The Kalman Filter

Control Engineering (Reguleringsteknik)

Christoffer Sloth

`chsl@mmmi.sdu.dk`

SDU Robotics
The Maersk Mc-Kinney Moller Institute
University of Southern Denmark

SDU

Agenda

Introduction

Curriculum for Reguleringsteknik (REG)

Matematiske og grafiske metoder til syntese af lineære tidsinvariante systemer:[1]

- diskret og kontinuert tilstandsbeskrivelse
- analyse i tid og frekvens
- stabilitet, reguleringshastighed, følsomhed og fejl
- digitale PI, PID, LEAD og LAG regulatorer (serieregulatorer)
- tilstandsregulering, pole-placement og tilstands-estimering (observer)
- optimal regulering (least squares) og optimal tilstands-estimation (Kalman-filter)

Færdigheder:
Efter gennemførelse af kurset kan den succesfulde studerende:

- kunne analysere, dimensionere og implementere såvel kontinuert som tidsdiskret regulering af lineære tidsinvariante og stokastiske systemer

Kompetencer:
Efter gennemførelse af kurset kan den succesfulde studerende:

- anvende og implementere klassiske og moderne reguleringsteknikker for at kunne styre og regulere en robot hurtig og præcist

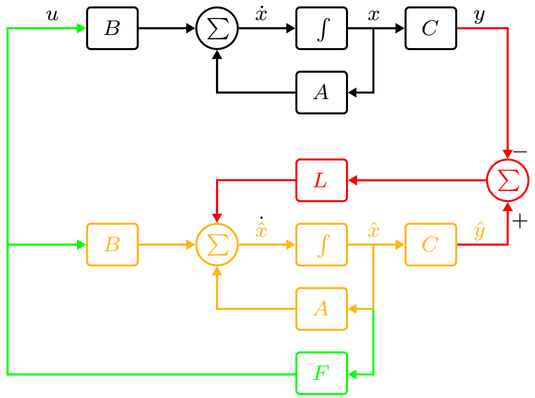
[1] Based on https://fagbesk.sam.sdu.dk/?fag_id=39673

The twelve lectures of the course are

- **Lecture 1**: Introduction to Linear Time-Invariant Systems
- **Lecture 2**: Stability and Performance Analysis
- **Lecture 3**: Introduction to Control
- **Lecture 4**: Design of PID Controllers
- **Lecture 5**: Root Locus
- **Lecture 6**: The Nyquist Plot
- **Lecture 7**: Dynamic Compensators and Stability Margins
- **Lecture 8**: Implementation
- **Lecture 9**: State Feedback
- **Lecture 10**: Observer Design
- **Lecture 11**: Optimal Control (Linear Quadratic Control)
- **Lecture 12**: The Kalman Filter

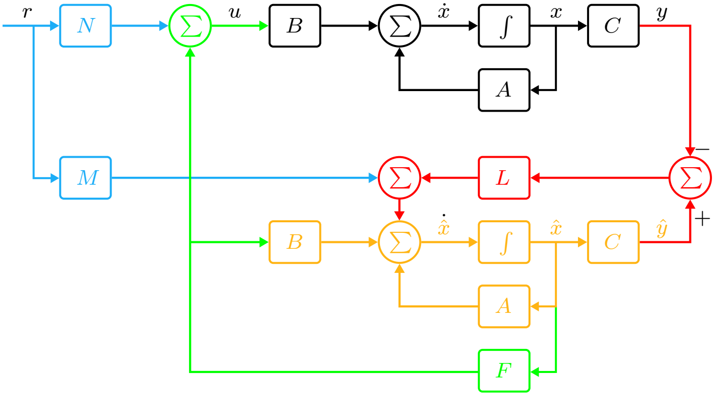
Introducing Reference Signals

Block Diagram



Introducing Reference Signals

Block Diagram



Introducing Reference Signals

System Description

System:

