# Filtering and Identification

## Day 2 - Lecture 2:
## Kalman Filtering

Michel Verhaegen

# Overview

- **The Stochastic Least Squares Problem**
- Asymptotic observer
- State reconstruction in the presence of noise
- The Conventional Kalman filter (SLS framework)
- Case Study: Level estimation in liquid tanks
- The Kalman filter problem as a weighted least squares problem.

# The SLS Problem

Given the prior on the RV $x \sim (\overline{x}, P)$ with $P \geq 0$ and given the observations:

$$y = Fx + L\epsilon \quad \epsilon \sim (0, I) \quad E[(x - \overline{x})\epsilon^T] = 0$$

with $F, L$ deterministic $L$ full rank. Then seek among the linear estimators:

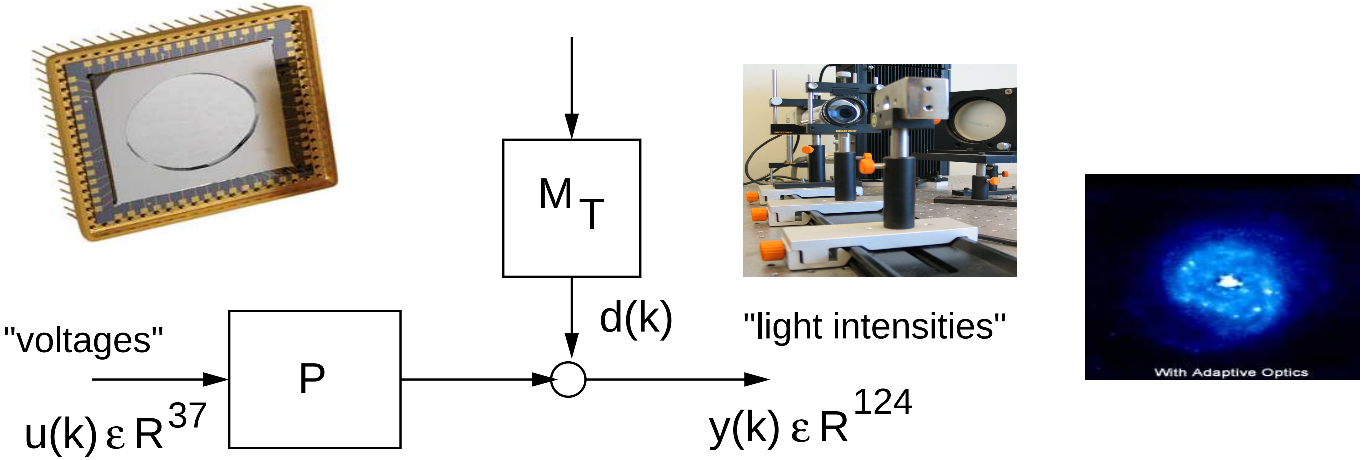
$$\widetilde{x} = \begin{bmatrix} M & N \end{bmatrix} \begin{bmatrix} y \\ \overline{x} \end{bmatrix}$$

The unbiased minimum variance estimate, i.e.:

$$E[(x - \widetilde{x})(x - \widetilde{x})^T] \quad \text{is minimized} \quad E[\widetilde{x}] = \overline{x}$$

# Smart Optics Systems

# The Smart Optics Problem




Later on in the course we will learn how to derive from a sequence of (light intensity) measurements $\{y(k)\}_{k=1}^{N(\text{very large} \approx 10000)}$ for $u(k) \equiv 0$ the following model:

$$
\begin{aligned}
x(k+1) &= Ax(k) + Ke(k) \quad e(k) \sim (0, R_e) \text{ \& white} \\
y(k) &= Cx(k) + e(k)
\end{aligned}
$$

In real-time: Can we process .., $y(k-1), y(k)$ to estimate $y(k+1)$?

# Notation

The estimate of $y(k+1)$ $\Big(x(k+1)\Big)$ given past measurements $.., y(k-1), y(k)$ is called the one-step predicted estimate $\tilde{y}(k|k-1)$ $\Big(\tilde{x}(k|k-1)\Big)$.

