Statistics and Sensor Data Fusion

6. The Kalman Filter

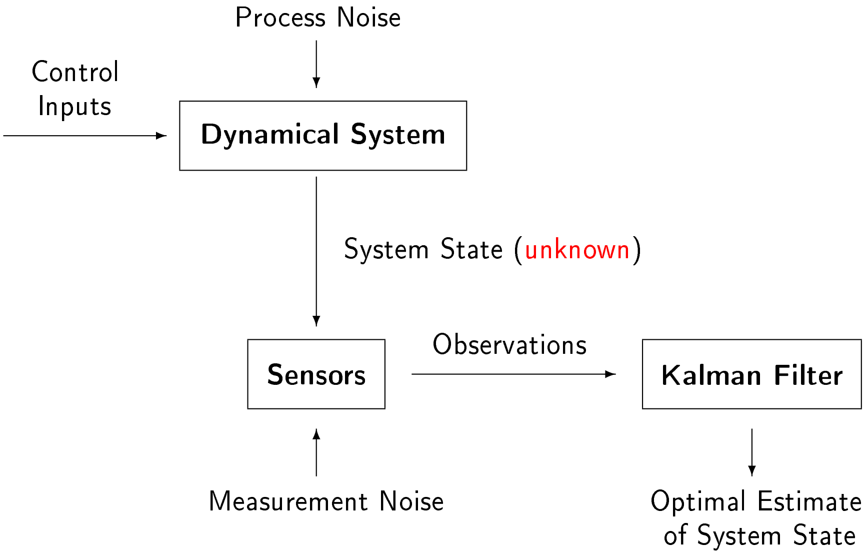
The Kalman Filter

Goal: State estimation of dynamical systems

6.1 State Space Description of Dynamical Systems

State Space Description of Dynamical Systems

State Estimation of Dynamical Systems:



State Space Description of Dynamical Systems

State Space Description:

- In order to control a **physical system** of any kind, whether it be a vehicle, a chemical process, or the national economy, one attempts to develop a **mathematical model** that represents the relevant aspects of system behaviour.
- The mathematical model which captures the interrelationships between certain variables of interest and their development in time is called a dynamical system.
- In order to observe the actual system behaviour, measurement devices are applied to generate output signals or observations related to the variables of interest.
- These observations and possible control inputs to the system are the only information available about the actual system state which itself is unknown.

State Space Description of Dynamical Systems

State Space Description:

The state of a dynamical system is usually defined as a **minimum set of variables** describing the behaviour of the system.

In the case that unpredictable random effects are involved, the system state is modeled by a time-dependent **random vector**

$$\boldsymbol{X}(t) = (X_1(t), \ldots, X_n(t))^T$$

Since frequently one is interested in the state of the system only at **discrete points in time** given by $t_k = t_0 + k \cdot \Delta t$ with $k \in \mathbb{N}$, we will denote the time-dependent state vector at time t_k according to

$$\boldsymbol{X}_k = \boldsymbol{X}(t_k), \quad k \in \mathbb{N}$$

The time evolution of the state vector $\boldsymbol{X}_k$ is determined by the so-called process model in terms of an **equation of motion**.

State Space Description of Dynamical Systems

State Space Description – Process Model:

Assumig the special case of a stochastic time-variant linear system, the process model is given by a **difference equation** according to

$$\boldsymbol{X}_k = \boldsymbol{F}_{k-1} \cdot \boldsymbol{X}_{k-1} + \boldsymbol{B}_{k-1} \cdot \boldsymbol{u}_{k-1} + \boldsymbol{W}_{k-1}, \quad k \in \mathbb{N}$$

In the above difference equation

- the matrix $\boldsymbol{F}_{k-1} \in \mathbb{R}^{n \times n}$ denotes the state transition matrix describing the transition between the state $\boldsymbol{X}_{k-1}$ and $\boldsymbol{X}_k$
- the vector $\boldsymbol{u}_{k-1} \in \mathbb{R}^{\ell}$ denotes the control vector at time $k-1$
- the matrix $\boldsymbol{B}_{k-1}$ denotes the control matrix describing the influence of the control vector $\boldsymbol{u}_{k-1}$ on the system state $\boldsymbol{X}_k$
- the random vector $\boldsymbol{W}_{k-1} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_{k-1})$ denotes the process noise vector at time $k-1$, capturing the uncertainties present in the process model (i.e. the deviations from reality)