$$\begin{aligned} \dot{x} &= Ax + B(F\hat{x} + Nr) \\ y &= Cx \end{aligned}$$

Observer:

$$\dot{\hat{x}} = A\hat{x} + BF\hat{x} + L(C\hat{x} - y) + Mr$$

Introducing Reference Signals

System Description

System:

$$\begin{aligned} \dot{x} &= Ax + B(F\hat{x} + Nr) \\ y &= Cx \end{aligned}$$

Observer:

$$\dot{\hat{x}} = A\hat{x} + BF\hat{x} + L(C\hat{x} - y) + Mr$$

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & BF \\ -LC & A + BF + LC \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} BN \\ M \end{bmatrix} r$$

$$y = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

Recall from Lecture 2 how to find the zeros of a state space model.

LEMMA. A square (#inputs=#outputs) system with a state space model of the form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

has a zero with value $z \in \mathbb{C}$ only if

$$\det \begin{bmatrix} A - zI & B \\ C & D \end{bmatrix} = 0$$

$$\det\left(\begin{bmatrix} A_{\mathsf{cl}} - zI & B_{\mathsf{cl}} \\ C_{\mathsf{cl}} & D_{\mathsf{cl}} \end{bmatrix}\right) = 0$$

Introducing Reference Signals

Zeros of Closed-Loop System

$$\det\left(\begin{bmatrix} A_{\text{cl}} - zI & B_{\text{cl}} \\ C_{\text{cl}} & D_{\text{cl}} \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & BF & BN \\ -LC & A + BF + LC - zI & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

Introducing Reference Signals
Zeros of Closed-Loop System

$$\det\left(\begin{bmatrix} A_{\text{cl}} - zI & B_{\text{cl}} \\ C_{\text{cl}} & D_{\text{cl}} \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & BF & BN \\ -LC & A + BF + LC - zI & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & BF - BNN^{-1}F & BN \\ -LC & A + BF + LC - zI - MN^{-1}F & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

Introducing Reference Signals

Zeros of Closed-Loop System

$$\det\left(\begin{bmatrix} A_{\text{cl}} - zI & B_{\text{cl}} \\ C_{\text{cl}} & D_{\text{cl}} \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & BF & BN \\ -LC & A + BF + LC - zI & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & BF - BNN^{-1}F & BN \\ -LC & A + BF + LC - zI - MN^{-1}F & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & 0 & BN \\ -LC & A + BF + LC - MN^{-1}F - zI & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

Introducing Reference Signals

Zeros of Closed-Loop System

$$\det\left(\begin{bmatrix} A - zI & BF & BN \\ -LC & A + BF + LC - zI & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & BF - BNN^{-1}F & BN \\ -LC & A + BF + LC - zI - MN^{-1}F & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & 0 & BN \\ -LC & A + BF + LC - MN^{-1}F - zI & M \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

$$\det\left(\begin{bmatrix} A - zI & 0 & B \\ -LC & A + BF + LC - \tilde{M}F - zI & \tilde{M} \\ C & 0 & 0 \end{bmatrix}\right) = 0$$

Introducing Reference Signals

Zeros of Closed-Loop System

$$\det\left(\begin{bmatrix} A-zI & BF-BNN^{-1}F & BN \\ -LC & A+BF+LC-zI-MN^{-1}F & M \\ C & 0 & 0 \end{bmatrix}\right)=0$$

$$\det\left(\begin{bmatrix} A-zI & 0 & BN \\ -LC & A+BF+LC-MN^{-1}F-zI & M \\ C & 0 & 0 \end{bmatrix}\right)=0$$

$$\det\left(\begin{bmatrix} A-zI & 0 & B \\ -LC & A+BF+LC-\tilde{M}F-zI & \tilde{M} \\ C & 0 & 0 \end{bmatrix}\right)=0$$

$$\begin{cases} \det\left(\begin{bmatrix} A-zI & B \\ C & 0 \end{bmatrix}\right)=0 \qquad \text{or} \\ \det\left(A+BF+LC-\tilde{M}F-zI\right)=0 \end{cases}$$

