

ARO Lecture notes - Kalman Filter

Ondřej Franek

Prerequisites

Def: Event \mathbf{A} is *conditionally independent* on \mathbf{C} given \mathbf{B} iff $p(\mathbf{A} | \mathbf{B}, \mathbf{C}) = p(\mathbf{A} | \mathbf{B})$

Def: State \mathbf{x}_{t-1} is *complete* iff future \mathbf{x}_t is conditionally independent on past given \mathbf{x}_{t-1}

Consequences:

- State-transition probability:

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) = p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t, x_{0:t-2}, z_{1:t-1}, u_{1:t-1}) \quad (1)$$

- Measurement probability:

$$p(\mathbf{z}_t | \mathbf{x}_t) = p(\mathbf{z}_t | \mathbf{x}_t, x_{0:t-1}, z_{1:t-1}, u_{1:t}) \quad (2)$$

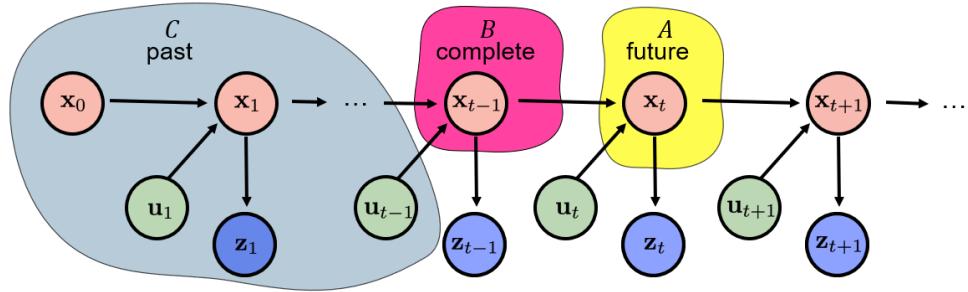


Figure 1: Complete state illustration.

Bayes Filter

The Bayes filter is an algorithm for probabilistic state estimation in dynamic systems. It predicts the state of a system over time, given control inputs and sensory measurements. The filter performs two main steps in a repetitive manner: prediction and measurement update.

- **Initialization:** The belief of the initial state is represented by $bel(\mathbf{x}_0)$ at time $t = 1$.
- **Prediction Step:** Given the control action u_t at time t , the prediction step computes a prior belief state:

$$\overline{bel}(\mathbf{x}_t) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) bel(\mathbf{x}_{t-1}) d\mathbf{x}_{t-1} \quad (3)$$

- **Measurement Update:** Upon receiving a new measurement z_t , the measurement update refines the predicted belief to produce a posterior belief state:

$$bel(\mathbf{x}_t) = \eta p(\mathbf{z}_t | \mathbf{x}_t) \overline{bel}(\mathbf{x}_t) \quad (4)$$

where η is a normalization constant that ensures the posterior belief is a valid probability distribution by integrating to 1.

Kalman Filter

The Kalman filter is a powerful algorithm used for estimating the state of a system over time. It operates on a simple yet effective principle: predict the future state, then update this prediction with new observations. It handles systems with uncertainties and noise quickly and effectively, making it widely utilized.

Linear System with Gaussian Noise

The Kalman filter assumes a linear system with Gaussian noise.

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t + \mathbf{w}_t \quad (5)$$

$$\mathbf{z}_t = \mathbf{C}_t \mathbf{x}_t + \mathbf{v}_t \quad (6)$$

where \mathbf{A}_t , \mathbf{B}_t , \mathbf{C}_t are matrices that define the system dynamics, \mathbf{u}_t is the control input, \mathbf{x}_t is state vector and $\mathbf{w} \sim \mathcal{N}(0, \mathbf{R}_t)$ and $\mathbf{v} \sim \mathcal{N}(0, \mathbf{Q}_t)$ are the transition and measurement noise with covariance matrices \mathbf{R}_t and \mathbf{Q}_t , respectively. This implies that all probability distributions involved are Gaussian, which simplifies computation. The state-transition and measurement probabilities are modeled as follows:

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) = \mathcal{N}(\mathbf{x}_t; \mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t, \mathbf{R}_t) \quad (7)$$

$$p(\mathbf{z}_t | \mathbf{x}_t) = \mathcal{N}(\mathbf{z}_t; \mathbf{C}_t \mathbf{x}_t, \mathbf{Q}_t) \quad (8)$$

