University of Michigan - NAME 568/EECS 568/ROB 530 **Winter 2022**

Note: Kalman Filtering

Lecturer and Author: Maani Ghaffari

1 Discrete-Time Gauss-Markov Process

The Markov property states that "the future is independent of the past if the present is known." A stochastic process that has this property is called a Markov process. The state of a dynamic system driven by white noise is a Markov process.

A discrete-time random process (random sequence), w_k, is called white noise if:

$$\mathbb{E}[w_k w_j^\mathsf{T}] = Q_k \delta_{kj},$$

where the Kronecker δ_{kj} is

$$\delta_{kj} = \begin{cases} 1 & \text{if } k = j \\ 0 & \text{if } k \neq j \end{cases}.$$

The state of a dynamic system excited by white noise,

$$x_{k+1} = f(x_k, w_k),$$

is a discrete-time Markov process or Markov sequence. The state of a linear dynamic system excited by white Gaussian noise,

$$x_{k+1} = F x_k + w_k,$$

is called a Gauss-Markov process. Assuming the initial condition is Gaussian, because of linearity x_k is Gaussian and because of the whiteness of the process noise it is Markov.

2 Kalman Filter Assumptions

The state, x_k, evolves according to a known linear dynamic equation with known inputs, u_k, an additive process noise, w_k, which is a zero-mean white (uncorrelated) process with known covariance Q_k;

$$x_k^- = F_k x_{k-1} + G_k u_k + w_k. \tag{1}$$

The measurement model is a known linear function of the state with an additive measurement noise, v_k, which is a zero-mean white (uncorrelated) process with known covariance R_k;

$$z_k = H_k x_k^- + v_k. \tag{2}$$

Initial state is assumed to be a random variable with known mean (initial estimate) and known covariance (initial uncertainty). Initial state and noises are all mutually uncorrelated.

In summary:

- Known initial state x_0 (with possibly given prior information z_0): mean $\mathbb{E}[x_0|z_0] = \hat{x}_0$ and covariance $\mathrm{Cov}[x_0|z_0] = P_0$.
- Process and measurement noise sequences are white with known covariances:

$$\mathbb{E}[w_k] = 0,\ \mathbb{E}[w_k w_j^{\mathsf{T}}] = Q_k \delta_{kj}, \quad \text{and} \quad \mathbb{E}[w_k] = 0,\ \mathbb{E}[v_k v_j^{\mathsf{T}}] = R_k \delta_{kj}.$$

- All the above are uncorrelated.

3 Probabilistic Derivation for the Multivariate Gaussian Case

In the prediction step, we use the joint distribution of the state at time steps k and $k-1$, $p(x_k, x_{k-1}|u_{1:k}, z_{1:k-1})$, and marginalize x_{k-1} as follows:

$$\underbrace{\begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix}}_{y_k} = \underbrace{\begin{bmatrix} F_k \\ I \end{bmatrix}}_{A_k} x_{k-1} + \underbrace{\begin{bmatrix} G_k \\ 0 \end{bmatrix}}_{B_k} u_k + \underbrace{\begin{bmatrix} I \\ 0 \end{bmatrix}}_{W_k} w_k,$$

$$y_k = A_k x_{k-1} + B_k u_k + W_k w_k.$$

We compute the mean and covariance as:

$$\mathbb{E}[y_k] = \mathbb{E}[A_k x_{k-1} + B_k u_k + W_k w_k] = A_k \mu_{k-1} + B_k u_k = \begin{bmatrix} F_k \mu_{k-1} + G_k u_k \\ \mu_{k-1} \end{bmatrix},$$

$$\begin{aligned} \mathrm{Cov}[y_k] &= \mathbb{E}[(y_k - \mathbb{E}[y_k])(y_k - \mathbb{E}[y_k])^{\mathsf{T}}] \\ &= \mathbb{E}[(A_k(x_{k-1} - \mu_{k-1}) + W_k w_k)(A_k(x_{k-1} - \mu_{k-1}) + W_k w_k)^{\mathsf{T}}] \\ &= A_k \Sigma_{k-1} A_k^{\mathsf{T}} + W_k Q_k W_k^{\mathsf{T}} = \begin{bmatrix} F_k \Sigma_{k-1} F_k^{\mathsf{T}} + Q_k & F_k \Sigma_{k-1} \\ \Sigma_{k-1} F_k^{\mathsf{T}} & \Sigma_{k-1} \end{bmatrix}. \end{aligned}$$