Zero Assignment

LEMMA. If $\tilde{M}$ is an 'observer gain' such that the characteristic polynomial of the matrix $A_{\text{za}} + \tilde{M}C_{\text{za}}$ has the characteristic polynomial

$$\det\left(sI - \left(A_{\text{za}} + \tilde{M}C_{\text{za}}\right)\right) = (s - z_1)\cdots(s - z_n)$$

with $A_{\text{za}} = A + BF + LC$ and $C_{\text{za}} = -F$, then the numbers $z_1, \ldots, z_n$ are all zeros of the closed loop transfer function from r to y.

1. Design $\tilde{M}$ assigning zeros close to the cut-off frequency of the Bode plot, such that the 'horizontal' part is extended.

1. Design $\tilde{M}$ assigning zeros close to the cut-off frequency of the Bode plot, such that the 'horizontal' part is extended.
2. Compute N such that the DC-value of the transfer function from r to y is unity:

$$N = -\left(C_{\text{cl}} A_{\text{cl}}^{-1} \tilde{B}_{\text{cl}}\right)^{-1}$$

where

$$A_{\text{cl}} = \begin{bmatrix} A & BF \\ -LC & A + BF + LC \end{bmatrix}, \quad \tilde{B}_{\text{cl}} = \begin{bmatrix} B \\ \tilde{M} \end{bmatrix}$$

$$C_{\text{cl}} = \begin{bmatrix} C & 0 \end{bmatrix}$$

1. Design $\tilde{M}$ assigning zeros close to the cut-off frequency of the Bode plot, such that the 'horizontal' part is extended.
2. Compute N such that the DC-value of the transfer function from r to y is unity:

   $$N = -\left(C_{\text{cl}} A_{\text{cl}}^{-1} \tilde{B}_{\text{cl}}\right)^{-1}$$

   where

   $$A_{\text{cl}} = \begin{bmatrix} A & BF \\ -LC & A + BF + LC \end{bmatrix}, \quad \tilde{B}_{\text{cl}} = \begin{bmatrix} B \\ \tilde{M} \end{bmatrix}$$
   $$C_{\text{cl}} = \begin{bmatrix} C & 0 \end{bmatrix}$$

3. Compute $M = MN^{-1}N = \tilde{M}N$.

We consider again the system

$$
\begin{aligned}
\dot{x} &= \begin{bmatrix} 2 & -3 \\ 4 & -5 \end{bmatrix} x + \begin{bmatrix} 2 \\ 3 \end{bmatrix} u \\
y &= \begin{bmatrix} -3 & 2 \end{bmatrix} x
\end{aligned}
$$

A state feedback F that assign poles in $\{-3, -4\}$ and an observer gain L that assigns poles in $\{-9, -12\}$ are given by:

$$
F = \begin{bmatrix} 22 & -16 \end{bmatrix}, \quad L = \begin{bmatrix} -122 \\ -192 \end{bmatrix}
$$

We would like to assign zeros from r to y in $\{-3, -4\}$ to cancel the poles from F.

Zero Assignment

Example (2)

With these values of F and L we obtain:

$$A_{\text{za}} = A + BF + LC = \begin{bmatrix} 412 & -279 \\ 646 & -437 \end{bmatrix}$$

$$C_{\text{za}} = -F = \begin{bmatrix} -22 & 16 \end{bmatrix}$$

An 'observer gain' that assigns poles in $\{-3, -4\}$ for $A_{\text{za}} + \tilde{M} C_{\text{za}}$ is

$$\tilde{M} = \begin{bmatrix} 7.0460 \\ 10.8133 \end{bmatrix}$$

Zero Assignment

Example (3)

N can be computed as:

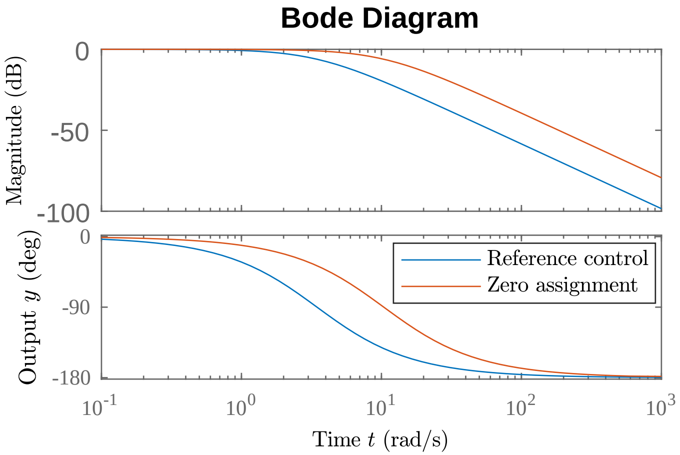
$$N = -\left(\begin{pmatrix} C & 0 \end{pmatrix} \begin{pmatrix} A & BF \\ -LC & A + BF + LC \end{pmatrix}^{-1} \begin{pmatrix} B \\ \tilde{M} \end{pmatrix}\right)^{-1}$$
$$= 108$$

M is obtained from:

$$M = \tilde{M} N = \begin{bmatrix} 7.0460 \\ 10.8133 \end{bmatrix} \cdot 108 = \begin{bmatrix} 760.97 \\ 1167.84 \end{bmatrix}$$

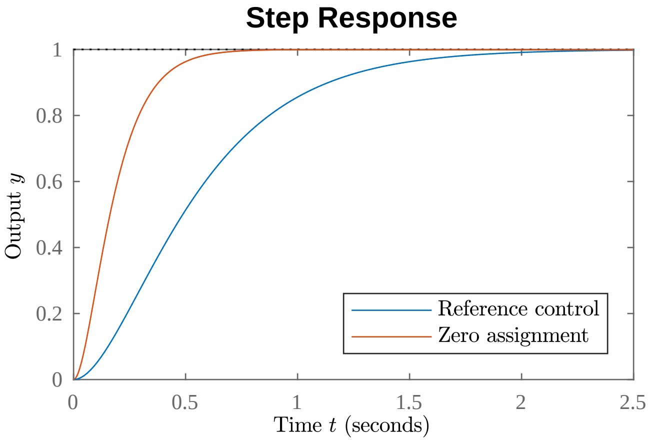
Zero Assignment

Example: Bode Plot

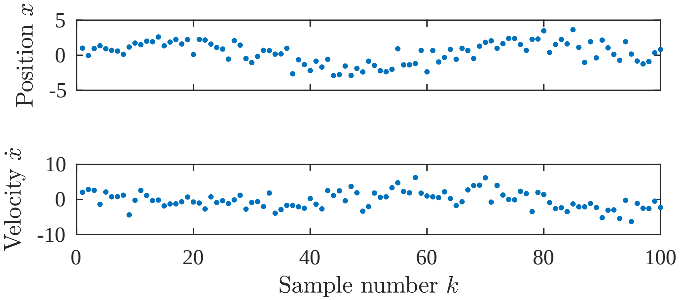


Zero Assignment

Example: Step Response



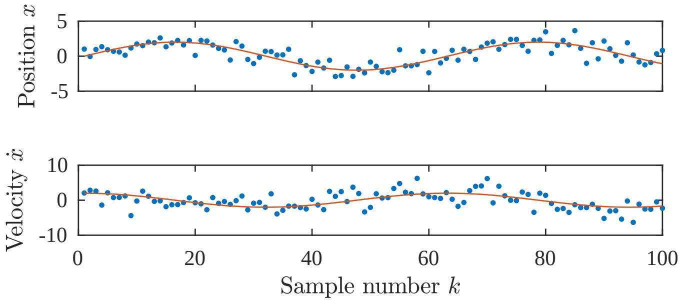
How is it possible to combine noisy measurements, and and an uncertain system model to provide a "good" state estimate?