State Space Description of Dynamical Systems

State Space Description – Measurement Model:

The system state $\boldsymbol{X}_k$ cannot be observed directly, instead an array of measurement devices or sensors generate **observations** $\boldsymbol{Z}_k$ which are linked to the system state $\boldsymbol{X}_k$ via the measurement model:

$$\boldsymbol{Z}_k = \boldsymbol{H}_k \cdot \boldsymbol{X}_k + \boldsymbol{V}_k, \quad k \in \mathbb{N}$$

In this equation

- the random vector $\boldsymbol{Z}_k$ denotes the observation vector at time k
- the matrix $\boldsymbol{H}_k \in \mathbb{R}^{m \times n}$ denotes the observation matrix which captures predictable distortions of the sensors at time k
- the random vector $\boldsymbol{V}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_k)$ denotes the measurement noise vector at time k

It is assumed that the process noise vector $\boldsymbol{W}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$ and the measurement noise vector $\boldsymbol{V}_k$ are **uncorrelated** for all times.

State Space Description of Dynamical Systems

State Space Description – Linear Time-Invariant System:

In the special case of a stochastic time-invariant linear system, the state space description consisting of both the process model and the measurement model is given by the equations

$$\boldsymbol{X}_k = \boldsymbol{F} \cdot \boldsymbol{X}_{k-1} + \boldsymbol{B} \cdot \boldsymbol{u}_{k-1} + \boldsymbol{W}, \quad k \in \mathbb{N}$$

$$\boldsymbol{Z}_k = \boldsymbol{H} \cdot \boldsymbol{X}_k + \boldsymbol{V}, \quad k \in \mathbb{N}$$

Here, the state transition matrix $\boldsymbol{F}$, the control matrix $\boldsymbol{B}$ as well as the observation matrix $\boldsymbol{H}$ do not change with the time $k \in \mathbb{N}$ but remain **constant** (therefore time-invariant).

Additionally, both the process noise vector $\boldsymbol{W} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$ and the measurement noise vector $\boldsymbol{V} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$ are characterized by **constant** covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively.

State Space Description of Dynamical Systems

The Filtering Problem:

In practical scenarios, sensor measurements at time k only provide a **single realization** $\boldsymbol{z}_k$ of the random observation vector $\boldsymbol{Z}_k$.

The problem of finding the **probability distribution** of the state vector $\boldsymbol{X}_k$ based on the series of measurements $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_k$ and the initial conditions is called the filtering problem.

Given the assumptions made in terms of the process model, the state vector $\boldsymbol{X}_k$ is a random vector distributed according to

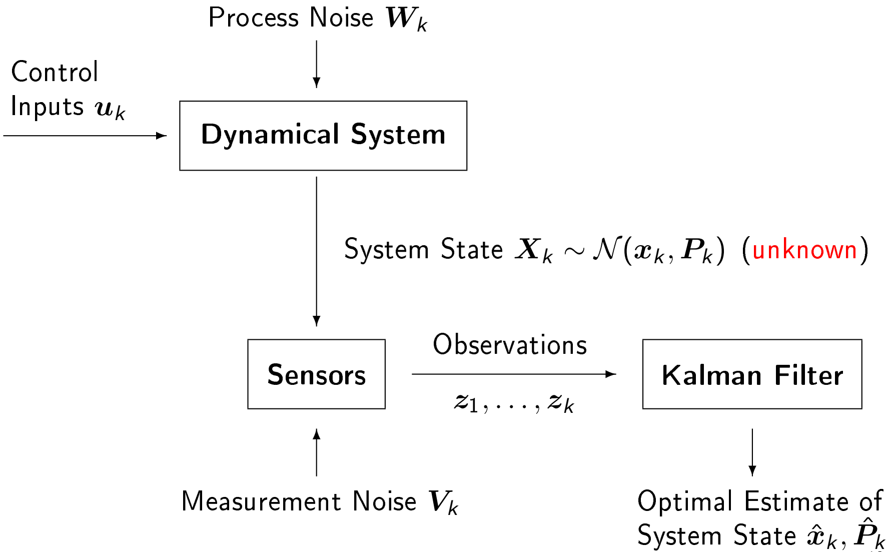
$$\boldsymbol{X}_k \sim \mathcal{N}(\boldsymbol{x}_k, \boldsymbol{P}_k), \quad k \in \mathbb{N}$$

