The Filtering Problem and Kalman Filtering

Roland Hostettler

November 14, 2018

Recap: Discretization of Nonlinear Models

- Nonlinear dynamic model:

$$\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t)) + \boldsymbol{B}_w(\boldsymbol{x}(t))\boldsymbol{w}(t)$$

- Discretization of the linearized model:

$$\boldsymbol{x}_n = \boldsymbol{x}_{n-1} + \int_{t_{n-1}}^{t_n} e^{\boldsymbol{A}_x(t_n - t)} \mathrm{d}t f(\boldsymbol{x}_{n-1}) + \boldsymbol{q}_n$$

  with

$$\boldsymbol{q}_n \sim \mathcal{N}(0, \boldsymbol{Q}_n),\ \boldsymbol{Q}_n \approx \int_{t_{n-1}}^{t_n} e^{\boldsymbol{A}_x(t_n-\tau)} \boldsymbol{B}_w \boldsymbol{\Sigma}_w \boldsymbol{B}_w^\mathsf{T} e^{\boldsymbol{A}_x^\mathsf{T}(t_n-\tau)} \mathrm{d}\tau$$

- Euler–Maruyama discretization:

$$\boldsymbol{x}_n = \boldsymbol{x}_{n-1} + \Delta t f(\boldsymbol{x}_{n-1}) + \boldsymbol{q}_n$$

  with

$$\boldsymbol{q}_n \sim \mathcal{N}(0, \boldsymbol{Q}_n),\ \boldsymbol{Q}_n \approx \Delta t \boldsymbol{B}_w(\boldsymbol{x}_{n-1}) \boldsymbol{\Sigma}_w \boldsymbol{B}_w(\boldsymbol{x}_{n-1})^\mathsf{T}.$$

Intended Learning Outcomes

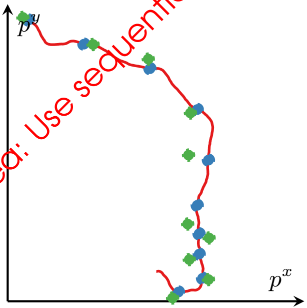
After this lecture, you will be able to:

- explain the relationship between the dynamic model, measurement model, and the filtering methodology,
- describe and employ the Kalman filter for linear state-space models.

Example: GPS Navigation (1/2)

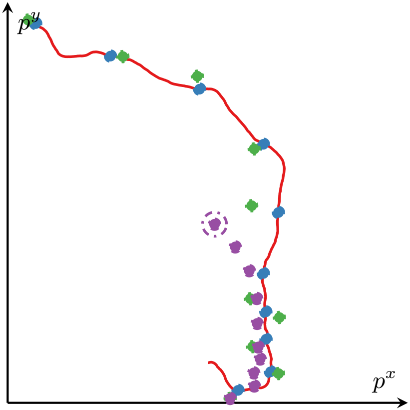
- Goal: Estimate the position $\boldsymbol{x}_n = \begin{bmatrix} p_n^x & p_n^y \end{bmatrix}^\mathsf{T}$ at each time step t_n
- Measurements: Noisy position ($\boldsymbol{G}_n = \boldsymbol{I}$):

$$\boldsymbol{y}_n = \boldsymbol{G}_n \boldsymbol{x}_n + \boldsymbol{r}_n$$



Idea: Use sequential least squares!

Example: GPS Navigation (2/2)



The Filtering Approach

Iterate the following two steps for all points in time:

1. Prediction: Predict the current state using the dynamic model (also called *time update*)
2. Measurement Update: Estimate the current state using the prediction and the new measurement

The Filtering Approach: Prediction

- Objective: Predict the current state $\boldsymbol{x}_n$ at t_n given all previous data $\boldsymbol{y}_{1:n-1} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_{n-1}\}$
- Notation:
  - $\hat{\boldsymbol{x}}_{n|n-1}$: Denotes the predicted value of $\boldsymbol{x}_n$ given the measurements $\boldsymbol{y}_{1:n-1} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_{n-1}\}$
  - $\boldsymbol{P}_{n|n-1}$: Denotes the covariance of the predicted $\boldsymbol{x}_n$
- Prediction adds uncertainty