Random Variables

Motivating Example

How is it possible to combine noisy measurements, and and an uncertain system model to provide a "good" state estimate?



The solution is a ***Kalman filter*** that relies on a stochastic model, and noisy measurements.

To introduce uncertainty and noise in the considered system models, we introduce ***random variables***.

To introduce uncertainty and noise in the considered system models, we introduce ***random variables***.

Let X be a random variable describing the outcome of rolling a fair dice. The fair dice is characterized by

- It has 6 different outcomes $\{1, 2, 3, 4, 5, 6\}$.
- The probability of getting each of the six outcomes is the same, i.e., $\Pr(X = 4) = \frac{1}{6}$.
- The outcome of each roll of the dice is independent.

To introduce uncertainty and noise in the considered system models, we introduce ***random variables***.

Let X be a random variable describing the outcome of rolling a fair dice. The fair dice is characterized by

- It has 6 different outcomes $\{1, 2, 3, 4, 5, 6\}$.
- The probability of getting each of the six outcomes is the same, i.e., $\Pr(X = 4) = \frac{1}{6}$.
- The outcome of each roll of the dice is independent.

To describe the above dice mathematically, a ***probability mass function*** p_X is associated to X that determines the probability that X equals x, i.e.,

$$p_X(x) = \Pr(\{X = x\})$$

To introduce uncertainty and noise in the considered system models, we introduce ***random variables***.

Let X be a random variable describing the outcome of rolling a fair dice. The fair dice is characterized by

- It has 6 different outcomes $\{1, 2, 3, 4, 5, 6\}$.
- The probability of getting each of the six outcomes is the same, i.e., $\Pr(X = 4) = \frac{1}{6}$.
- The outcome of each roll of the dice is independent.

To describe the above dice mathematically, a ***probability mass function*** p_X is associated to X that determines the probability that X equals x, i.e.,

$$p_X(x) = \Pr(\{X = x\})$$

and for the fair dice

$$\Pr(\{X = 1\}) = \Pr(\{X = 2\}) = \cdots = \Pr(\{X = 6\}) = \frac{1}{6}$$

The ***expected value*** (mean value) of a random variable X with n outcomes $\{x_1, x_2, \ldots, x_n\}$ can be determined from the probability mass function p_X as

$$E[X] \equiv \sum_{i=1}^{n} x_i p_X(x_i).$$

The ***expected value*** (mean value) of a random variable X with n outcomes $\{x_1, x_2, \ldots, x_n\}$ can be determined from the probability mass function p_X as

$$E[X] \equiv \sum_{i=1}^{n} x_i p_X(x_i).$$

The ***variance*** quantifies how much a random variable is varying around the mean value and is defined as

$$\text{Var}(X) = E[(X-\mu)^2]$$

where $\mu = E(X)$.

A continuous random variable X often has zero probability of being one particular value; thus, its outcome is described with a ***probability density function*** f_X as

$$\Pr(\{a \leq X \leq b\}) = \int_a^b f_X(x)dx.$$

A continuous random variable X often has zero probability of being one particular value; thus, its outcome is described with a ***probability density function*** f_X as

$$\text{Pr}(\{a \leq X \leq b\}) = \int_a^b f_X(x)dx.$$

This means that the probability of the random value being in a particular range $[a\, b]$ can be determined as shown above.

The ***expected value*** (mean value) of a continuous random variable X can be determined from the probability density function f_X as

$$E[X] \equiv \int_{-\infty}^{\infty} x f_X(x) dx.$$

The ***expected value*** (mean value) of a continuous random variable X can be determined from the probability density function f_X as

$$E[X] \equiv \int_{-\infty}^{\infty} x f_X(x) dx.$$

The ***variance*** of X is defined as

$$\text{Var}(X) = E[(X - \mu)^2]$$

where $\mu = E(X)$.

