Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Introduction

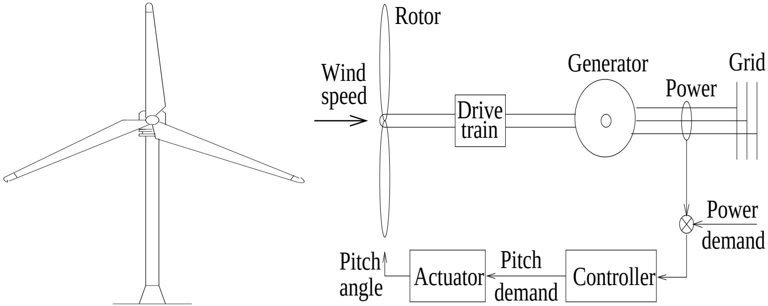
Objective (L1)

- What is in the course
- Dynamic systems

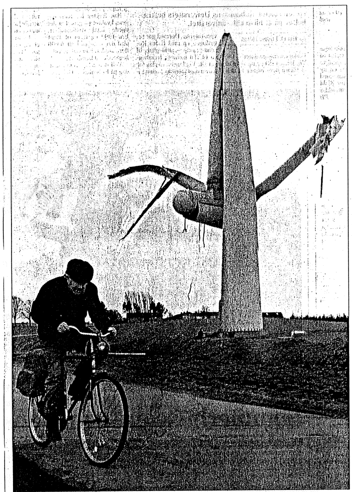
Course Contents

- Stochastic process and systems
- Filter and Control design (SS and trf)
- System identification
- Adaptive control

Wind turbine



Wind turbine

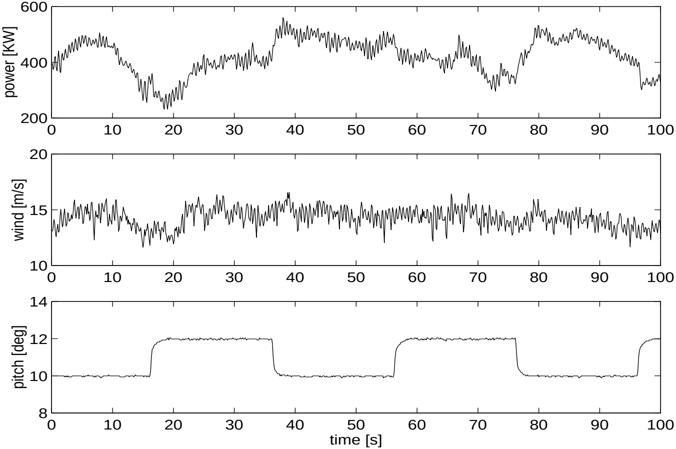


Mølle fik et knæk: Den kraftige storm greb i går fat i en af Nordvestsjællands Energiforsynings tre vindmøller i Gierslev i Vestsjælland. Mølletårnet knækkede på midten, og vingerne fik slået stumper af spidserne. Eksperter fra Risø Vindmøllepark skal nu klarlægge årsagen til kæmpemøllens sørgelige endeligt. Isenkram for tre mill. kr. lå i går spredt ud over det forblæste område – selv generatoren er ìaldet af den tidligere energikilde. *hil*

Foto: Søren Steffen

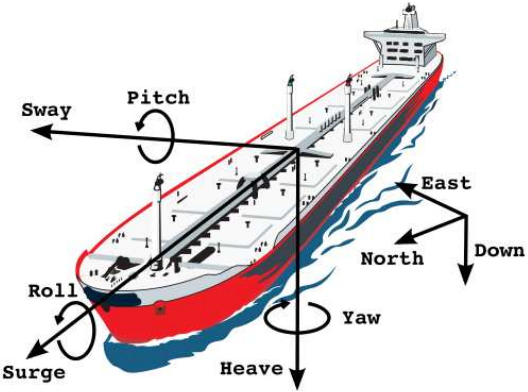
containsverbatim]

Wind turbine



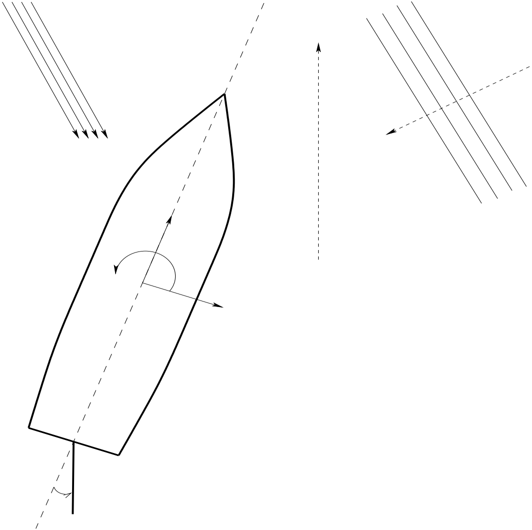
Stochastic caused by weather

Surface Vessel



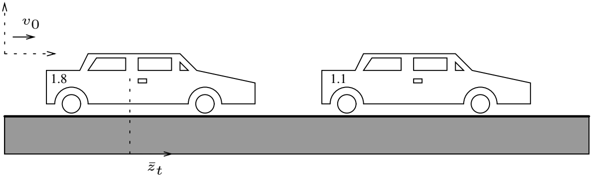
From: C. Holden, Roberto Galeazzi, C. Rodrguez, T. Perez, T. I. Fossen, M. Blanke, M. A. S. Neves. Nonlinear Container Ship Model for the Study of Parametric Roll Resonance Modeling, Identification and Control, 28, pp. 87-113, 2007.