The estimate of $x(k)$ given past measurements $.., y(k-1), y(k)$ to estimate $x(k)$ is called the filtered estimate $\tilde{x}(k|k)$.

# Overview

- The Stochastic Least Squares Problem
- **Asymptotic observer**
- State reconstruction in the presence of noise
- The Conventional Kalman filter (SLS framework)
- Case Study: Level estimation in liquid tanks
- The Kalman filter problem as a weighted least squares problem.

# The asymptotic observer

Given the state space model

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

with $x(k) \in \mathbb{R}^n, y(k) \in \mathbb{R}^\ell, u(k) \in \mathbb{R}^m$
For the case $x(0)$ unknown, we seek a 'state-observer' of the form,

$$\underbrace{\widehat{x}(k+1) = A\widehat{x}(k) + Bu(k)}_{\text{Predictor}} + \underbrace{K\Big(y(k) - C\widehat{x}(k) - Du(k)\Big)}_{\text{Corrector}}$$

such that $\lim_{k \to \infty} \widehat{x}(k) - x(k) = 0$

# The asymptotic observer

Given the dynamics of the system and observer state,

$$
\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) \\
\widehat{x}(k+1) &= A\widehat{x}(k) + Bu(k) + K\Big(y(k) - C\widehat{x}(k) - Du(k)\Big) \\
&= A\widehat{x}(k) + Bu(k) + K\Big(Cx(k) + Du(k) - C\widehat{x}(k) - Du(k)\Big) \\
&= A\widehat{x}(k) + Bu(k) - KC\Big(\widehat{x}(k) - x(k)\Big)
\end{aligned}
$$

The dynamics of the state error, $x_e(k) = \widehat{x}(k) - x(k)$, satisfies,

$$x_e(k+1) = (A - KC)x_e(k)$$

the goal is to design $K$, such that:

$(A - KC)$ is **asymptotically stable**

# Design of the asymptotic observer

Recall the observer structure,

$$\widehat{x}(k+1) = (A-KC)\widehat{x}(k)+Bu(k)+K\Big(y(k)-Du(k)\Big)$$

Theorem: If the pair $(A, C)$ is observable, then there exists a matrix $K \in \mathbb{R}^{n\times\ell}$ such that $A - KC$ is asymptotically stable. Design in matlab: Given the pair $(A, C)$ and a desired spectrum $Sp$ (spectrum) of the poles of the matrix $A - KC$ then we can determine (in matlab) the gain matrix $K$ by,

```
K = place(A',C',Sp)';
```

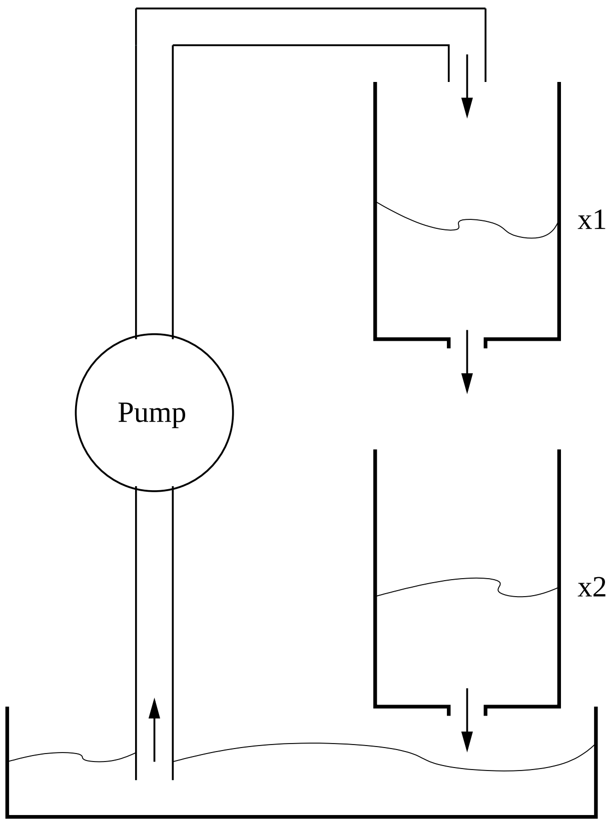
# Acker_test.m

# Example: Double tank

Linearized discrete-time state-space model:

$$x(k+1) = \begin{bmatrix} 0.9512 & 0 \\ 0.0476 & 0.9512 \end{bmatrix} x(k) + \begin{bmatrix} 0.0975 \\ 0.0024 \end{bmatrix} u(k)$$