The random variable X is said to be normally distributed if it has probability density function

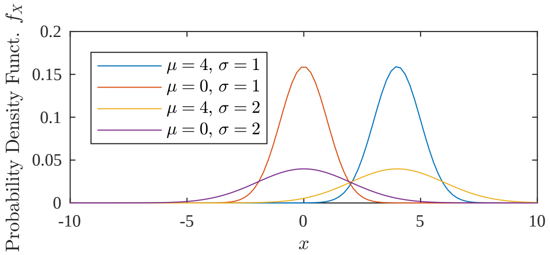
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

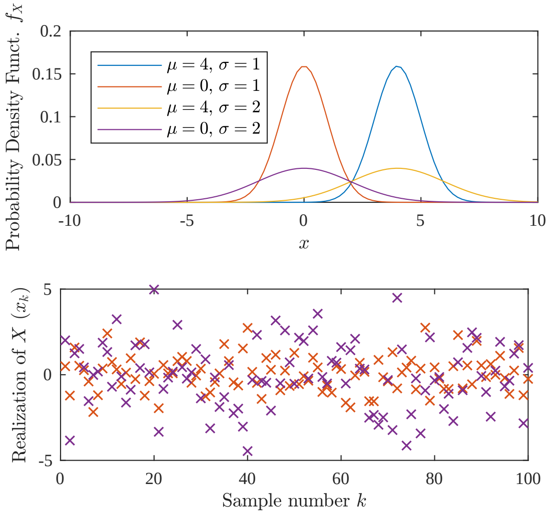
where μ is the expectation of X and σ is the standard deviation of X (the standard deviation is defined from $\sigma^2 = \text{Var}(X)$).

The random variable X is said to be normally distributed if it has probability density function

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ is the expectation of X and σ is the standard deviation of X (the standard deviation is defined from $\sigma^2 = \text{Var}(X)$).




Probability Density Funct. f_X
0.2
0.15
0.1
0.05
0
$\mu = 4,\ \sigma = 1$
$\mu = 0,\ \sigma = 1$
$\mu = 4,\ \sigma = 2$
$\mu = 0,\ \sigma = 2$
-10
-5
0
5
10
x
Realization of X (x_k)
5
0
-5
0
20
40
60
80
100
Sample number k

For a multivariate random variable

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

the ***covariance matrix*** is

$$\Sigma = E\left[(X - E[X])(X - E[X])^T\right]$$

For a multivariate random variable

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

the ***covariance matrix*** is

$$\Sigma = E\left[(X - E[X])(X - E[X])^T\right]$$

We consider random variables that are ***independent***, i.e., for a dice the probability of getting a 6 is the same independent on the previous outcome.

Recall from Lecture 1 that a deterministic discrete-time state space model is given by

$$
\begin{aligned}
x_{k+1} &= \Phi x_k + \Gamma u_k \\
y_k &= C x_k + D u_k
\end{aligned}
$$

Recall from Lecture 1 that a deterministic discrete-time state space model is given by

$$
\begin{aligned}
x_{k+1} &= \Phi x_k + \Gamma u_k \\
y_k &= C x_k + D u_k
\end{aligned}
$$

Now the following ***stochastic discrete-time state space model*** is considered

$$
\begin{aligned}
x_{k+1} &= \Phi x_k + \Gamma u_k + w_k \\
y_k &= C x_k + D u_k + v_k
\end{aligned}
$$

where w_k is the ***process noise*** (drawn from a zero mean normal distribution with covariance matrix Q_k) and v_k is the ***measurement noise*** (drawn from a zero mean normal distribution with covariance matrix R_k).