Surface Vessel



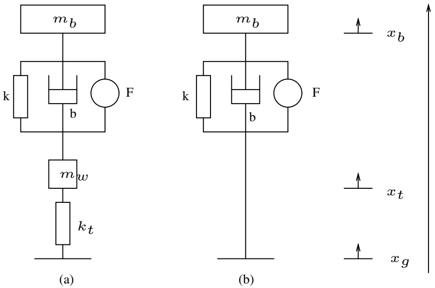
Stochastic caused by weather

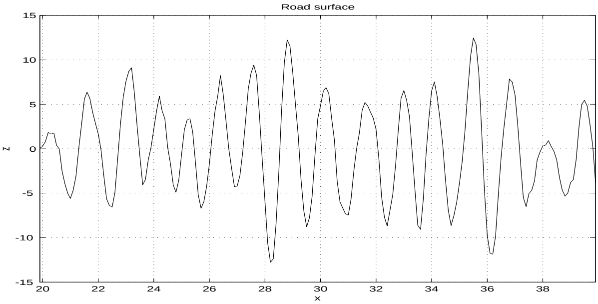
Follow that car - but don't hit it

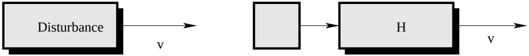


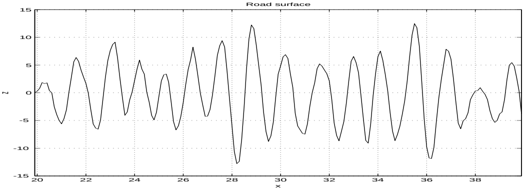
Stochastics caused by human activity

Active suspension



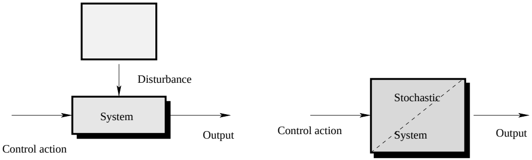


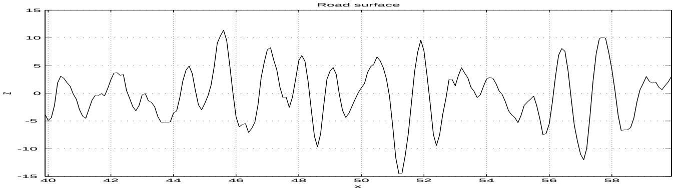




$$\frac{d}{dt_c}x_t = Ax_t + v_t \qquad y_t = H(\frac{d}{dt_c})e_t$$

$$y_t = Cx_t + e_t$$

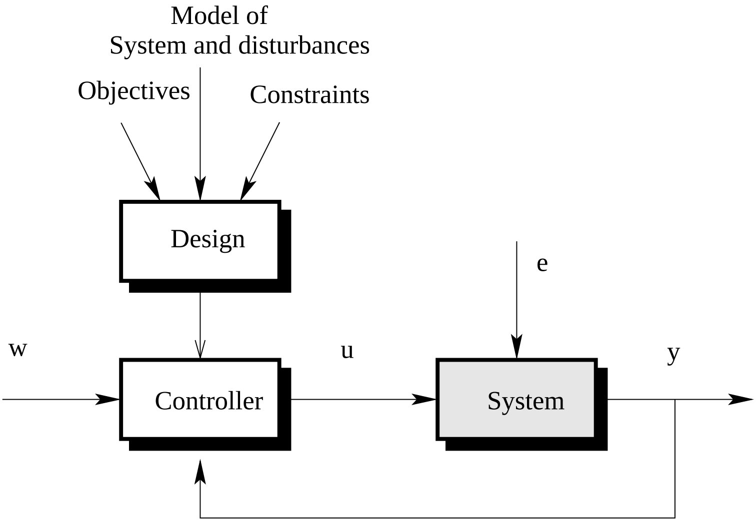


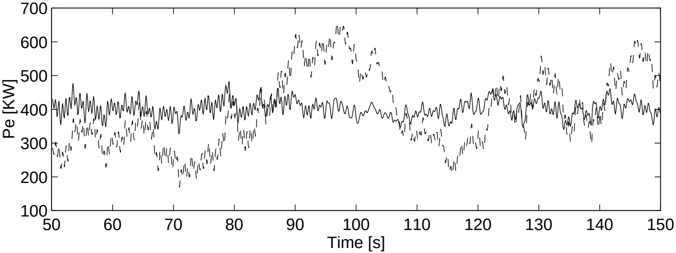


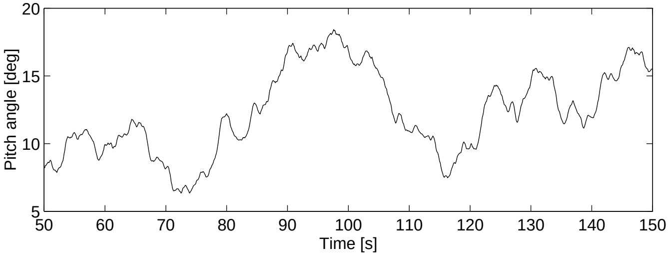
$$\frac{d}{dt_c} x_t = A x_t + B u_t + v_t \qquad y_t = G\left(\frac{d}{dt_c}\right) u_t + H\left(\frac{d}{dt_c}\right) e_t$$

$$y_t = C x_t + e_t$$

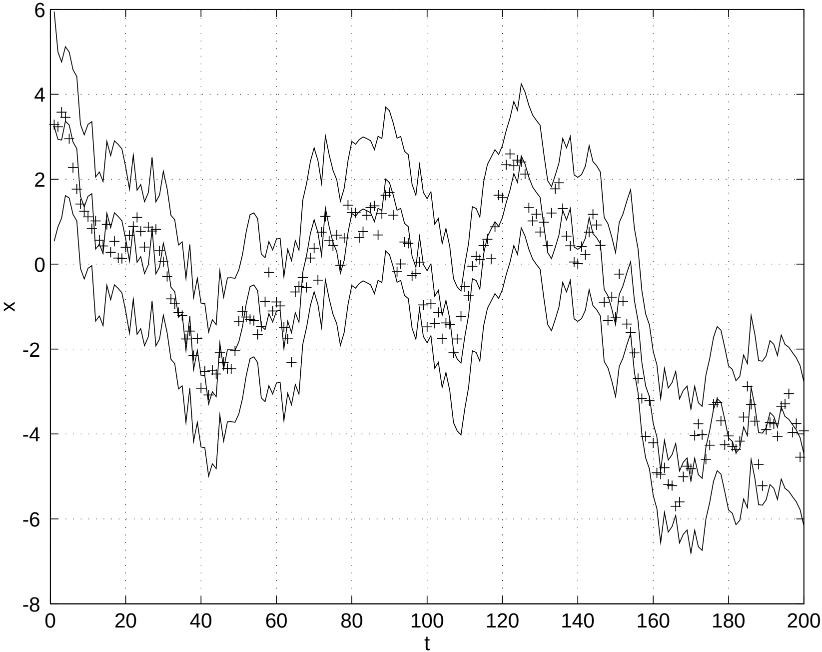
Control design



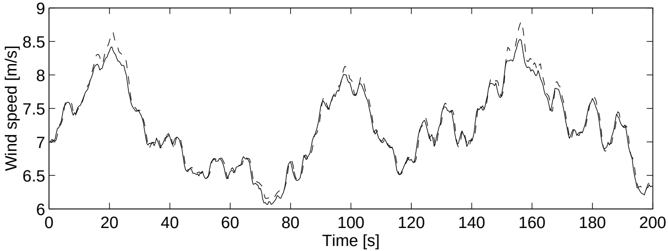

700
600
500
400
300
200
100
Pe [KW]
50 60 70 80 90 100 110 120 130 140 150
Time [s]

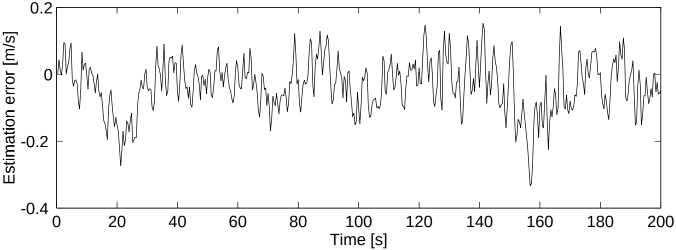

20
15
10
5
Pitch angle [deg]
50 60 70 80 90 100 110 120 130 140 150
Time [s]

State estimation (Kalman filtering)