Implementation of Kalman Filter

Combining the principles of the Bayes filter (equations 3 and 4), with prerequisites (equations 1 and 2) and the properties of linear Gaussian systems (equations 7 and 8), we can derive the Kalman filter update equations:

- **Prediction step** (new action \mathbf{u}_t):

$$\bar{\boldsymbol{\mu}}_t = \mathbf{A}_t \boldsymbol{\mu}_{t-1} + \mathbf{B}_t \mathbf{u}_t \quad (9)$$

$$\bar{\boldsymbol{\Sigma}}_t = \mathbf{A}_t \boldsymbol{\Sigma}_{t-1} \mathbf{A}_t^\top + \mathbf{R}_t \quad (10)$$

$$\overline{\text{bel}}(\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_t, \bar{\boldsymbol{\Sigma}}_t) \quad (11)$$

- **Measurement update** (new measurement \mathbf{z}_t):

$$\mathbf{K}_t = \bar{\boldsymbol{\Sigma}}_t \mathbf{C}_t^\top (\mathbf{C}_t \bar{\boldsymbol{\Sigma}}_t \mathbf{C}_t^\top + \mathbf{Q}_t)^{-1} \quad (12)$$

$$\boldsymbol{\mu}_t = \bar{\boldsymbol{\mu}}_t + \mathbf{K}_t (\mathbf{z}_t - \mathbf{C}_t \bar{\boldsymbol{\mu}}_t) \quad (13)$$

$$\boldsymbol{\Sigma}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{C}_t) \bar{\boldsymbol{\Sigma}}_t \quad (14)$$

$$\text{bel}(\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) \quad (15)$$

where $\boldsymbol{\mu}_t$ is the estimated mean, $\boldsymbol{\Sigma}_t$ is the estimated uncertainty (covariance), and \mathbf{K}_t is the Kalman gain which minimizes the estimated uncertainty.

In the following picture 2, we can see the propagation of the state estimate and its uncertainty through a linear function in the prediction step. The **Green Gaussian** represents the prior belief, and the **Red Gaussian** represents the posterior belief.

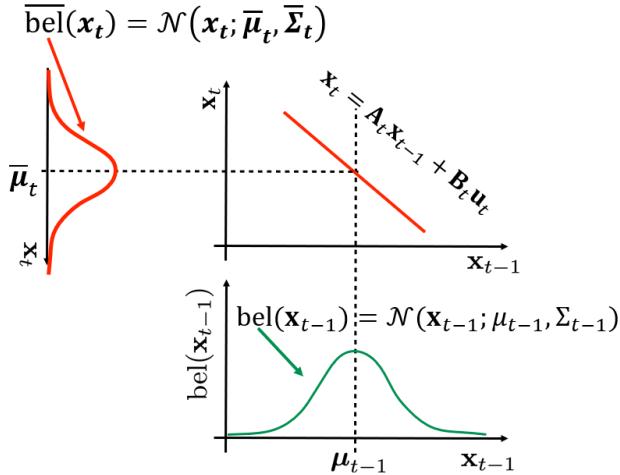


Figure 2: Linear system propagation.

Kalman Filter Example: Squeezes Gaussian

In this example of a Kalman Filter, the state of a system is represented by position and velocity of the robot. The state can be expressed as a vector $\mathbf{x} = \begin{bmatrix} x \\ v \end{bmatrix}$. In this example the state-transition probability is defined as

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) = \mathcal{N}(\mathbf{x}_t; \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{\mathbf{A}_t} \mathbf{x}_{t-1} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{B}_t} \mathbf{u}_t, \underbrace{\begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}}_{\mathbf{R}_t})$$

and measurement probability is

$$p(\mathbf{z}_t | \mathbf{x}_t) = \mathcal{N}(\mathbf{z}_t; \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{\mathbf{C}_t} \mathbf{x}_{t-1}, \underbrace{\begin{bmatrix} 0.3 \end{bmatrix}}_{\mathbf{Q}_t}).$$

When there are no measurements available the probability distribution (modeled as a Gaussian) tends to become more squeezed and skewed over time. This is due to the growing uncertainty in predictions as time progresses. Additionally, the skewing of the distribution is caused by the linear dependence of position on velocity. The sequence of images (Figure 3) demonstrates how the distribution evolves from x_0 to x_5 , visually representing the increasing uncertainty at different time steps.

