} \quad \pi_{j+1} = \mathcal{A}_j \widehat{\pi}_{j+1}$$

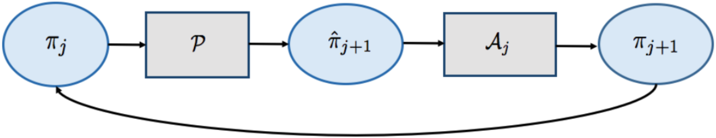


Figure 7.1 Prediction and analysis steps combined.

Filtering Problem Posterior Distribution

$$\begin{aligned}
\widehat{\pi}_{j+1}\left(v_{j+1}\right) &= \mathbb{P}\left(v_{j+1} \mid Y_j\right) \\
&= \int_{\mathbb{R}^d} \mathbb{P}\left(v_{j+1} \mid Y_j, v_j\right) \mathbb{P}\left(v_j \mid Y_j\right) dv_j \\
&= \int_{\mathbb{R}^d} \mathbb{P}\left(v_{j+1} \mid v_j\right) \mathbb{P}\left(v_j \mid Y_j\right) dv_j \\
&= \int_{\mathbb{R}^d} \mathbb{P}\left(v_{j+1} \mid v_j\right) \pi_j\left(v_j\right) dv_j \\
&= \frac{1}{(2\pi)^{d/2}(\det \Sigma)^{1/2}} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}\left|v_{j+1} - \Psi\left(v_j\right)\right|_{\Sigma}^2\right) \pi_j\left(v_j\right) dv_j.
\end{aligned}$$

Filtering Problem Posterior Distribution

$$
\begin{aligned}
\pi_{j+1}\left(v_{j+1}\right) &= \mathbb{P}\left(v_{j+1} \mid Y_{j+1}\right) \\
&= \mathbb{P}\left(v_{j+1} \mid Y_j, y_{j+1}\right) \\
&= \frac{\mathbb{P}\left(y_{j+1} \mid v_{j+1}, Y_j\right) \mathbb{P}\left(v_{j+1} \mid Y_j\right)}{\mathbb{P}\left(y_{j+1} \mid Y_j\right)} \\
&= \frac{\mathbb{P}\left(y_{j+1} \mid v_{j+1}\right) \mathbb{P}\left(v_{j+1} \mid Y_j\right)}{\mathbb{P}\left(y_{j+1} \mid Y_j\right)} \\
&= \frac{\exp\left(-\frac{1}{2}\left|y_{j+1} - h\left(v_{j+1}\right)\right|_\Gamma^2\right) \widehat{\pi}_{j+1}\left(v_{j+1}\right)}{\int_{\mathbb{R}^d} \exp\left(-\frac{1}{2}\left|y_{j+1} - h\left(v_{j+1}\right)\right|_\Gamma^2\right) \widehat{\pi}_{j+1}\left(v_{j+1}\right) dv_{j+1}}.
\end{aligned}
$$

Well-Posedness of Smoothing

Thm 7.6

If $\exists R \geq 0$ such that data Y, Y' and observation function h satisfy

- $|Y|, |Y'| \leq R$
- $\mathbb{E}^{\rho}\varphi^2(V) < \infty$ where $\varphi(V) := \left(\sum_{j=1}^{J} |h(v_j)|^2\right)^{1/2}$,

then $\exists \kappa \in [0, \infty)$ independent of Y, Y' such that

$$d_H(\Pi, \Pi') \leq \kappa |Y - Y'|$$

$$d_H(\Pi, \Pi') = \left\| \sqrt{\Pi} - \sqrt{\Pi'} \right\|_{L^2}$$

Well-Posedness of Filtering

Cor 7.7

If $\exists R \geq 0$ such that data Y, Y' and observation function h satisfy

- $|Y|, |Y'| \leq R$
- $\mathbb{E}^{\rho}\varphi^2(V) < \infty$ where $\varphi(V) := \left(\sum_{j=1}^{J} |h(v_j)|^2\right)^{1/2}$,

then $\exists \kappa = \kappa(R) \in [0, \infty)$ such that

$$d_{\text{TV}}(\pi_J, \pi_J') \leq \kappa |Y - Y'|$$

$$d_{\text{TV}}(\pi_J, \pi_J') = \left\|\pi_J - \pi_J'\right\|_{L^1}$$

Kalman Filter (Linear-Gaussian Setting)