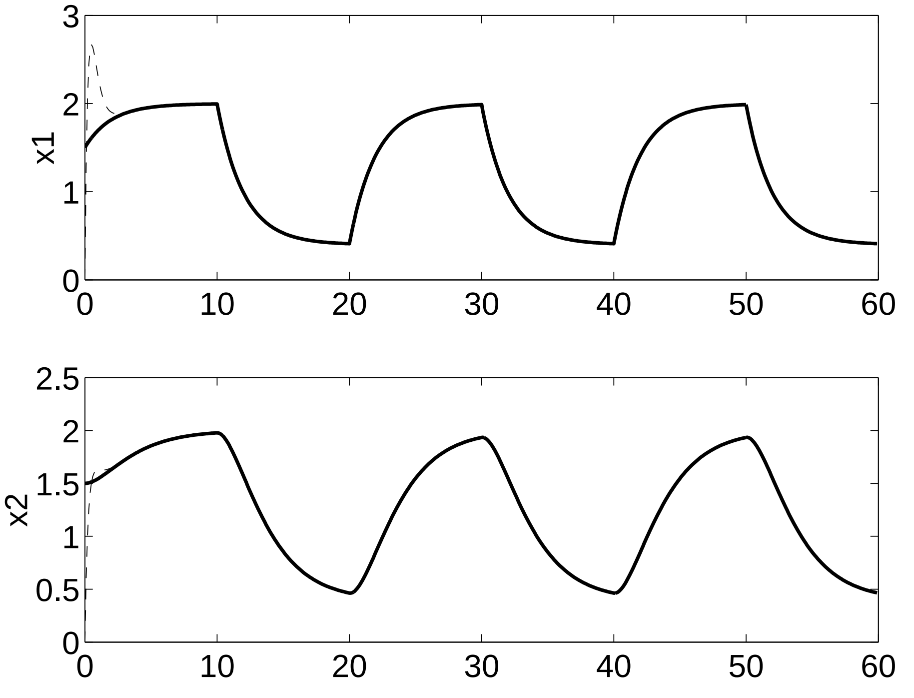
$$y(k) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(k)$$




Determine levels in both tanks: estimate states.

# Example: asymptotic observer

Selecting the eigenvalues of the asymptotic observer equal to $\{0.7, 0.8\}$ with initial state vector $[1.5, 1.5]^T$ yields,



# Example: Double tank with noise

Signal generating model:

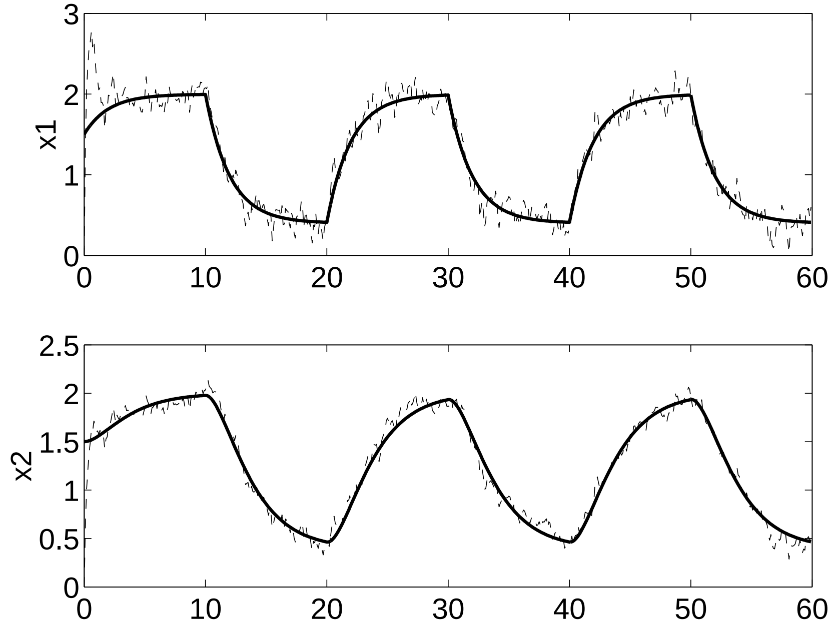
$$
\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) + \begin{bmatrix} 0.0975 & 0 \\ 0.0024 & 0.0975 \end{bmatrix} w(k) \\
y(k) &= \begin{bmatrix} 0 & 1 \end{bmatrix} x(k) + v(k)
\end{aligned}
$$