The Filtering Approach: Measurement Update

- Objective: Estimate the current value of $\boldsymbol{x}_n$ given the new measurement $\boldsymbol{y}_n$, taking the prediction into account
- Notation:
  - $\hat{\boldsymbol{x}}_{n|n}$: Denotes the estimated value at t_n, given the measurements $\boldsymbol{y}_{1:n} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n\}$
  - $\boldsymbol{P}_{n|n}$: Denotes the covariance at t_n
- The measurement update reduces uncertainty

Linear State-Space Model

- Linear state-space model:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{w}(t)$$
$$\boldsymbol{y}_n = \boldsymbol{G}_n\boldsymbol{x}(t_n) + \boldsymbol{r}_n$$

- Discrete-time equivalent:

$$\boldsymbol{x}_n = \boldsymbol{F}_n\boldsymbol{x}_{n-1} + \boldsymbol{q}_n$$
$$\boldsymbol{y}_n = \boldsymbol{G}_n\boldsymbol{x}_n + \boldsymbol{r}_n$$

  with

$$\mathrm{E}\{\boldsymbol{q}_n\} = 0,\ \mathrm{Cov}\{\boldsymbol{q}_n\} = \boldsymbol{Q}_n,$$
$$\mathrm{E}\{\boldsymbol{r}_n\} = 0,\ \mathrm{Cov}\{\boldsymbol{r}_n\} = \boldsymbol{R}_n$$

- Initial conditions:

$$\mathrm{E}\{\boldsymbol{x}_0\} = \boldsymbol{m}_0$$
$$\mathrm{Cov}\{\boldsymbol{x}_0\} = \boldsymbol{P}_0$$

Linear Model: Measurement Update (1/3)

- Assume that the prediction yields the prior knowledge $\hat{\boldsymbol{x}}_{n|n-1}$, $\boldsymbol{P}_{n|n-1}$
- $\boldsymbol{y}_n$ provides the new information of the state
- We can use regularized least squares to estimate $\boldsymbol{x}_n$:

$$J_{\text{ReLS}}(\boldsymbol{x}_n) = (\boldsymbol{y}_n - \boldsymbol{G}_n \boldsymbol{x}_n)^\mathsf{T} \boldsymbol{R}_n^{-1} (\boldsymbol{y}_n - \boldsymbol{G}_n \boldsymbol{x}_n) + (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1})^\mathsf{T} \boldsymbol{P}_{n|n-1}^{-1} (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1})$$

and solve

$$\hat{\boldsymbol{x}}_{n|n} = \underset{\boldsymbol{x}_n}{\operatorname{argmin}}\, J_{\text{ReLS}}(\boldsymbol{x}_n)$$

Linear Model: Measurement Update (2/3)

- Regularized least squares problem:

$$\hat{\boldsymbol{x}}_{n|n} = \underset{\boldsymbol{x}_n}{\operatorname{argmin}}\, (\boldsymbol{y}_n - \boldsymbol{G}_n \boldsymbol{x}_n)^\mathsf{T} \boldsymbol{R}_n^{-1} (\boldsymbol{y}_n - \boldsymbol{G}_n \boldsymbol{x}_n) \\ + (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1})^\mathsf{T} \boldsymbol{P}_{n|n-1}^{-1} (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_{n|n-1})$$

- Solution (see Lecture 3 / Chapter 2.4):

$$\boldsymbol{K}_n = \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^\mathsf{T} (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^\mathsf{T} + \boldsymbol{R}_n)^{-1}$$
$$\hat{\boldsymbol{x}}_{n|n} = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n (\boldsymbol{y}_n - \boldsymbol{G}_n \hat{\boldsymbol{x}}_{n|n-1})$$

- Covariance of $\hat{\boldsymbol{x}}_{n|n}$:

$$\boldsymbol{P}_{n|n} = \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^\mathsf{T} + \boldsymbol{R}_n) \boldsymbol{K}_n^\mathsf{T}$$

- $\boldsymbol{K}_n$ is called the Kalman gain

Linear Model: Measurement Update (3/3)

- Measurement update:

$$
\begin{aligned}
\boldsymbol{K}_n &= \boldsymbol{P}_{n|n-1}\boldsymbol{G}_n^{\mathsf{T}}(\boldsymbol{G}_n\boldsymbol{P}_{n|n-1}\boldsymbol{G}_n^{\mathsf{T}} + \boldsymbol{R}_n)^{-1} \\
\hat{\boldsymbol{x}}_{n|n} &= \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n(\boldsymbol{y}_n - \boldsymbol{G}_n\hat{\boldsymbol{x}}_{n|n-1}) \\
\boldsymbol{P}_{n|n} &= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n(\boldsymbol{G}_n\boldsymbol{P}_{n|n-1}\boldsymbol{G}_n^{\mathsf{T}} + \boldsymbol{R}_n)\boldsymbol{K}_n^{\mathsf{T}}
\end{aligned}
$$