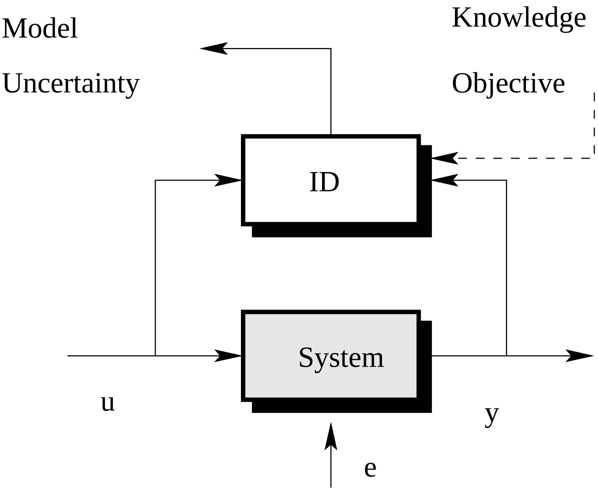
Kalman filtering



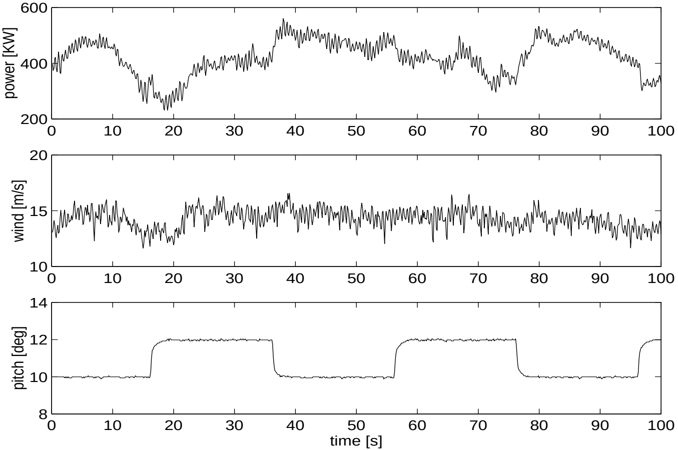


System identification

Controller for stochastic systems do require information on the system and the stochastic disturbances.



Wind turbine



$$\frac{d}{dt_c}x_t = Ax_t + Bu_t + v_t$$

$$y_t = Cx_t + e_t$$

$$y_t = G(\frac{d}{dt_c})u_t + H(\frac{d}{dt_c})e_t$$

System identification

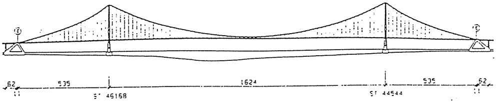
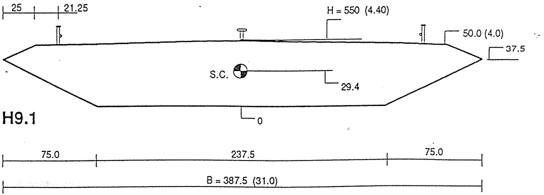
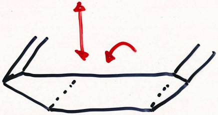
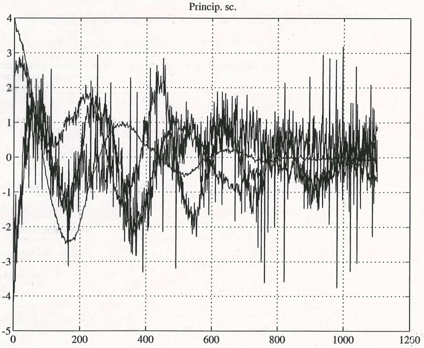


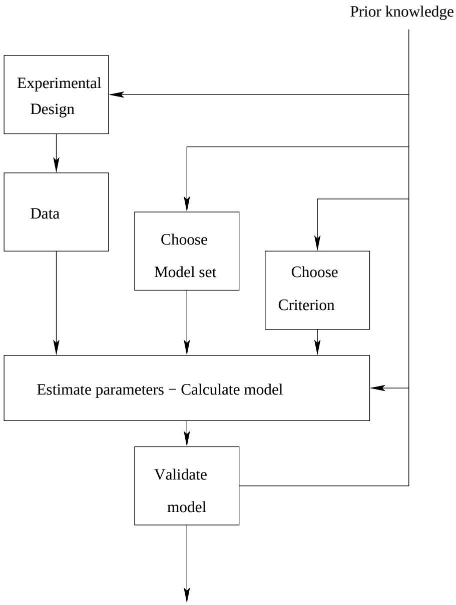
Figure 1. Schematic Drawing of Great Belt East Bridge.



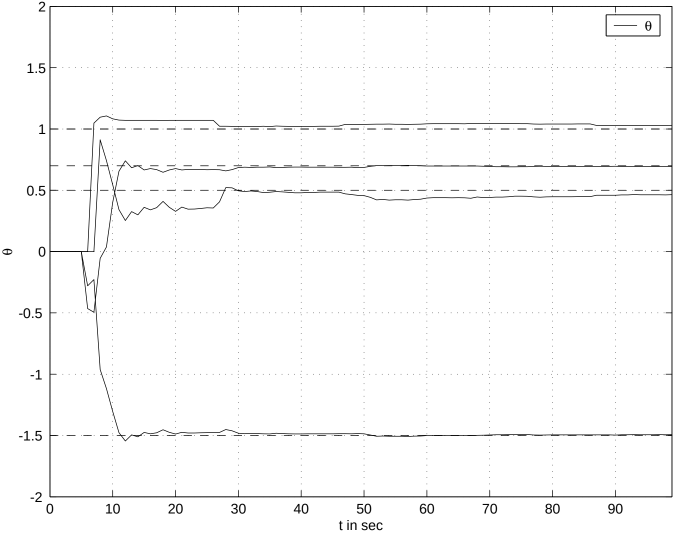
Model dimensions are given in mm.



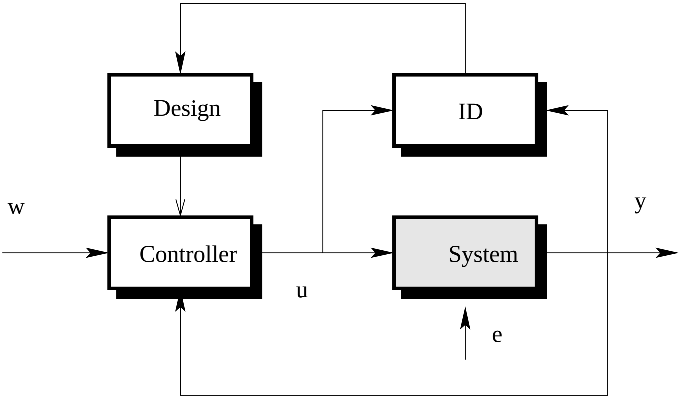

Princip. sc.
4
3
2
1
0
-1
-2
-3
-4
-5
0
200
400
600
800
1000
1200

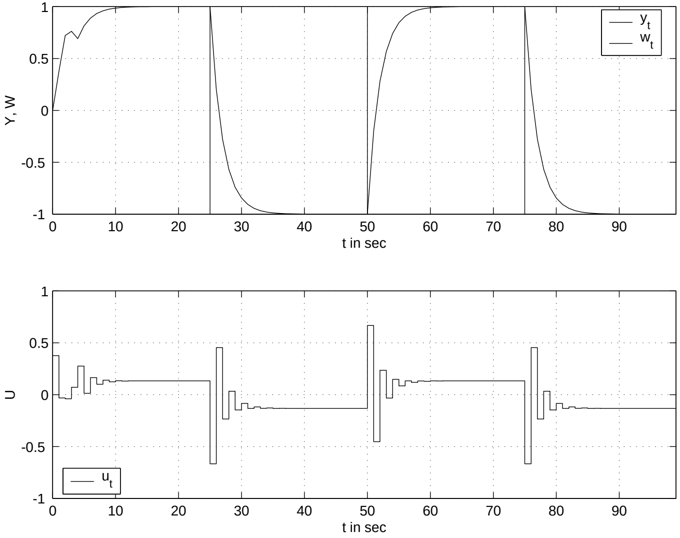

Prior knowledge
Experimental Design
Data
Choose Model set
Choose Criterion
Estimate parameters − Calculate model
Validate model

System identification



Time variation, nonlinearities





1
0.5
0
-0.5
-1
Y, W
0
10
20
30
40
50
60
70
80
90
t in sec
y_t
w_t
1
0.5
0
-0.5
-1
U
0
10
20
30
40
50
60
70
80
90
t in sec
u_t

02421: Stochastic Adaptive Control.