with $w(k)$ and $v(k)$ zero-mean white noise with,

$$
E\left[\begin{array}{c} v(k) \\ \hline w(k) \end{array}\right]\left[\begin{array}{c|c} v(j)^T & w(j)^T \end{array}\right] = \left[\begin{array}{c|cc} 0.0125 & 0 & 0.005 \\ \hline 0 & 0.01 & 0 \\ 0.005 & 0 & 0.01 \end{array}\right] \Delta(k-j)
$$

# Example: asymptotic observer

Selecting the eigenvalues of the asymptotic observer equal to $\{0.7, 0.8\}$ with initial state vector $[1.5, 1.5]^T$ yields,



# Overview

- The Stochastic Least Squares Problem
- Asymptotic observer
- **State reconstruction in the presence of noise**
- The Conventional Kalman filter (SLS framework)
- Case Study: Level estimation in liquid tanks
- The Kalman filter problem as a weighted least squares problem.

## The signal generating model including noise perturbations

State reconstruction is considered for the following signal generating model (SGM):

$$
\begin{aligned}
x(k+1) &= A(k)x(k) + B(k)u(k) + w(k) \\
y(k) &= C(k)x(k) + v(k) \quad \text{with} \quad x(0) \sim \Big(\widehat{x}(0|-1), P(0|-1)\Big) \quad \text{and}
\end{aligned}
$$

$$
\begin{bmatrix} v(k) \\ w(k) \end{bmatrix} \sim \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R(k) & S(k)^T \\ S(k) & Q(k) \end{bmatrix} \geq 0 \right) \; \& \; \text{white} \qquad R(k) > 0
$$

The Asymptotic observer

$\hat{x}(k+1) = A(k)\hat{x}(k) + B(k)u(k) + K(k)(y(k) - C\hat{x}(k))$ yields the state reconstruction error $x_e(k) = x(k) - \hat{x}(k)$,

$$
x_e(k+1) = (A(k) - K(k)C(k))x_e(k) + w(k) - K(k)v(k)
$$

The error $x_e(k)$ is an RV $\Rightarrow$ does not converge to zero.

## The Kalman Filter Problem

Let the SGM be given and let at time $k-1$ an estimate of the state $x(k)$ be given as $\sim (\hat{x}(k|k-1), P(k|k-1) \geq 0)$ with $E[\Big(x(k) - \hat{x}(k|k-1)\Big) \left[v^T(k) \;\; w^T(k)\right]^T] = 0$. Finally, let at time $k$ become available the measurements $u(k), y(k)$, then we seek a linear estimate of $x(k)$ and $x(k+1)$ of the form,

$$\begin{bmatrix} \hat{x}(k|k) \\ \hat{x}(k+1|k) \end{bmatrix} = M \begin{bmatrix} y(k) \\ -B(k)u(k) \end{bmatrix} + N\hat{x}(k|k-1)$$

such that both estimates are minimum variance unbiased estimates, i.e. estimates with the following properties,

$$E[\hat{x}(k|k)] = E[x(k)] \quad E[\hat{x}(k+1|k)] = E[x(k+1)]$$

$E\Big[\Big(x(k)-\widehat{x}(k|k)\Big)\Big(x(k)-\widehat{x}(k|k)\Big)^T\Big]$ and $E\Big[\Big(x(k+1)-\widehat{x}(k+1|k)\Big)\Big(x(k+1)-\widehat{x}(k+1|k)\Big)^T\Big]$
are minimal.