- It can be shown that it holds that:

$$
\begin{aligned}
\mathrm{E}\{\boldsymbol{x}_n \mid \boldsymbol{y}_{1:n}\} &= \hat{\boldsymbol{x}}_{n|n} \\
\mathrm{Cov}\{\boldsymbol{x}_n \mid \boldsymbol{y}_{1:n}\} &= \boldsymbol{P}_{n|n}
\end{aligned}
$$

Linear Model: Prediction (1/2)

- Linear dynamic model:

$$\boldsymbol{x}_n = \boldsymbol{F}_n \boldsymbol{x}_{n-1} + \boldsymbol{q}_n, \quad \mathrm{Cov}\{\boldsymbol{q}_n\} = \boldsymbol{Q}_n$$

- Given: $\mathrm{E}\{\boldsymbol{x}_{n-1} \mid \boldsymbol{y}_{1:n-1}\} = \hat{\boldsymbol{x}}_{n-1|n-1}$, $\mathrm{Cov}\{\boldsymbol{x}_{n-1} \mid \boldsymbol{y}_{1:n-1} =\}\boldsymbol{P}_{n-1|n-1}$
- Predicted mean:

$$\hat{\boldsymbol{x}}_{n|n-1} = \mathrm{E}\{\boldsymbol{x}_n \mid \boldsymbol{y}_{1:n-1}\}$$

Linear Model: Prediction (2/2)

- Linear dynamic model:

$$\boldsymbol{x}_n = \boldsymbol{F}_n \boldsymbol{x}_{n-1} + \boldsymbol{q}_n, \quad \mathrm{Cov}\{\boldsymbol{q}_n\} = \boldsymbol{Q}_n$$

- Given: $\mathrm{E}\{\boldsymbol{x}_{n-1} \mid \boldsymbol{y}_{1:n-1}\} = \hat{\boldsymbol{x}}_{n-1|n-1}$, $\mathrm{Cov}\{\boldsymbol{x}_{n-1} \mid \boldsymbol{y}_{1:n-1} =\}\boldsymbol{P}_{n-1|n-1}$
- Covariance:

$$\begin{aligned}\boldsymbol{P}_{n|n-1} &= \mathrm{Cov}\{\boldsymbol{x}_n \mid \boldsymbol{y}_{1:n-1}\} \\ &= \mathrm{E}\{(\boldsymbol{x}_n - \mathrm{E}\{\boldsymbol{x}_n \mid \boldsymbol{y}_{1:n-1}\})(\boldsymbol{x}_n - \mathrm{E}\{\boldsymbol{x}_n \mid \boldsymbol{y}_{1:n-1}\})^{\mathsf{T}} \mid \boldsymbol{y}_{1:n-1}\}\end{aligned}$$

Linear Model: Summary

- Prediction:

$$
\begin{aligned}
\hat{\boldsymbol{x}}_{n|n-1} &= \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1|n-1} \\
\boldsymbol{P}_{n|n-1} &= \boldsymbol{F}_n \boldsymbol{P}_{n-1|n-1} \boldsymbol{F}_n^\mathsf{T} + \boldsymbol{Q}_n
\end{aligned}
$$

- Measurement update:

$$
\begin{aligned}
\boldsymbol{K}_n &= \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^\mathsf{T} (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^\mathsf{T} + \boldsymbol{R}_n)^{-1} \\
\hat{\boldsymbol{x}}_{n|n} &= \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n (\boldsymbol{y}_n - \boldsymbol{G}_n \hat{\boldsymbol{x}}_{n|n-1}) \\
\boldsymbol{P}_{n|n} &= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^\mathsf{T} + \boldsymbol{R}_n) \boldsymbol{K}_n^\mathsf{T}
\end{aligned}
$$

- Initialization of the recursion:

$$
\begin{aligned}
\hat{\boldsymbol{x}}_{0|0} &= \boldsymbol{m}_0 \\
\boldsymbol{P}_{0|0} &= \boldsymbol{P}_0
\end{aligned}
$$