Spring 2011:

The purpose of 02421 Stochastic Adaptive Control (former 04332 and 0414) is to give the students knowledge of methods for modelling and control of Dynamic Stochastic Systems. In more details the focus is on:

- *Modelling and analysis of stochastic systems (ie. dynamic systems which are influenced by stochastic disturbances).*
- *Control of stochastic systems.*
- *Identification of stochastic systems (i.e. estimation of unknown parameters) and, finally*
- *Adaptive control of stochastic systems (i.e. simultaneous identification and control).*

The Lectures will take place in F3 (Tuesdays 8-12 and Friday 12-17) in room 205, build 305 and the exercises in the G-bar (room 115 and 221, building 305).

Lecture: Niels Kjølstad Poulsen (nkp), DTU Informatics.
Teaching assistant: Mahmood Mirzaei (mmir), DTU Informatics.

Further information:

- *Introduction*
- *Course description (in english) and in danish*
- *Course schedule (Timetable)*
- *Lectured material in english and danish. This list will develop during term.*
- *Course material*
- *Foils and slides*
- *Exercises and projects*
- *Toolbox as zip file. (Last updated 2.12.2009)*

Information from an earlier version of the course:

- *Exercises (and solutions from E00) Not part of the course any longer.*

Further information available at: Niels Kjølstad Poulsen ,IMM Bldg. 321, DTU
Tel.: 4525 3356/ Fax: 4588 2673, E-mail: nkp@imm.dtu.dk

Last update Jan 29 2010

Time table - 02421

1: **Topic:** Introduction, System theory
Danish Litt.: 5-8 (Introduction), 46-52 (LTI systems).
English Litt.: Å+W: Chapter 2
Excercise: 1.

2: **Topic:** System theory
Danish Litt.: 52-72 (State transformations, Poles and zeroes).
V:487-488 (Z-transform)
English Litt.: Å+W: Chapter 2
Excercise: 2.

3: **Topic:** System theory
Danish Litt.: 71-72 (Stability), 73-86 (Control- andobservability).
502 (definite matrices)
English Litt.: Å+W: Chapter 3, p. 77-79 (Stability),
93-103 (controlability and observability).
Excercise: 3

4: **Topic:** Stochastic Processes I
Danish Litt.: 124-135 (Stochastic variable). Not theorem 3.4.
135-146 mid (Stochastic vectors)
English Litt.: Jz: 8-42. Go easy on page 37-38 and
on characteristic functions
Excercise: 4

5: **Topic:** Stochastic Processes II
Danish Litt.: 176-240 (stochastic processes),
209-216 (Internal process models).
English Litt.: Jz: p. 47-56 (stochastic processes),
p.85-90 (Internal process models).
Excercise: 5

6: **Topic:** Stochastic Systems and State Estimation
Danish Litt.: 229-240 (Stochastic Systems),
146-153 (Projection Theorem) [not proof of Theorem 3.19, 3.20],
249-260 (State Estimation and Kalman filter) [not proof of
Theorem 7.7, 7.8]
265-273 (Estimation error and Stationarity)
English Litt.: Å70: 210-215 (Filtering) [not proof of theorem 2.1],
218-221 (Projection Theorem) [not proofs and not Theorem 3.3],
225-233 (Kalman filter) [not algebraic proof of theorem 4.2].

Lectured material - 02421
2008

1: (30.1.2007): **Topic:** Introduction, System theory
Danish Litt.: 5-8 (Introduction), 46-52 (LTI systems).
English Litt.: Å+W: Chapter 2
Excercise: 1.

ÅW: K.J. Åström and B. Wittenmark (1984): Computer Controlled Systems

Å+W: K.J. Åström and B. Wittenmark (1997): Computer Controlled Systems, Theory and design

T-book K.J. Åström and B. Wittenmark (1995): Adaptive Control

DV: M.H.A. Davis and R.B. Vinter (1985): Stochastic Modelling and Control

Jz: A. H. Jazwinsky (1970): Stochastic Processes and Filtering Theory.

TS: Torsten Söderström and Petre Stoica (1989): System Identification

Å70: K.J.Åström (1970): Introduction to Stochastic Control Theory.

LL: Lennart Ljung (1999): System Identification - Theory for the user

xreg: Niels Kjølstad Poulsen (2004): Stochastic Control, External models

Course material for Stochastic Adaptive Control (02421)

In danish:

- Lærebog: Stokastisk adaptiv regulering af Niels Kjølstad Poulsen, Lyngby 2007.
  In B5 format in ps and in pdf. **Notice page references are given with respect to the B5 version.**
  In A4 format in ps and in pdf.

The book (in B5 format) can be bought at the IMM book store (300 kr) or be downloaded (for free).
Check opening hours for the IMM book store

In english:

- Stochastic control theory, External System description (B5, pdf), (B5, ps), (A4, pdf), (A4, ps),
- Chapters from the Litteratur (will be announced at the lectures)

Suplementary material:

- Mark Gockenback: A Practical Introduction to MATLAB (as ps) or (as html)

For further information contact, Niels Kjølstad Poulsen, IMM Build. 321, DTU
Phone: +45 4588 3356 / Fax: +45 4588 2673, E-mail: nkp@imm.dtu.dk

Stochastic Adaptive Control

You can downlaod foils and slides from the following (developing) list. Please notice, not all foils are here and not all foils here are a part of the curriculum. The slides are in pdf as slideshow or in ps (2 by 2) for printing.

1. **Introduction (pdf, ps)**
2. **Dynamic systems (pdf ps)**
3. **Stochastic vetors and variable (pdf ps)**
4. **Internal process models and systems (pdf ps)**
5. **State estimation and Kalman filtering (pdf ps)**
6. **Control of systems (internal or state space models) (pdf ps)**
7. **External process models and systems (pdf ps)**
8. **Control of systems (external or transfer function models) (pdf ps)**
9. **System identification (pdf ps)**
10. **Model validation (pdf ps)**
11. **Recursive estimation (pdf ps)**
12. **Adaptive control (pdf ps)**
13. **Closed loop identification (pdf ps)**
14. **Experimental design (pdf ps)**

- **The total slidesshow collected in one file (pdf ps)**

Stochastic Adaptive Control
02421 - Exercises and Projects in F2006.

For each exercise (or project) there exists a problem formulation, some m-files in a distribution (ie. a zipped directory) and in most cases a solution (in terms of a set of m-files). A total collection of m files and solutions can be found below.

A practicel notice in connection to the solutions: Have the distribution (where you work with the exercise) and the solution in separate directories. When examining the solutions (be in that directory and) include the distribution with the **path** command. For a automatic view of the solutions, just type **solution** (in Matlab and in the directory containing the downloaded solutions).