## Why The Kalman Filter Problem is not "precisely" an SLS?

The collected data $u(k), y(k)$ define the following DATA EQUATION:

$$\underbrace{\begin{bmatrix} y(k) \\ -B(k)u(k) \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} C(k) & 0 \\ A(k) & -I \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} x(k) \\ x(k+1) \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} v(k) \\ w(k) \end{bmatrix}}_{L(k)\epsilon \quad \epsilon \sim (0,I)}$$

This would be an SLS provided,

1. initial estimates of both $x(k)$ and $x(k+1)$ would be given.
2. We seek to minimize the joint covariance matrix,

$$E\Big[\Big(\begin{bmatrix} x(k) - \widehat{x}(k|k) \\ x(k+1) - \widehat{x}(k+1|k) \end{bmatrix}\Big)\Big(\begin{bmatrix} x(k) - \widehat{x}(k|k) \\ x(k+1) - \widehat{x}(k+1|k) \end{bmatrix}\Big)^T\Big]$$

# Overview

- The Stochastic Least Squares Problem
- Asymptotic observer
- State reconstruction in the presence of noise
- **The Conventional Kalman filter (SLS framework)**
- Case Study: Level estimation in liquid tanks
- The Kalman filter problem as a weighted least squares problem.

## The conventional KF: measurement update

Theorem: Let the conditions of KF problem hold, then the minimum variance unbiased estimate for $x(k)$ is given as,

$$\hat{x}(k|k) = P(k|k-1)C(k)^T\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)^{-1}y(k) +$$

$$\Big\{I - P(k|k-1)C(k)^T\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)^{-1}C(k)\Big\}\hat{x}(k|k-1)$$

and has covariance matrix,

$$E\Big[\Big(x(k) - \widehat{x}(k|k)\Big)\Big(x(k) - \widehat{x}(k|k)\Big)^T\Big] =$$

$$P(k|k-1) - P(k|k-1)C(k)^T\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)^{-1}C(k)P(k|k-1)$$

## The conventional KF: time & measurement update

Theorem: Let the conditions of KF problem hold, then the minimum variance unbiased estimate for $x(k+1)$ is given as $\hat{x}(k+1|k) =$,

$$= \underbrace{\Big(A(k)P(k|k-1)C(k)^T + S(k)\Big)\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)^{-1}}_{K(k)} y(k) +$$

$$B(k)u(k) + \Big(A(k) - K(k)C(k)\Big)\hat{x}(k|k-1)$$

& has covariance matrix $E\Big[\Big(x(k+1) - \widehat{x}(k+1|k)\Big)\Big(x(k+1) - \widehat{x}(k+1|k)\Big)^T\Big]$,

$$P(k+1|k) = A(k)P(k|k-1)A(k)^T + Q(k)$$

$$-\Big(A(k)P(k|k-1)C(k)^T + S(k)\Big)\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)^{-1}\Big(\bullet\Big)^T$$

## Sketch of the proof (step 1)

1. Unbiased estimation property: Using the DATA EQUATION,

$$\begin{bmatrix} \hat{x}(k|k) \\ \hat{x}(k+1|k) \end{bmatrix} = \begin{bmatrix} M_{11}C(k) + M_{12}A(k) & -M_{12} \\ M_{21}C(k) + M_{22}A(k) & -M_{22} \end{bmatrix} \begin{bmatrix} x(k) \\ x(k+1) \end{bmatrix}$$

$$+ \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} v(k) \\ w(k) \end{bmatrix} + \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} \hat{x}(k|k-1)$$

Taking the mean, and using notation $E[x(k)] = \overline{x}_k, C(k) = C$, etc.