The Kalman Filter

Algorithm 1 Kalman Filter

1: Initialize $\hat{\boldsymbol{x}}_{0|0} = \boldsymbol{m}_0$, $\boldsymbol{P}_{0|0} = \boldsymbol{P}_0$

2: **for** $n = 1, 2, \dots$ **do**

3: Prediction (time update):

$$\begin{aligned}\hat{\boldsymbol{x}}_{n|n-1} &= \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1|n-1} \\ \boldsymbol{P}_{n|n-1} &= \boldsymbol{F}_n \boldsymbol{P}_{n-1|n-1} \boldsymbol{F}_n^{\mathsf{T}} + \boldsymbol{Q}_n\end{aligned}$$

4: Measurement update:

$$\begin{aligned}\boldsymbol{K}_n &= \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^{\mathsf{T}} (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n + \boldsymbol{R}_n)^{-1} \\ \hat{\boldsymbol{x}}_{n|n} &= \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n (\boldsymbol{y}_n - \boldsymbol{G}_n \hat{\boldsymbol{x}}_{n|n-1}) \\ \boldsymbol{P}_{n|n} &= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n + \boldsymbol{R}_n) \boldsymbol{K}_n^{\mathsf{T}}\end{aligned}$$

5: **end for**

Example: GPS Navigation (1/3)

- Goal: Estimate the kinematic state at each time t_n
- Dynamic model: 2D Wiener velocity model:

$$\begin{bmatrix} \dot{p}^x(t) \\ \dot{p}^y(t) \\ \dot{v}^x(t) \\ \dot{v}^y(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p^x(t) \\ p^y(t) \\ v^x(t) \\ v^y(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$$

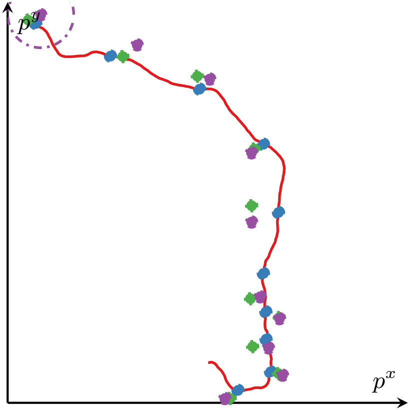
- Measurements: Noisy position ($\boldsymbol{G}_n = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \end{bmatrix}$):

$$\boldsymbol{y}_n = \boldsymbol{G}_n \boldsymbol{x}_n + \boldsymbol{r}_n$$

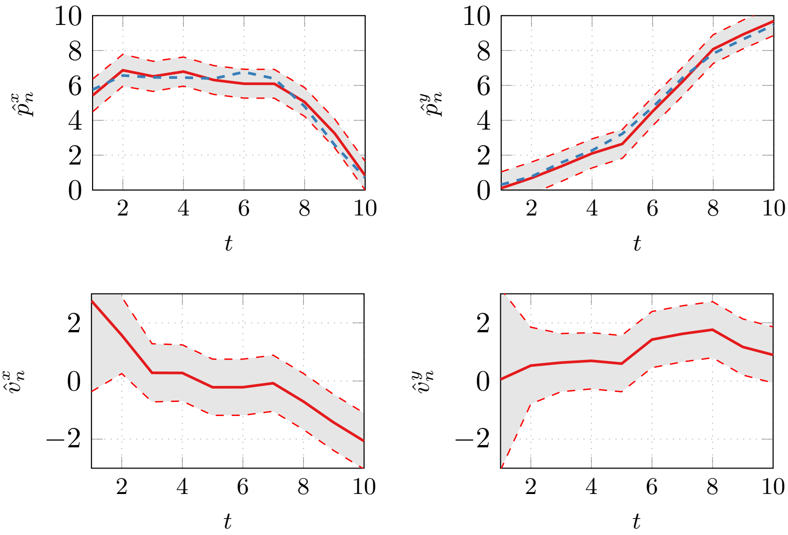
- Discrete-time linear state-space model:

$$\begin{aligned} \boldsymbol{x}_n &= \boldsymbol{F}_n \boldsymbol{x}_{n-1} + \boldsymbol{q}_n \\ \boldsymbol{y}_n &= \boldsymbol{G}_n \boldsymbol{x}_n + \boldsymbol{r}_n \end{aligned}$$

Example: GPS Navigation (2/3)



Example: GPS Navigation (3/3)



Performance Evaluation

A few questions:

- Why not just use the measurements $\boldsymbol{y}_n$ as the position estimate ($\hat{\boldsymbol{p}}_n = \boldsymbol{y}_n$)?
- How should we assess the performance of the algorithm?
- One possible criterion: The root mean squared error (RMSE):

$$e_{\mathsf{RMSE}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\hat{\boldsymbol{x}}_{n|n} - \boldsymbol{x}_n)^{\mathsf{T}}(\hat{\boldsymbol{x}}_{n|n} - \boldsymbol{x}_n)}$$

Example: GPS Navigation

- RMSE for the primitive approach ($\hat{\boldsymbol{p}}_n = \boldsymbol{y}_n$):

$$e_{\text{RMSE}} = 0.41$$

- RMSE for Kalman filter:

$$e_{\text{RMSE}} = 0.29$$

- The prior knowledge imposed by the dynamic model significantly improves performance!

Measurement Update: Some Observations

- Measurement update:

$$
\begin{aligned}
\boldsymbol{K}_n &= \boldsymbol{P}_{n|n-1}\boldsymbol{G}_n^{\mathsf{T}}(\boldsymbol{G}_n\boldsymbol{P}_{n|n-1}\boldsymbol{G}_n + \boldsymbol{R}_n)^{-1} \\
\hat{\boldsymbol{x}}_{n|n} &= \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n(\boldsymbol{y}_n - \boldsymbol{G}_n\hat{\boldsymbol{x}}_{n|n-1}) \\
\boldsymbol{P}_{n|n} &= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n(\boldsymbol{G}_n\boldsymbol{P}_{n|n-1}\boldsymbol{G}_n + \boldsymbol{R}_n)\boldsymbol{K}_n^{\mathsf{T}}
\end{aligned}
$$

- Prediction of the output and covariances:

$$
\begin{aligned}
\mathrm{E}\{\boldsymbol{y}_n \mid \boldsymbol{y}_{1:n-1}\} &= \boldsymbol{G}_n\hat{\boldsymbol{x}}_{n|n-1} \\
\mathrm{Cov}\{\boldsymbol{y}_n \mid \boldsymbol{y}_{1:n-1}\} &= \boldsymbol{G}_n\boldsymbol{P}_{n|n-1}\boldsymbol{G}_n^{\mathsf{T}} + \boldsymbol{R}_n \\
\mathrm{Cov}\{\boldsymbol{x}_n, \boldsymbol{y}_n \mid \boldsymbol{y}_{1:n-1}\} &= \boldsymbol{P}_{n|n-1}\boldsymbol{G}_n^{\mathsf{T}}
\end{aligned}
$$

Summary

- The filtering approach iterates between two steps:
  1. Prediction: $\hat{\boldsymbol{x}}_{n-1|n-1}$, $\boldsymbol{P}_{n-1|n-1} \Rightarrow \hat{x}_{n|n-1}$, $\boldsymbol{P}_{n|n-1}$
  2. Measurement update: $\hat{x}_{n|n-1}$, $\boldsymbol{P}_{n|n-1} \Rightarrow \hat{x}_{n|n}$, $\boldsymbol{P}_{n|n}$
- The Kalman filter is the optimal filter for linear state-space models
  1. Prediction:

$$\begin{aligned} \hat{\boldsymbol{x}}_{n|n-1} &= \boldsymbol{F}_n \hat{\boldsymbol{x}}_{n-1|n-1} \\ \boldsymbol{P}_{n|n-1} &= \boldsymbol{F}_n \boldsymbol{P}_{n-1|n-1} \boldsymbol{F}_n^\mathsf{T} + \boldsymbol{Q}_n \end{aligned}$$

  2. Measurement update:

$$\begin{aligned} \boldsymbol{K}_n &= \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n^\mathsf{T} (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n + \boldsymbol{R}_n)^{-1} \\ \hat{\boldsymbol{x}}_{n|n} &= \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n (\boldsymbol{y}_n - \boldsymbol{G}_n \hat{\boldsymbol{x}}_{n|n-1}) \\ \boldsymbol{P}_{n|n} &= \boldsymbol{P}_{n|n-1} - \boldsymbol{K}_n (\boldsymbol{G}_n \boldsymbol{P}_{n|n-1} \boldsymbol{G}_n + \boldsymbol{R}_n) \boldsymbol{K}_n^\mathsf{T} \end{aligned}$$