The task of computing **estimates** $\hat{\boldsymbol{x}}_k$ and $\hat{\boldsymbol{P}}_k$ for the mean vector $\boldsymbol{x}_k$ and the covarince matrix $\boldsymbol{P}_k$ of $\boldsymbol{X}_k$ based on the recorded observations $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k$ is solved by the so-called Kalman filter.

6.2 Structure and Function of the Kalman Filter

Structure and Function of the Kalman Filter

State Estimation of Dynamical Systems:



Structure and Function of the Kalman Filter

Idea of the Kalman Filter:

The core idea of the Kalman filter is to formulate the estimate $\hat{\boldsymbol{x}}_k$ of the system state at time k in terms of a **linear combination** of the previous estimate $\hat{\boldsymbol{x}}_{k-1}$ and the new observation $\boldsymbol{z}_k$.

This is possible since the estimate $\hat{\boldsymbol{x}}_{k-1}$ at time $k-1$ contains **all information** about the observations $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{k-1}$, thus allowing for a recursive formulation of the state estimation problem.

The **main steps** of the Kalman filter algorithm are

1. Initialization
2. Prediction
3. Correction

where the latter two steps (Prediction and Correction) are iterated.

Structure and Function of the Kalman Filter

Initialization Step:

The Kalman filter algorithm requires **starting conditions** in terms of a prior estimate of the system state at time $k = 0$.

Frequently, the values

$$\hat{\boldsymbol{x}}_0 = \mathbf{0} \quad \text{and} \quad \hat{\boldsymbol{P}}_0 = \sigma^2 \cdot \boldsymbol{I}$$

are chosen as starting conditions for $k = 0$, where the matrix $\boldsymbol{I}$ is given by the n-dimensional **identity matrix**

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

and $\sigma^2 > 0$ represents a suitable variance.

Structure and Function of the Kalman Filter

Prediction Step:

In the first step of the filtering procedure, the previous estimates $\hat{\boldsymbol{x}}_{k-1}$ and $\hat{\boldsymbol{P}}_{k-1}$ for time $k-1$ are fed into the **process model** in order to obtain a prediction for the system state at time k.

The prediction of the **mean vector** is obtained as

$$\boxed{\hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}_{k-1}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{B}_{k-1}\boldsymbol{u}_{k-1}}$$

and the prediction of the **covariance matrix** yields

$$\boxed{\hat{\boldsymbol{P}}_{k|k-1} = \boldsymbol{F}_{k-1}\hat{\boldsymbol{P}}_{k-1}\boldsymbol{F}_{k-1}^{T} + \boldsymbol{Q}_{k-1}}$$

where $\boldsymbol{Q}_{k-1}$ denotes the covariance matrix of the process noise vector $\boldsymbol{W}_{k-1}$, capturing the uncertainties of the process model.

Structure and Function of the Kalman Filter

Correction Step:

In the subsequent step of the filtering procedure, the predictions $\hat{\boldsymbol{x}}_{k|k-1}$ and $\hat{\boldsymbol{P}}_{k|k-1}$ for the system state at time k are subject to a correction based on the current **observation** $\boldsymbol{z}_k$.