$$\begin{bmatrix} \overline{x}_k \\ \overline{x}_{k+1} \end{bmatrix} = \begin{bmatrix} M_{11}C + M_{12}A & -M_{12} \\ M_{21}C + M_{22}A & -M_{22} \end{bmatrix} \begin{bmatrix} \overline{x}_k \\ \overline{x}_{k+1} \end{bmatrix} + \begin{bmatrix} N_1 \\ N_2 \end{bmatrix} \overline{x}_k$$

$$\Leftrightarrow \begin{bmatrix} I - M_{11}C - M_{12}A - N_1 & M_{12} \\ -M_{21}C - M_{22}A - N_2 & I + M_{22} \end{bmatrix} \begin{bmatrix} \overline{x}_k \\ \overline{x}_{k+1} \end{bmatrix} = 0$$

## Conclusion from the Unbiasedness Condition (step 1)

$$\Leftrightarrow \begin{bmatrix} I - M_{11}C - M_{12}A - N_1 & M_{12} \\ -M_{21}C - M_{22}A - N_2 & I + M_{22} \end{bmatrix} \begin{bmatrix} \overline{x}_k \\ \overline{x}_{k+1} \end{bmatrix} = 0$$

1. $M_{12} = 0$
2. $I - M_{11}C - N_1 = 0 \quad \Rightarrow N_1 = I - \boxed{M_{11}}C$
3. $M_{22} = -I$
4. $-M_{21}C + A - N_2 = 0 \quad \Rightarrow N_2 = A - \boxed{M_{21}}C$

## Sketch of the proof (step 2)

**2.** Expression for Covariance matrices to be minimized:

$$\begin{bmatrix} x(k) - \hat{x}(k|k) \\ x(k+1) - \hat{x}(k+1|k) \end{bmatrix} = \begin{bmatrix} I - M_{11}C(k) \\ A(k) - M_{21}C(k) \end{bmatrix} \Big(x(k) - \hat{x}(k|k-1)\Big)$$

$$- \begin{bmatrix} M_{11} & 0 \\ M_{21} & -I \end{bmatrix} \begin{bmatrix} v(k) \\ w(k) \end{bmatrix}$$

Since $\epsilon$ and $\Big(x(k) - \hat{x}(k|k-1)\Big)$ are uncorrelated,

$$E[\begin{bmatrix} x(k) - \hat{x}(k|k) \\ x(k+1) - \hat{x}(k+1|k) \end{bmatrix} \begin{bmatrix} x(k) - \hat{x}(k|k) \\ x(k+1) - \hat{x}(k+1|k) \end{bmatrix}^T] =$$

## Sketch of the proof (step 2, Ct'd)

$$\begin{bmatrix} I - M_{11}C(k) \\ A(k) - M_{21}C(k) \end{bmatrix} P(k|k-1) \begin{bmatrix} I - M_{11}C(k) \\ A(k) - M_{21}C(k) \end{bmatrix}^T$$

$$+ \begin{bmatrix} M_{11} & 0 \\ M_{21} & -I \end{bmatrix} \begin{bmatrix} R(k) & S(k)^T \\ S(k) & Q(k) \end{bmatrix} \begin{bmatrix} M_{11}^T & M_{21}^T \\ 0 & -I \end{bmatrix}$$

Therefore, the covariance matrix to be minimized for the *measurement update* is,

$$E\Big[(x(k) - \hat{x}(k|k))(x(k) - \hat{x}(k|k))^T\Big] =$$

$$P(k|k-1) - M_{11}C(k)P(k|k-1) - P(k|k-1)C(k)^T M_{11}^T +$$

$$M_{11}\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)M_{11}^T$$

## Sketch of the proof (step 3, Ct'd)

**3.** Minimizing the covariance matrix w.r.t. $M_{11}$ ($M_{12}$) by an application of the C.O.S. argument. This yields,

$$
\begin{aligned}
M_{11} &= P(k|k-1)C(k)^T\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)^{-1} \\
M_{21} &= \Big(A(k)P(k|k-1)C(k)^T + S(k)\Big) \\
&\quad \times\Big(C(k)P(k|k-1)C(k)^T + R(k)\Big)^{-1}
\end{aligned}
$$