State Space Model

$$
\begin{aligned}
\text{Dynamics models:} \quad & v_{j+1} = \Psi\left(v_j\right) + \xi_j,\ j \in \mathbb{Z}^+ \\
\text{Data Model:} \quad & y_{j+1} = h(v_{j+1}) + \eta_{j+1},\ j \in \mathbb{Z}^+ \\
\text{Probabilistic Structure:} \quad & v_0 \sim \mathcal{N}\left(m_0, C_0\right), \xi_j \sim \mathcal{N}(0, \Sigma), \eta_j \sim \mathcal{N}(0, \Gamma) \text{ i.i.d.} \\
\text{Probabilistic Structure:} \quad & v_0 \perp \{\xi_j\} \perp \{\eta_j\}
\end{aligned}
$$

Kalman Filter (Linear-Gaussian Setting)

Linear dynamics and linear observations, $M \in \mathbb{R}^{d\times d}$, $H \in \mathbb{R}^{k\times d}$

State Space Model

$$\begin{aligned}
\text{Dynamics models:} \quad & v_{j+1} = M v_j + \xi_j,\ j \in \mathbb{Z}^+ \\
\text{Data Model:} \quad & y_{j+1} = H v_{j+1} + \eta_{j+1},\ j \in \mathbb{Z}^+ \\
\text{Probabilistic Structure:} \quad & v_0 \sim \mathcal{N}(m_0, C_0), \xi_j \sim \mathcal{N}(0,\Sigma), \eta_j \sim \mathcal{N}(0,\Gamma) \text{ i.i.d.} \\
\text{Probabilistic Structure:} \quad & v_0 \perp \{\xi_j\} \perp \{\eta_j\}
\end{aligned}$$

Kalman Filter (Linear-Gaussian Setting)

Thm 8.2 and 8.3

For $j \in \mathbb{Z}^+$, $\pi_0, \{\hat{\pi}_{j+1}\}, \{\pi_{j+1}\}$ are all Gaussian distributions and C_j is positive definite.

$$\widehat{\pi}_{j+1} = \mathbb{P}(v_{j+1} \mid Y_j) = \mathcal{N}\left(\widehat{m}_{j+1}, \widehat{C}_{j+1}\right), \quad \text{(prediction)}$$

$$\pi_{j+1} = \mathbb{P}(v_{j+1} \mid Y_{j+1}) = \mathcal{N}(m_{j+1}, C_{j+1}), \quad \text{(analysis)}$$

$$\begin{aligned}
\widehat{m}_{j+1} &= M m_j, \\
\widehat{C}_{j+1} &= M C_j M^\top + \Sigma, \\
C_{j+1}^{-1} &= \left(M C_j M^\top + \Sigma\right)^{-1} + H^\top \Gamma^{-1} H, \\
C_{j+1}^{-1} m_{j+1} &= \left(M C_j M^\top + \Sigma\right)^{-1} M m_j + H^\top \Gamma^{-1} y_{j+1}.
\end{aligned}$$

Proof (if there's time)

Proof (if there's time)

Proof (if there's time)

Kalman Filter Algorithm

Algorithm 1: Kalman Filter Algorithm

Input: $\{y_j\}_{j=1}^{J}$, $\pi_0 = \mathcal{N}(m_0, C_0)$, Σ, Γ

Result: $\widehat{\pi}_{j+1} = \mathcal{N}\left(\widehat{m}_{j+1}, \widehat{C}_{j+1}\right)$ and $\pi_{j+1} = \mathcal{N}\left(m_{j+1}, C_{j+1}\right)$

for $j \leftarrow 0$ **to** $J-1$ **do**

Prediction:

$$\widehat{m}_{j+1} = M m_j$$
$$\widehat{C}_{j+1} = M C_j M^\top + \Sigma$$

Analysis:

$$m_{j+1} = \widehat{m}_{j+1} + K_{j+1} d_{j+1}$$
$$C_{j+1} = (I - K_{j+1} H)\,\widehat{C}_{j+1}$$

where

$$d_{j+1} = y_{j+1} - H\widehat{m}_{j+1}$$
$$S_{j+1} = H\widehat{C}_{j+1} H^\top + \Gamma$$
$$K_{j+1} = \widehat{C}_{j+1} H^\top S_{j+1}^{-1}$$

end