The Kalman filter finds an ***unbiased state estimate*** $\hat{x}_k$ of x_k ($E[x_k - \hat{x}_k] = 0$) ***with minimal variance***, by exploiting

- a model of the system
- a noise model

The Kalman filter finds an ***unbiased state estimate*** $\hat{x}_k$ of x_k $(E[x_k - \hat{x}_k] = 0)$ ***with minimal variance***, by exploiting

- a model of the system
- a noise model

The Kalman filter is similar to the observer that was introduced in Lecture 10; however, the observer did not take into account the process noise w_k and the measurement noise v_k.

The Kalman filter consists of two stages

1. Prediction: $\begin{cases} \hat{x}_{k+1|k} & = \Phi \hat{x}_{k|k} + \Gamma u_k \\ P_{k+1|k} & = \Phi P_{k|k} \Phi^T + Q_k \end{cases}$

2. Update: $\begin{cases} \hat{x}_{k+1|k+1} & = \hat{x}_{k+1|k} + K_k (y_k - C \hat{x}_{k+1|k}) \\ P_{k+1|k+1} & = P_{k+1|k} - K_k C P_{k+1|k} \end{cases}$

where the ***Kalman gain*** is given by

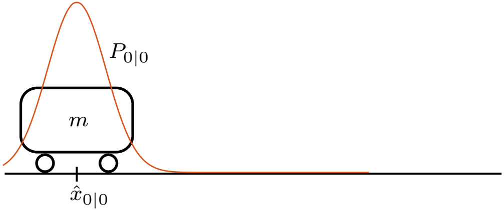
$$K_k = P_{k+1|k} C^T (C P_{k+1|k} C^T + R_k)^{-1}$$

$k = 0$

Initialize: Provide $\hat{x}_{0|0}$ and $P_{0|0}$.

$k = 0$

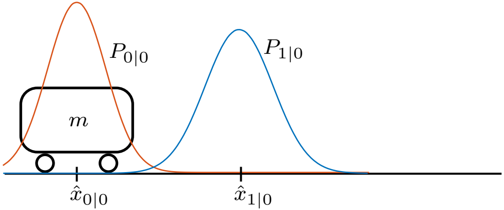
Initialize: Provide $\hat{x}_{0|0}$ and $P_{0|0}$.



$k = 0$

Initialize: Provide $\hat{x}_{0|0}$ and $P_{0|0}$.

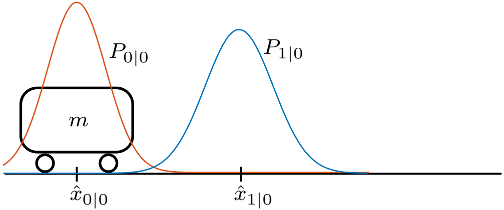
Predict : $\begin{cases} \hat{x}_{1|0} = \Phi \hat{x}_{0|0} + \Gamma u_0 \\ P_{1|0} = \Phi P_{0|0} \Phi^T + Q \end{cases}$



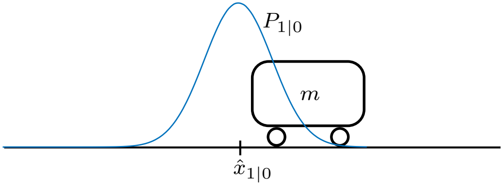
$k = 0$

Initialize: Provide $\hat{x}_{0|0}$ and $P_{0|0}$.

Predict : $\begin{cases} \hat{x}_{1|0} = \Phi \hat{x}_{0|0} + \Gamma u_0 \\ P_{1|0} = \Phi P_{0|0} \Phi^T + Q \end{cases}$



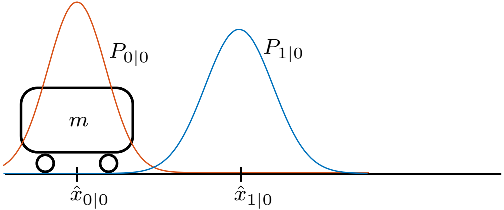
$k = 1$



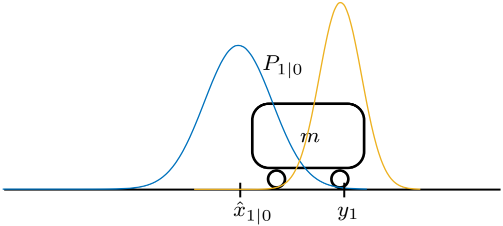
$k = 0$

Initialize: Provide $\hat{x}_{0|0}$ and $P_{0|0}$.