## Recap: The conventional KF

Theorem: Let the conditions of KF problem hold, then the minimum variance unbiased estimate for $x(k+1)$ is given as $\hat{x}(k+1|k) =$,

$$= \underbrace{\left(A(k)P(k|k-1)C(k)^T + S(k)\right)\left(C(k)P(k|k-1)C(k)^T + R(k)\right)^{-1}}_{K(k)} y(k) +$$

$$B(k)u(k) + \left(A(k) - K(k)C(k)\right)\hat{x}(k|k-1)$$

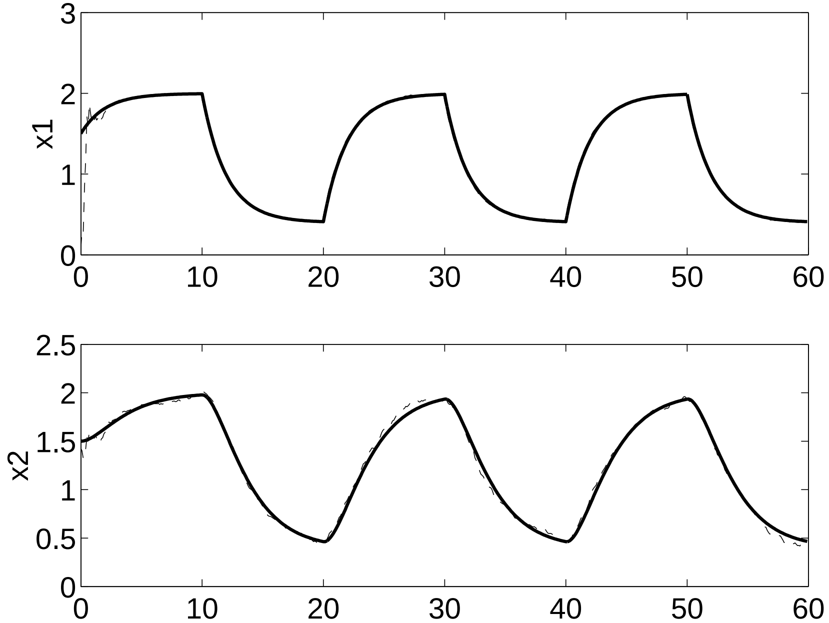
& has covariance matrix $E\left[\left(x(k+1) - \widehat{x}(k+1|k)\right)\left(x(k+1) - \widehat{x}(k+1|k)\right)^T\right]$,

$$P(k+1|k) = A(k)P(k|k-1)A(k)^T + Q(k)$$

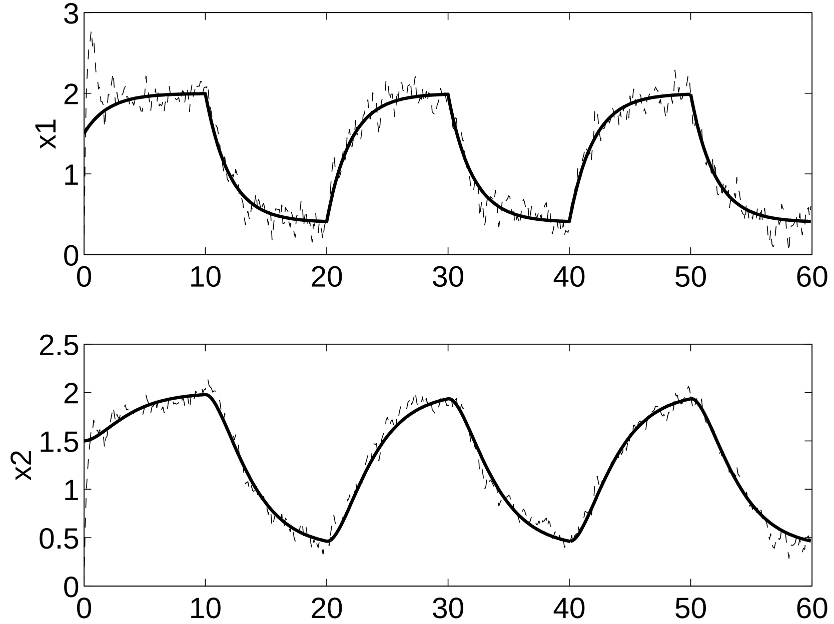
$$-\left(A(k)P(k|k-1)C(k)^T + S(k)\right)\left(C(k)P(k|k-1)C(k)^T + R(k)\right)^{-1}\left(\bullet\right)^T$$

# Example: Double tank with noisy data

Reliable Kalman Filter




Asymptotic observer



## What have we achieved?

The application of the “completion of the squares” principle to derive the convential Kalman filter recursions.