Released exercises: 1-24 (the rest is just for reference and from last term)

1. *Exercise part 1: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
2. *Exercise part 2: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
3. *Exercise part 3: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
4. *Exercise part 4: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
5. *Exercise part 5: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
6. *Exercise part 6: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
7. *Exercise part 7: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
8. *Exercise part 8: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
9. *Exercise part 9: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
10. *Exercise part 10: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
11. *Exercise part 11: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
12. ***Project** part 12: Text (in ps format), m-files for downloading (zipped m-files).*
13. *Exercise part 13: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
14. *Exercise part 14: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
15. *Exercise part 15: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
16. *Exercise part 16: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
17. *Exercise part 17: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
18. *Exercise part 18: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
19. *Exercise part 19: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
20.
21. *Exercise part 21: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
22. *Exercise part 22: Text (in ps format), m-files for downloading and solutions (zipped m-files).*
23.
24. ***Project** part 24: Text (in ps format), m-files for downloading (zipped m-files).*
25.
26.

What to do (in order to fetch the m-files on a UNIX platform):

1. *Press SHIFT LEFT_MOUSE_BUTTOM (ie. click on m-files)*
2. *unzip -a dist1.zip*

Course outline

- Lecture 2 hours (approx)
- Exercise 2 hours (approx)

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

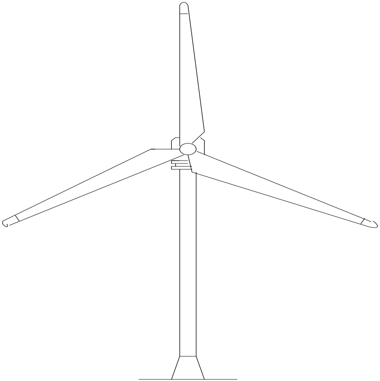
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Dynamic Systems

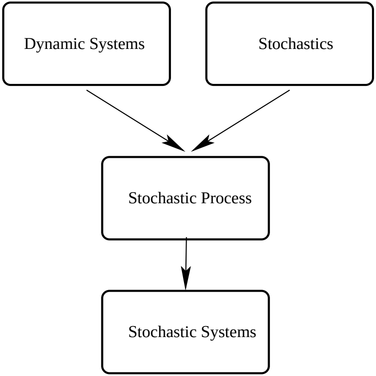
Dynamic Systems



▸ L2
▸ L3

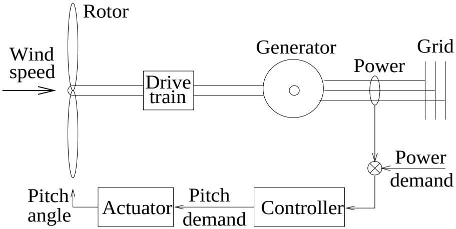
DTU

The local view (ie. the horizon until Stochastic Systems).

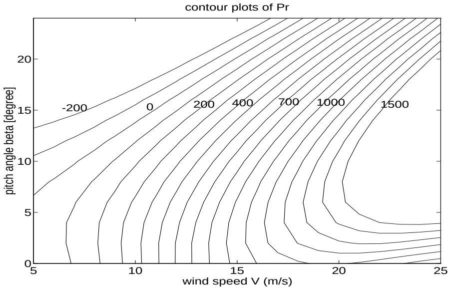


Example: Wind Turbine

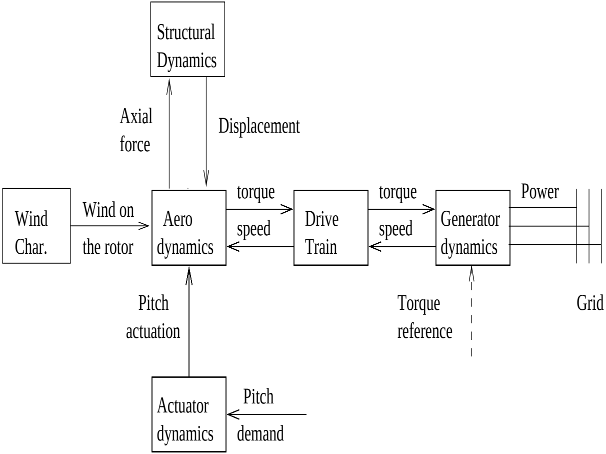
Fixed speed mode:



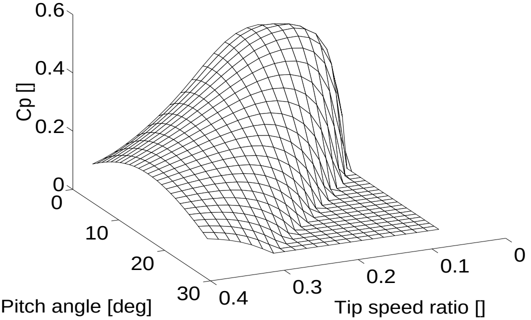
Control Objective:



Wind turbine



$$P_r = \frac{1}{2}\rho\pi R^2 v^3 \;\; C_p(\lambda, \beta) \qquad \lambda = \frac{v}{\omega_r R} \qquad T_r = \frac{P_r}{\omega_r}$$



Lesson learned: Wind turbine is a nonlinear system.

Turbine model

$$P_r = \frac{1}{2}\rho\pi R^2 v^3 C_p(\lambda, \beta) \qquad \lambda = \frac{v}{\omega_r R} \qquad T_r = \frac{P_r}{\omega_r}$$

$$P_e = \omega_g T_g$$

$$J_r\dot{\omega}_r = {\color{red}T_r} - N_g T_g \qquad \dot{\omega}_r = f(\omega_r, \begin{bmatrix} \beta \\ T_g \end{bmatrix}, v)$$

$$y = \begin{bmatrix} P_e \\ \omega_g \end{bmatrix} \qquad y_m = \omega_g$$

Lesson learned: Wind turbine is a first order nonlinear system (in the simple version).

Continuous time systems

States	x			
inputs	u	Controls	or	Disturbance
outputs	y	Measured outputs	or	controlled outputs

$$\begin{aligned} \dot{x} &= f(x,u) \qquad x_0 = \underline{x}_0 \\ y &= g(x,u) \end{aligned}$$

$x \in \mathbb{R}^n \qquad y \in \mathbb{R}^m \qquad u \in \mathbb{R}^p$

Stationary point and Linearization

Assume U_0 is constant. Then X_0 is a stationary point if:

$$\begin{aligned} 0 &= f(X_0, U_0) \\ Y_0 &= g(X_0, U_0) \end{aligned}$$

Let $x = X - X_0$ be the deviation away from the stationary point.