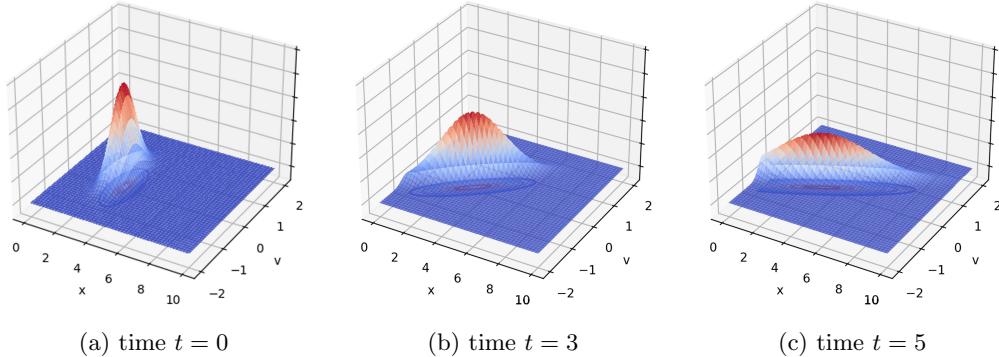


Figure 3: Visualization of state squeezing and skewing at different time steps.

If the robot receives some measurements, the probability distribution undergoes a significant correction. Each measurement provides additional information that helps refine the estimates of the robot's state, reducing uncertainty and compensating for the previous skewing and squeezing effects. The corrected probability Gaussian is depicted in Figure 4.

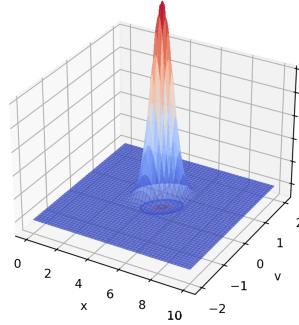


Figure 4: Probability after the measurement step.

Conclusion

The Kalman Filter is a fundamental tool in control systems and signal processing, efficiently estimating states in linear systems with Gaussian noise. Its recursive structure is well-suited for real-time applications in various fields like navigation and tracking. However, the effectiveness of the Kalman Filter is constrained by its assumption that both system dynamics and measurement functions are linear. This limitation restricts its applicability to linear systems, as it struggles to accurately address deviations arising from non-linear behaviors in real-world scenarios. To manage non-linearities, adaptations such as the Extended Kalman Filter or alternative approaches like the Particle Filter are commonly employed.

Extended Kalman Filter

The Extended Kalman Filter (EKF) is an adaptation of the Kalman Filter for non-linear systems. While the Kalman Filter is restricted to linear models, the EKF allows for a broader range of applications by linearizing about the current estimate.

Non-Linear System with Gaussian Noise

The state and measurement models for a non-linear system can be expressed as:

$$\mathbf{x}_t = \mathbf{g}(\mathbf{x}_{t-1}, \mathbf{u}_t) + \mathbf{w} \quad (16)$$

$$\mathbf{z}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{v} \quad (17)$$

where \mathbf{g} and \mathbf{h} are non-linear functions of the state and control inputs, and $\mathbf{w} \sim \mathcal{N}(0, \mathbf{R}_t)$ and $\mathbf{v} \sim \mathcal{N}(0, \mathbf{Q}_t)$ represent the transition and measurement noise with covariance matrices \mathbf{R}_t and \mathbf{Q}_t , respectively.

In the following picture we can see p

Linearization

The EKF approximates these non-linear functions using the first-order Taylor expansion around the current estimate:

$$\mathbf{g}(\mathbf{u}_t, \mathbf{x}_{t-1}) \approx \mathbf{g}(\mathbf{u}_t, \boldsymbol{\mu}_{t-1}) + \mathbf{G}_t(\mathbf{x}_{t-1} - \boldsymbol{\mu}_{t-1}) \quad (18)$$

$$\mathbf{h}(\mathbf{x}_t) \approx \mathbf{h}(\boldsymbol{\mu}_t) + \mathbf{H}_t(\mathbf{x}_t - \boldsymbol{\mu}_t), \quad (19)$$

where \mathbf{G}_t and \mathbf{H}_t are defined as follows:

$$\mathbf{G}_t = \frac{\partial \mathbf{g}(\mathbf{x} = \boldsymbol{\mu}_{t-1}, \mathbf{u} = \mathbf{u}_t)}{\partial \mathbf{x}} \quad (20)$$

$$\mathbf{H}_t = \frac{\partial \mathbf{h}(\mathbf{x} = \boldsymbol{\mu}_t)}{\partial \mathbf{x}} \quad (21)$$