More needs to be (can be) said.

1. Is the Kalman filter a weighted least squares problem?
2. A reliable (square root) implementation?
3. Is their more “insight” to be gained from the algorithmic framework?

# Overview

- The Stochastic Least Squares Problem
- Asymptotic observer
- State reconstruction in the presence of noise
- The Conventional Kalman filter (SLS framework)
- Case Study: Level estimation in liquid tanks
- **The Kalman filter problem as a weighted least squares problem**

## The KF in the WLS framework (1)

Recall the Data Equation of the Kalman filter problem:

$$\underbrace{\begin{bmatrix} y(k) \\ -B(k)u(k) \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} C(k) & 0 \\ A(k) & -I \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} x(k) \\ x(k+1) \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} v(k) \\ w(k) \end{bmatrix}}_{L(k)\epsilon(k) \quad \epsilon(k)\sim(0,I)}$$

What is the matrix $L(k)$? This follows from the covariance expression:

$$E[\begin{bmatrix} v(k) \\ w(k) \end{bmatrix}] = \begin{bmatrix} R(k) & S(k)^T \\ S(k) & Q(k) \end{bmatrix} = L(k)E[\epsilon(k)\epsilon(k)^T]L(k)^T$$

So $L(k)$ is the “square root” of the measurement-, process noise covariance matrix!

## Determining the "square root" of a matrix

The square root $L(k)$ of the measurement-, process noise covariance matrix,

$$\begin{bmatrix} R(k) & S(k)^T \\ S(k) & Q(k) \end{bmatrix} = L(k)L(k)^T \geq 0 \quad \text{and} \quad R(k) > 0$$

is given by the solution of Exercise 5.2.

$$L(k) = \begin{bmatrix} R(k)^{1/2} & 0 \\ X(k) & Q_x(k)^{1/2} \end{bmatrix}$$

with $X(k)$ and $Q_x(k)$ satisfying

$$\begin{aligned} X(k) &= S(k)R(k)^{-T/2}, \\ Q_x(k) &= Q(k) - S(k)R(k)^{-1}S(k)^T \end{aligned}$$

# The KF in the WLS framework (2)

The prior on $x(k)$, denoted as:

$$x(k) = \hat{x}(k|k-1) - P(k|k-1)^{1/2}\tilde{x}(k) \quad \tilde{x}(k) \sim (0, I)$$

in combination with the Kalman filter DATA EQUATION leads to the constraint of a weighted least squares problem.

$$\begin{bmatrix} \widehat{x}(k|k-1) \\ y(k) \\ -B(k)u(k) \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ C(k) & 0 \\ A(k) & -I_n \end{bmatrix} \begin{bmatrix} x(k) \\ x(k+1) \end{bmatrix}$$

$$+ \begin{bmatrix} P(k|k-1)^{1/2} & 0 & 0 \\ 0 & R(k)^{1/2} & 0 \\ 0 & X(k) & Q_x(k)^{1/2} \end{bmatrix} \underbrace{\begin{bmatrix} \widetilde{x}(k) \\ \widetilde{v}(k) \\ \widetilde{w}(k) \end{bmatrix}}_{\mu(k)} (Eq.1)$$

Then the weighted least squares problem is denoted as,

$$\min_{x(k), x(k+1)} \mu(k)^T \mu(k) \quad \text{subject to} \quad Eq.(1).$$

# Preparation for this afternoon

Preparation:

1. Study Chapter 4 (4.5.3 - 4.5.4) for the SLS problem.
2. Study Chapter 5 (5.1-5.5.1) for the KF problem.

Download Homework 2