Linear description

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad\qquad x_0 = \underline{x}_0 \\ y &= Cx + Du \end{aligned}$$

Linearization

$$\begin{aligned} A &= \frac{\partial}{\partial x} f \qquad\qquad B = \frac{\partial}{\partial u} f \\ C &= \frac{\partial}{\partial x} g \qquad\qquad D = \frac{\partial}{\partial u} g \end{aligned}$$

DTU

Solution to ODE (LTI system)

For known input function, the LTI system

$$\dot{x}_t = Ax_t + Bu_t \qquad x_0 = \bar{x}_0$$

has the solution:

$$x_t = e^{At}\bar{x}_0 + \int_0^{t} e^{A(t-s)} Bu_s ds$$

$$e^M = I + M + \frac{1}{2}M^2 + ... + \frac{1}{n!}M^n + ...$$ Definition of Matrix exponential

---

Check:

$$\frac{d}{dt}x_t = Ae^{At}\bar{x}_0 + Bu_t + A\int_0^t e^{A(t-s)}Bu_s ds = Ax_t + Bu_t$$

$$y_t = Cx_t + Du_t$$

$$y_t = Ce^{At}\bar{x}_0 + \int_0^t Ce^{A(t-s)}Bu_s ds + Du_t$$

Impulseresponse

$$
\begin{aligned}
y_t &= Ce^{At}\bar{x}_0 + \int_0^t Ce^{A(t-s)}Bu_s ds + Du_t \\
&= Ce^{At}\bar{x}_0 + \int_0^t h_{t-s}u_s ds \\
&= Ce^{At}\bar{x}_0 + h_t \star u_t
\end{aligned}
$$

Impulse response :

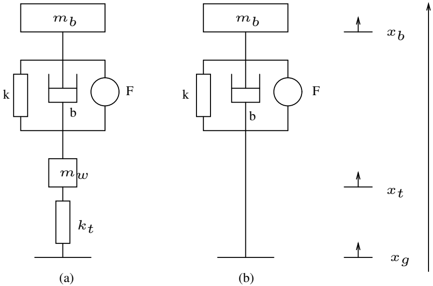
$$
h_t = \begin{cases} 0 & \text{for } t < 0 \\ Ce^{At}B & \text{for } t > 0 \end{cases}
$$

$$
h_t = D\delta(t) + Ce^{At}B \;\; u_h(t)
$$

Step response :

$$
s_t = \int_0^t h_s ds
$$

The quarter car model



External description (apply Newton II):

$$m_b \ddot{y}_b + b\dot{y}_b + ky_b = k_1 u$$

From external to internal and back

External description:

$$m_b\ddot{y}_b + b\dot{y}_b + ky_b = k_1 u$$

From external to internal description: Choose a set of system states.

$$x_1 = y_b \qquad x_2 = \dot{y}_b \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\dot{x}_1 = \dot{y}_b = x_2 \tag{1}$$

$$\dot{x}_2 = \ddot{y}_b = -\frac{k}{m_b}y_b - \frac{b}{m_b}\dot{y}_b + \frac{k_1}{m_b}u \tag{2}$$

$$= -\frac{k}{m_b}x_1 - \frac{b}{m_b}x_2 + \frac{k_1}{m_b}u \tag{3}$$

Internal description

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m_b} & -\frac{b}{m_b} \end{pmatrix} x + \begin{pmatrix} 0 \\ \frac{k_1}{m_b} \end{pmatrix} u \tag{4}$$

$$y = \begin{pmatrix} 1 & 0 \end{pmatrix} x \tag{5}$$

From internal to external: 1) eliminate system states. 2) See later (for a formal method).

From external to internal and back

Transfer operator, H

$$m_b\ddot{y}_b + b\dot{y}_b + ky_b = k_1 u$$

$$\left[m_b \frac{d^2}{dt^2} + b\frac{d}{dt} + k\right] y_b = k_1 u$$

$$y_b = H(\frac{d}{dt})u \qquad H(\frac{d}{dt}) = \frac{k_1}{m_b\frac{d^2}{dt^2} + b\frac{d}{dt} + k}$$

Descriptions of dynamic systems

In time domain, ODE

$$\frac{d}{dt}$$

$$y = H(\tfrac{d}{dt})u$$

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

Laplace transform

$$Y(s) = \mathcal{L}[y(t)] = \int_0^\infty y(t)e^{-st}dt$$

$$\mathcal{L}[\frac{df(t)}{dt}] = sF(s) - f(0) \qquad \mathcal{L}[f_1(t) * f_2(t)] = F_1(s)F_2(s)$$

Fourier transform, Bode plot, frequency gain and phase

$$Y(\omega) = \int_{-\infty}^\infty y(t)e^{-j\omega t}dt \qquad Y(\omega) \leftrightarrow y(t)$$

$$\mathcal{F}\left[\frac{d}{dt}f(t)\right] = j\omega F(\omega) \qquad \mathcal{F}[f_1 * f_2] = F_1(\omega)F_2(\omega)$$

Time domain

$$\frac{d^n}{dt^n} y + a_1 \frac{d^{n-1}}{dt^{n-1}} y + \ \dots\ + a_n y = b_1 \frac{d^{n-1}}{dt^{n-1}} u + \ \dots\ + b_n u$$

$$y = H(\frac{d}{dt})\, u$$

Frequency domain

$$H_f(\omega) = \frac{b_1 (j\omega)^{n-1} + \ \dots\ + b_n}{(j\omega)^n + a_1 (j\omega)^{n-1} + \ \dots\ + a_n}$$

Laplace domain

$$H(s) = \frac{b_1 s^{n-1} + \ \dots\ + b_n}{s^n + a_1 s^{n-1} + \ \dots\ + a_n}$$

$$H(\frac{d}{dt}) = C \left[\frac{d}{dt} I - A \right]^{-1} B + D$$

Eigenvalues and poles

Timedomain - external

$$y = H(\frac{d}{dt})u$$

Laplace - external

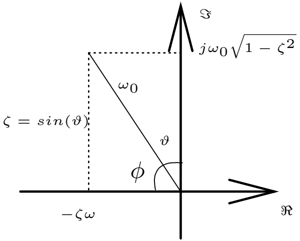
$$H(s) = \frac{b_1 s^{n-1} + \ \dots \ b_n}{s^n + a_1 s^{n-1} + \ \dots \ a_n}$$

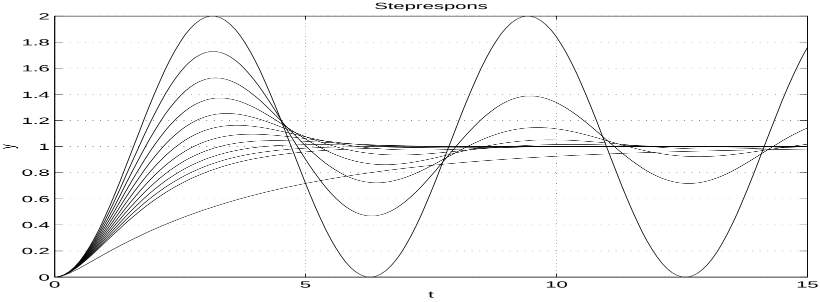
Internal (state space)

$$\frac{d}{dt}x = Ax + Bu$$
$$y = Cx + Du$$

NB: Same system.

$$\begin{aligned} \mathbf{C}(s) &= Det(sI - A) = s^n + a_1 s^{n-1} + \ \dots \ a_n \\ &= \prod_i (1 + s\tau_i) \prod_k (s^2 + 2\zeta_k \omega_k s + \omega_k^2) \end{aligned}$$


$\Im$
$j\omega_0\sqrt{1-\zeta^2}$
ω_0
$\zeta = sin(\vartheta)$
ϑ
ϕ
$-\zeta\omega$
$\Re$


Steprespons
y
t
0
5
10
15
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

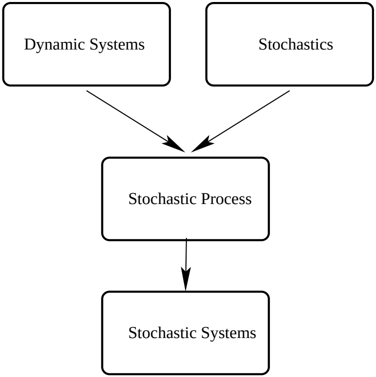
Lesson learned

- Stationary points
- Linearization
- Solution to LTI systems
- Impulse, step response
- Internal <-> external description
- Time domain, Frequency domain and Laplace domain
- Pole/eigen value <-> reponse

End of L1

The roadmap

The local view (ie. the horizon until Stochastic Systems).