Based on the so-called Kalman gain matrix according to

$$\boldsymbol{K}_k = \hat{\boldsymbol{P}}_{k|k-1}\boldsymbol{H}_k^T \cdot \left(\boldsymbol{H}_k\hat{\boldsymbol{P}}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k\right)^{-1}$$

the corrected **mean vector** is obtained as

$$\boxed{\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \cdot \left(\boldsymbol{z}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}\right)}$$

and the corrected **covariance matrix** is given by

$$\boxed{\hat{\boldsymbol{P}}_k = \hat{\boldsymbol{P}}_{k|k-1} - \boldsymbol{K}_k\left(\boldsymbol{H}_k\hat{\boldsymbol{P}}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k\right)\boldsymbol{K}_k^T}$$

Structure and Function of the Kalman Filter

Frequently, the correction step of the Kalman filter algorithm is formulated by using **auxiliary quantities**, the so-called innovation and the residual covariance matrix:

Correction Step – Innovation:

The innovation at time k is given by

$$\boxed{\tilde{\boldsymbol{y}}_k = \boldsymbol{z}_k - \boldsymbol{H}_k \, \hat{\boldsymbol{x}}_{k|k-1}}$$

The innovation $\tilde{\boldsymbol{y}}_k$ indicates how well the predicted mean $\hat{\boldsymbol{x}}_{k|k-1}$ obtained by the process model describes the current observation $\boldsymbol{z}_k$.

A bad prediction $\hat{\boldsymbol{x}}_{k|k-1}$ corrresponds to a significant innovation $\tilde{\boldsymbol{y}}_k$ and vice versa. In terms of the Kalman gain matrix $\boldsymbol{K}_k$ and the innovation $\tilde{\boldsymbol{y}}_k$, the corrected mean vector takes the form

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \tilde{\boldsymbol{y}}_k$$

Structure and Function of the Kalman Filter

Correction Step – Residual Covariance Matrix:

The residual covariance matrix at time k is given by

$$\boxed{\boldsymbol{S}_k = \boldsymbol{H}_k \hat{\boldsymbol{P}}_{k|k-1} \boldsymbol{H}_k^T + \boldsymbol{R}_k}$$

The residual covariance matrix $\boldsymbol{S}_k$ captures **both** the resulting covariance from the predicted covariance matrix $\hat{\boldsymbol{P}}_{k|k-1}$ subject to the measurement model as well as the covariance $\boldsymbol{R}_k$ of the measurement noise vector.

With the help of the residual covariance matrix $\boldsymbol{S}_k$, the Kalman gain matrix $\boldsymbol{K}_k$ can be written as

$$\boldsymbol{K}_k = \hat{\boldsymbol{P}}_{k|k-1} \boldsymbol{H}_k^T \boldsymbol{S}_k^{-1}$$

Structure and Function of the Kalman Filter

Remarks:

- As can be seen in the correction step, the estimate $\hat{\boldsymbol{x}}_k$ of the mean vector at time k depends **linearly** on the observation $\boldsymbol{z}_k$, i.e. the Kalman algorithm implements a linear filter.
- As the number of observations $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3, \ldots$ increase, the estimates $\hat{\boldsymbol{x}}_k$ and $\hat{\boldsymbol{P}}_k$ approach the true values $\boldsymbol{x}_k$ and $\boldsymbol{P}_k$ **arbitrarily close**.
- In statistical terms, the Kalman filter implements an unbiased and consistent estimator which has **minimum variance**.
- Therefore, the Kalman filter is an **optimal linear filter**, even more general nonlinear filter do not provide better results for the considered state space model.

6.3 Kalman Filter – Exercise

Kalman Filter – Exercise

Consider the linear time-invariant system with the process model

$$\boldsymbol{X}_k = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \boldsymbol{X}_{k-1} + \boldsymbol{W}, \quad k \in \mathbb{N}$$

where the process noise vector $\boldsymbol{W}$ is distributed according to

$$\boldsymbol{W} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}), \quad \boldsymbol{Q} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

The measurement model should be given by

$$Z_k = \begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \boldsymbol{X}_k + V, \quad k \in \mathbb{N}$$

where the measurement noise V is distributed according to

$$V \sim \mathcal{N}(0, 1)$$

Kalman Filter – Exercise

The starting conditions of the Kalman filter algorithm should be given by

$$\hat{\boldsymbol{x}}_0 = \mathbf{0} \quad \text{and} \quad \hat{\boldsymbol{P}}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Determine the state estimates $\hat{\boldsymbol{x}}_k$ if the available observation sequence is

$$(z_1, z_2, z_3, z_4, \ldots) = (4, -1, 2, 3, \ldots)$$