The state-transition and measurement probabilities are modeled as follows:

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) \approx \mathcal{N}(\mathbf{x}_t; \mathbf{g}(\boldsymbol{\mu}_{t-1}, \mathbf{u}_t) + \mathbf{G}_t(\mathbf{x}_{t-1} - \boldsymbol{\mu}_{t-1}), \mathbf{R}_t) \quad (22)$$

$$p(\mathbf{z}_t | \mathbf{x}_t) \approx \mathcal{N}(\mathbf{z}_t; \mathbf{h}(\bar{\boldsymbol{\mu}}_t) + \mathbf{H}_t(\mathbf{x}_t - \bar{\boldsymbol{\mu}}_t), \mathbf{Q}_t) \quad (23)$$

In the following figures, 5 and 6, we can observe the propagation of the prior belief, represented by the **Green Gaussian**, through a non-linear system (Figure 5) and a linearized system (Figure 6).

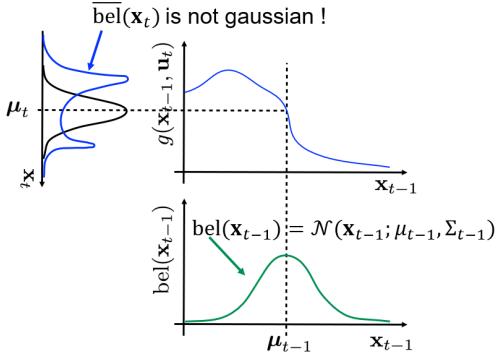


Figure 5: Non-linear system propagation.

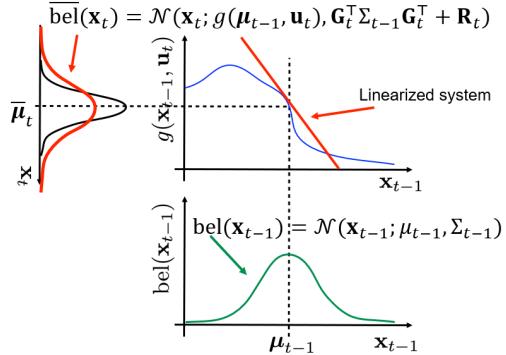


Figure 6: Linearized system propagation.

Implementation of Extended Kalman Filter

Combining the principles of the Bayes filter (equations 3 and 4), with prerequisites (equations 1 and 2) and the properties of linearized Gaussian systems (equations 22 and 23), we can derive the Kalman filter update equations:

- **Prediction step** (new action \mathbf{u}_t):

$$\bar{\boldsymbol{\mu}}_t = \mathbf{g}(\boldsymbol{\mu}_{t-1}, \mathbf{u}_t) \quad (24)$$

$$\bar{\Sigma}_t = \mathbf{G}_t \Sigma_{t-1} \mathbf{G}_t^\top + \mathbf{R}_t \quad (25)$$

$$\overline{\text{bel}}(\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_t; \bar{\boldsymbol{\mu}}_t, \bar{\Sigma}_t) \quad (26)$$

- **Measurement update** (new measurement \mathbf{z}_t):

$$\mathbf{K}_t = \bar{\Sigma}_t \mathbf{H}_t^\top (\mathbf{H}_t \bar{\Sigma}_t \mathbf{H}_t^\top + \mathbf{Q}_t)^{-1} \quad (27)$$

$$\boldsymbol{\mu}_t = \bar{\boldsymbol{\mu}}_t + \mathbf{K}_t (\mathbf{z}_t - \mathbf{h}(\bar{\boldsymbol{\mu}}_t)) \quad (28)$$

$$\Sigma_t = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \bar{\Sigma}_t \quad (29)$$

$$\text{bel}(\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_t, \Sigma_t) \quad (30)$$

Conclusion

The Extended Kalman Filter extends the Kalman Filter to non-linear systems by linearizing about the current state estimate. However, this approach introduces errors due to approximation, particularly when the system dynamics or measurements are highly non-linear. These errors can lead to inaccuracies and potential divergence of the filter, especially with poor initial estimates or significant non-linear deviations. Thus, for systems with strong non-linearities, more robust methods like the Unscented Kalman Filter or Particle Filter may provide better performance