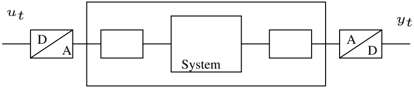
Lesson learned in L1

Continuous time descriptions:

- Stationary points
- Linearization
- Solution to LTI systems
- Impulse, step response
- Internal <-> external description (in continuous time)
- Time domain, Frequency domain and Laplace domain
- Pole/eigen value <-> reponse

An illustrive time axis.

Sampling



Continuous time ($\tau \in \mathbb{R}$) description (of system, actuators, sensors and filters)

$$\begin{aligned} \frac{d}{d\tau} x_\tau &= A_c x_\tau + B_c u_\tau; \qquad x_{\tau_0} = \underline{x}_0 \\ y_\tau &= C x_\tau + D u_\tau \end{aligned}$$

Solution

$$x_\tau = e^{A_c(\tau - \tau_0)} x_0 + \int_{\tau_0}^{\tau} e^{A_c(\tau - s)} B_c u_s ds$$

Let $t \in \mathbb{Z}$, and let $\tau = (t+1)T_s$, $\tau_0 = tT_s$ and using ZOH.

Discrete time description:

$$\begin{aligned} x_{t+1} &= e^{A_c T_s} x_t + \left(\int_0^{T_s} e^{As} B_c ds \right) u_t \\ y_t &= C x_t + D u_t \end{aligned}$$

DTU

Discrete time systems

$$x_{t+1} = Ax_k + Bu_t \qquad x_0 = \underline{x}_0 \tag{6}$$

$$y_t = Cx_t + Du_t \tag{7}$$

where:

$$A = e^{A_c T_s} \qquad B = \int_0^{T_s} e^{A_c s} B_c ds$$

$$C = C_c \qquad D = \begin{cases} D_c & \text{if controlled output} \\ 0 & \text{if measured output} \end{cases}$$

Notice: same order - and same states (at sample instants)

Notice: c2d i.e. $[A, B] = c2d(A_c, B_c, T_s)$ in matlab or tables.

$$N_r = \frac{t_r}{T_s} \in [2;4] \qquad > \frac{1}{\pi} \quad (Shannon\ \ limit)$$

Claude Elwood Shannon

- Born: April 30, 1916 Petoskey, Michigan, United States
- Died: February 24, 2001 (aged 84) Medford, Massachusetts, United States
- Residence: United States
- Institutions: Bell Laboratories Massachusetts Institute of Technology Institute for Advanced Study
- Alma mater: University of Michigan Massachusetts Institute of Technology

Assume we can go from an internal

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t \qquad x_0 = \underline{x}_0 \\ y_t &= Cx_t + Du_t \end{aligned}$$

to external description (we will later see how).

Here u_t and y_t are sequences of numbers (vectors), and t in y_t is an integer.

An example on an External description

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} = b_0 u_{t-1} + b_1 u_{t-2}$$

Introduce the shift operator

$$q y_t = y_{t+1} \qquad q^{-1} y_t = y_{t-1}$$

$$y_t + a_1 q^{-1} y_t + a_2 q^{-2} y_t = b_0 q^{-1} u_t + b_1 q^{-2} u_t$$

$$y_t + a_1 q^{-1} y_t + a_2 q^{-2} y_t = b_0 q^{-1} u_t + b_1 q^{-2} u_t$$ (just copied from last slide)

$$\left(1 + a_1 q^{-1} + a_2 q^{-2}\right) y_t = \left(b_0 q^{-1} + b_1 q^{-2}\right) u_t$$

$$A_p(q^{-1}) y_t = \bar{B}_p(q^{-1}) u_t$$

$$y_t = \frac{\bar{B}_p(q^{-1})}{A_p(q^{-1})} u_t = \frac{b_0 q^{-1} + b_1 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}} = H(q) u_t$$

$$\bar{B}_p(q^{-1}) = b_0 q^{-1} + b_1 q^{-2} = q^{-1}\left(b_0 + b_1 q^{-1}\right)$$

$$\bar{B}_p(q^{-1}) = q^{-k} B_p(q^{-1})$$

But

$$\left(1 + a_1 q^{-1} + a_2 q^{-2}\right) y_t = \left(b_0 q^{-1} + b_1 q^{-2}\right) u_t \qquad \text{(just copied from last slide)}$$

can also be written as:

$$\left(q^2 + a_1 q + a_2\right) y_t = \left(b_0 q + b_1\right) u_t$$

or in short:

$$\mathbf{A}_p(q) y_t = \mathbf{B}_p(q) u_t$$

or as a transfer operator

$$y_t = H(q) u_t$$

where

$$H(q) = \frac{\mathbf{B}_p(q)}{\mathbf{A}_p(q)} = \frac{b_0 q + b_1}{q^2 + a_1 q + a_2}$$

Notice:

$$H(q) = \frac{\mathbf{B}_p(q)}{\mathbf{A}_p(q)} = \frac{\bar{B}_p(q^{-1})}{A_p(q^{-1})} \qquad \frac{b_0 q + b_1}{q^2 + a_1 q + a_2} = \frac{b_0 q^{-1} + b_1 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}}$$

The Z-transform

Definition:

$$Y(z) = \mathcal{Z}(y_t) = \sum_{t=0}^{\infty} y_t z^{-t}$$

Actually only just a special version of the LaPlace transform, but for discrete time signals.

$$z = e^{sT_s} \qquad s = \sigma + j\omega$$

Properties

$$\mathcal{Z}(q y_t) = zY(z) - y_0$$

$$\mathcal{Z}(h_t \star u_t) = H(z)U(z)$$

$$\mathcal{Z}(a y_t + u_t) = aY(z) + U(z)$$

$$y_t = H(q)u_t \qquad Y(z) = H(z)U(z)$$

$$H_f(\omega) = H(e^{j\omega T_s})$$

zeroes, poles and time delays

Fourier transform

Definition:

$$Y_f(\omega) = \mathcal{F}(y_t) = \sum_{t=-\infty}^{\infty} y_t e^{-j\omega T_s t}$$

$$z = e^{sT_s} \qquad s = \sigma + j\omega$$

Properties

$$\mathcal{F}(qy_t) = e^{j\omega T_s}\; Y_f(\omega)$$

$$\mathcal{F}(h_t \star u_t) = H_f(\omega)U_f(\omega)$$

$$\mathcal{F}(ay_t + u_t) = aY_f(\omega) + U_f(\omega)$$

$$y_t = H(q)u_t \qquad Y_f(\omega) = H_f(\omega)U_f(\omega)$$

$$H_f(\omega) = H(e^{j\omega T_s}) \qquad H_f(f) = H_c(j\omega)$$

DC-gain (amplification)