Predict : $\begin{cases} \hat{x}_{1|0} = \Phi \hat{x}_{0|0} + \Gamma u_0 \\ P_{1|0} = \Phi P_{0|0} \Phi^T + Q \end{cases}$



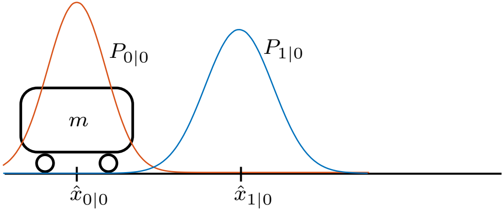
$k = 1$ (Get measurement y_1)



$k = 0$

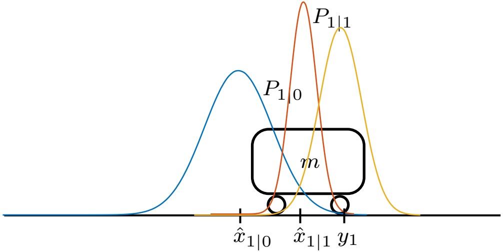
Initialize: Provide $\hat{x}_{0|0}$ and $P_{0|0}$.

Predict : $\begin{cases} \hat{x}_{1|0} = \Phi \hat{x}_{0|0} + \Gamma u_0 \\ P_{1|0} = \Phi P_{0|0} \Phi^T + Q \end{cases}$



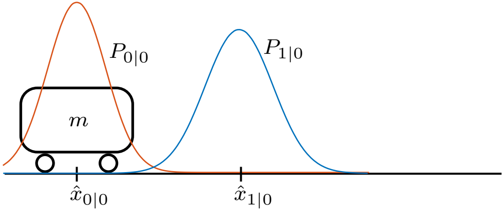
$k = 1$ (Get measurement y_1)

Update : $\begin{cases} \hat{x}_{1|1} = \hat{x}_{1|0} + K_k(y_1 - C\hat{x}_{1|0}) \\ P_{1|1} = P_{1|0} - K_k C P_{1|0} \end{cases}$



$k = 0$
Initialize: Provide $\hat{x}_{0|0}$ and $P_{0|0}$.

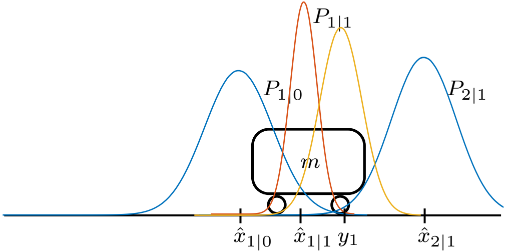
Predict : $\begin{cases} \hat{x}_{1|0} = \Phi \hat{x}_{0|0} + \Gamma u_0 \\ P_{1|0} = \Phi P_{0|0} \Phi^T + Q \end{cases}$



$k = 1$ (Get measurement y_1)

Update : $\begin{cases} \hat{x}_{1|1} & = \hat{x}_{1|0} + K_k(y_1 - C\hat{x}_{1|0}) \\ P_{1|1} & = P_{1|0} - K_k C P_{1|0} \end{cases}$

Predict : $\begin{cases} \hat{x}_{2|1} & = \Phi \hat{x}_{1|1} + \Gamma u_1 \\ P_{2|1} & = \Phi P_{1|1} \Phi^T + Q \end{cases}$



Evaluation of the Course

What has been good and bad about the course?