Using the marginalization property of jointly Gaussian random vectors we have:

$$\mathbb{E}[x_k] =: \mu_k^- = F_k \mu_{k-1} + G_k u_k \tag{3}$$

$$\mathrm{Cov}[x_k] =: \Sigma_k^- = F_k \Sigma_{k-1} F_k^{\mathsf{T}} + Q_k \tag{4}$$

In the correction step, we form the joint distribution $p(x_k, z_k|u_{1:k}, z_{1:t-1})$ and then condition on z_k.

$$\underbrace{\begin{bmatrix} x_k \\ z_k \end{bmatrix}}_{s_k} = \underbrace{\begin{bmatrix} I \\ H_k \end{bmatrix}}_{C_k} x_k + \underbrace{\begin{bmatrix} 0 \\ I \end{bmatrix}}_{V_k} v_k,$$

$$s_k = C_k x_k + V_k v_k.$$

$$\mathbb{E}[s_k] = \mathbb{E}[C_k x_k + V_k v_k] = C_k \mu_k^- = \begin{bmatrix} \mu_k^- \\ H_k \mu_k^- \end{bmatrix},$$

$$\begin{aligned} \mathrm{Cov}[s_k] &= \mathbb{E}[(s_k - \mathbb{E}[s_k])(s_k - \mathbb{E}[s_k])^\mathsf{T}] \\ &= C_k \Sigma_k^- C_k^\mathsf{T} + V_k R_k V_k^\mathsf{T} = \begin{bmatrix} \Sigma_k^- & \Sigma_k^- H_k^\mathsf{T} \\ H_k \Sigma_k^- & H_k \Sigma_k^- H_k^\mathsf{T} + R_k \end{bmatrix}. \end{aligned}$$

Using the conditioning property of jointly Gaussian random vectors we have:

$$\mathbb{E}[x_k | z_k] =: \mu_k = \mu_k^- + \Sigma_k^- H_k^\mathsf{T} (H_k \Sigma_k^- H_k^\mathsf{T} + R_k)^{-1} (z_k - H_k \mu_k^-), \tag{5}$$

$$\begin{aligned} \mathrm{Cov}[x_k | z_k] =: \Sigma_k &= \Sigma_k^- - \Sigma_k^- H_k^\mathsf{T} (H_k \Sigma_k^- H_k^\mathsf{T} + R_k)^{-1} H_k \Sigma_k^- \\ &= (I - \Sigma_k^- H_k^\mathsf{T} (H_k \Sigma_k^- H_k^\mathsf{T} + R_k)^{-1} H_k) \Sigma_k^-. \end{aligned} \tag{6}$$

4 Kalman Filter Algorithm

Algorithm 1 `Kalman-filter`

Require: belief mean μ_{k-1}, belief covariance Σ_{k-1}, action u_k, measurement z_k;

$\mu_k^- \leftarrow F_k \mu_{k-1} + G_k u_k$ ▷ predicted mean

$\Sigma_k^- \leftarrow F_k \Sigma_{k-1} F_k^\mathsf{T} + Q_k$ ▷ predicted covariance

$\nu_k \leftarrow z_k - H_k \mu_k^-$ ▷ innovation

$S_k \leftarrow H_k \Sigma_k^- H_k^\mathsf{T} + R_k$ ▷ innovation covariance

$K_k \leftarrow \Sigma_k^- H_k^\mathsf{T} S_k^{-1}$ ▷ filter gain

$\mu_k \leftarrow \mu_k^- + K_k \nu_k$ ▷ corrected mean

$\Sigma_k \leftarrow (I - K_k H_k) \Sigma_k^-$ ▷ corrected covariance

$//\Sigma_k \leftarrow (I - K_k H_k) \Sigma_k^- (I - K_k H_k)^\mathsf{T} + K_k R_k K_k^\mathsf{T}$ ▷ numerically stable form

return μ_k, Σ_k

5 Overview and Limitations of Kalman Filter

Under the Gaussian assumption for the initial state (or initial state error) and all the noises entering into the system, the Kalman filter is the optimal Minimum Mean Square Error (MMSE) state estimator. In other words, for a linear system with Gaussian noise, the Kalman filter equations are the Best Linear Unbiased Estimator (BLUE); this means they are performing right at the Cramér-Rao lower bound [1]. If these random

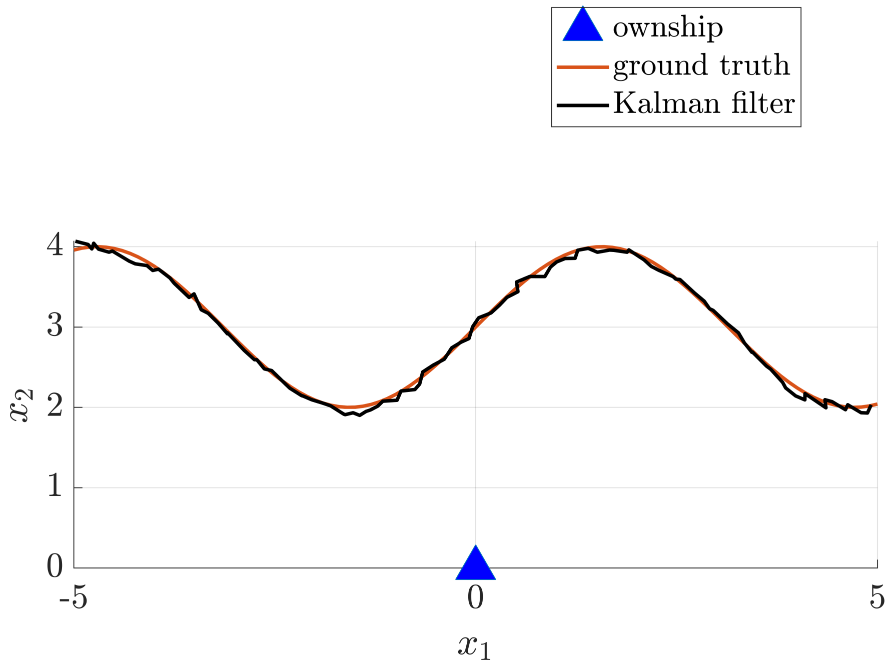

Figure 1: 2D target tracking using a Kalman filter.

variables are not Gaussian and one has only their first two moments, then the Kalman filter algorithm is the best linear state estimator in the sense of MMSE [2, page 207].

In practice, we come across many interesting problems that violate our assumptions. These are include:

- nonlinear motion (process) and measurement models;
- unknown control inputs or mode changes;
- data association uncertainty;
- auto-correlated or cross-correlated noise sequences.

Example 1 (Kalman Filter Target Tracking)**.** *A target is moving in a 2D plane. The ownship position is known and fixed at the origin. We have access to noisy measurements that directly observe the target 2D coordinates at any time step.*

$$F_k = I_2,\ G_k = 0_{2\times 2},\ H_k = I_2,\ Q_k = 0.001\, I_2,\ R_k = 0.05^2\, I_2$$

We estimate the target position using a Kalman filter as shown in Figure 1. See `kf_single_target.m` *(or Python version) for code. Try to change the parameters and study the filter's behavior.*

References

[1] T. D. Barfoot, *State estimation for robotics*. Cambridge University Press, 2017.

[2] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2001.