Internal and External descriptions

We can go from an external (transfer function) description such as:

$$y_t + a_1 y_{t-1} + a_2 y_{t-2} = b_0 u_{t-1}$$

or as

$$y_{t+1} + a_1 y_t + a_2 y_{t-1} = b_0 u_t$$

to an internal description by using a proper choice of states:

$$x^1 = y_t \qquad x^2 = y_{t-1}$$

We have

$$\begin{aligned} qx^1 = y_{t+1} &= b_0 u_t - a_1 y_t - a_2 y_{t-1} \\ &= b_0 u_k - a_1 x^1 - a_2 x^2 \end{aligned}$$

$$qx^2 = y_t = x^1$$

or the internal (state space) description:

$$\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_{t+1} = \begin{bmatrix} -a_1 & -a_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_t + \begin{bmatrix} b_0 \\ 0 \end{bmatrix} u_t$$

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_t + 0\, u_t$$

$$\begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_{t+1} = \begin{bmatrix} -a_1 & -a_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_t + \begin{bmatrix} b_0 \\ 0 \end{bmatrix} u_t \qquad \text{(just a copy)}$$

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}_t + 0\, u_t$$

$$x_{t+1} = Ax_t + Bu_t$$

$$y_t = Cx_t + Du_t$$

The other way, from internal (state space) to external (transfer function) model is more straight forward

$$H(q) = \frac{B_p}{A_p} = C(qI - A)^{-1}B + D$$

$$y_t = H(q)u_t$$

Matlab, Canonical forms

DTU

State space transformations

Consider the description

$$x_{t+1} = Ax_t + Bu_t$$

$$y_t = Cx_t + Du_t$$

and the transformation:

$$z_t = \mathbf{T}x_t$$

assume $\mathbf{T}^{-1}$ *exists*

Then the following model has the same input output relation

$$z_{t+1} = \mathbf{T}A\mathbf{T}^{-1}z_t + \mathbf{T}Bu_t$$

$$y_t = C\mathbf{T}^{-1}z_t + Du_t$$

and the transformation is denoted a similarity transformation :

$$\begin{aligned} A_z &= \mathbf{T}A\mathbf{T}^{-1} \qquad B_z = \mathbf{T}B \\ C_z &= C\mathbf{T}^{-1} \qquad D_z = D \end{aligned}$$

It preserves eigenvalues, orders a.o.

Lot of state space descriptions has the same dynamic input-output relation.

Diagonal form

$$\mathbf{T}A\mathbf{T}^{-1} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \qquad \mathbf{T}B = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

$$\begin{aligned} z^1_{t+1} &= \lambda_1 z^1_t + \beta_1 u_t \\ z^2_{t+1} &= \lambda_2 z^2_t + \beta_2 u_t \\ &\vdots \\ z^n_{t+1} &= \lambda_n z^n_t + \beta_n u_t \end{aligned}$$

$$y_t = c_1 z_1(t) + c_2 z_2(t) + \dots + c_n z_n(t)$$

When does $\mathbf{T}$ exists. Jordan forms. How can λ_i be found. Implication in relation to open loop stability.

Canonical forms

Consider the external description:

$$y_t + a_1 y_{t-1} + ... + a_n y_{t-n} = \bar{b}_1 u_{t-1} + ... + \bar{b}_n u_{t-n}$$

or the transfer operator:

$$H(q) = \frac{\bar{b}_1 q^{-1} + ... + \bar{b}_n q^{-n}}{1 + a_1 q^{-1} + ... + a_n q^{-n}} = \sum_{i=1}^{\infty} h_i q^{-i}$$ *(expansion or pulse response)*

We will now discuss how to realize such a transfer function, i.e. to a find a suitable internal model.

We will espically discuss Minimal realizations and properties such as observability, controlability and stability.

The direct term $b_0 = h_0$ is here set to zero for simplicity.

Controller canonical form

$$A_c = \begin{bmatrix} -a_1 & \dots & -a_{n-1} & -a_n \\ 1 & \dots & 0 & 0 \\ & \ddots & & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix} \qquad B_c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C_c = (\bar{b}_1, \bar{b}_2, \dots, \bar{b}_n) \qquad D_c = 0$$

For $\bar{b}_0 \neq 0$:

$$A_c = \begin{bmatrix} -a_1 & \dots & -a_{n-1} & -a_n \\ 1 & \dots & 0 & 0 \\ & \ddots & & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix} \qquad B_c = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C_c = (\bar{b}_1 - \bar{b}_0 a_1, \bar{b}_2 - \bar{b}_0 a_2, \dots, \bar{b}_n - \bar{b}_0 a_n) \qquad D_c = b_0$$

Observer canonical form

for $\bar{b}_0 = 0$

$$A_o = \begin{bmatrix} -a_1 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ -a_{n-1} & 0 & \dots & 1 \\ -a_n & 0 & \dots & 0 \end{bmatrix} \quad B_o = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \vdots \\ \bar{b}_n \end{bmatrix}$$

$$C_o = (1, 0, \dots, 0) \qquad D_o = 0$$

Controlability canonical form

for $\bar{b}_0 = 0$

$$A_{co} = \begin{bmatrix} 0 & \dots & 0 & -a_n \\ 1 & \dots & 0 & -a_{n-1} \\ & \ddots & & \vdots \\ 0 & \dots & 1 & -a_1 \end{bmatrix} \qquad B_{co} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C_{co} = (h_1, h_2, \dots, h_n) \qquad D_{co} = 0$$

Observability canonical form

for $\bar{b}_0 = 0$

$$A_{ob} = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & \dots & -a_1 \end{bmatrix} \quad B_{ob} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$$

$$C_c = (1, 0, \dots, 0) \qquad D_{ob} = 0$$

$$A_o = A_c^T \qquad B_o = C_c \qquad C_o = B_c$$

$$A_{ob} = A_{co}^T \qquad B_{ob} = C_{co} \qquad C_{ob} = B_{co}$$

$$D_c = D_o = D_{co} = D_{ob} = \bar{b}_0 = h_0$$

Direct realization

$$y_t + a_1 y_{t-1} + ... + a_n y_{t-n} = \bar{b}_1 u_{t-1} + ... + \bar{b}_n u_{t-n}$$

$$y_t = -a_1 y_{t-1} - ... - a_n y_{t-n} + \bar{b}_1 u_{t-1} + ... + \bar{b}_n u_{t-n}$$

$$x_t^T = (-y_{t-1}, ..., -y_{t-n}, u_{t-1}, ..., u_{t-n}) \qquad C_d = (a_1, ..., a_n, \bar{b}_1, ..., \bar{b}_n)$$

$$y_t = C_d x_t$$

$$x_t^T = (-y_{t-1}, -y_{t-2} ..., -y_{t-n}, u_{t-1}, u_{t-2} ..., u_{t-n})$$

$$\Longrightarrow$$

$$x_{t+1}^T = (-y_t, -y_{t-1}, ..., -y_{t+1-n}, u_t, u_{t-1} ..., u_{t+1-n})$$

$$\Uparrow \qquad \Uparrow$$

DTU

Direct realization

$$A_d = \left\{ \begin{array}{cccc|cccc} a_1 & \dots & a_{n-1} & a_n & \bar{b}_1 & \dots & \bar{b}_{n-1} & \bar{b}_n \\ 1 & & 0 & 0 & 0 & \dots & 0 & 0 \\ & \ddots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & & 1 & 0 & 0 & \dots & 0 & 0 \\ \hline 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \hline 0 & \dots & 0 & 0 & 1 & & 0 & 0 \\ \vdots & & \vdots & \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & & 1 & 0 \end{array} \right\} \qquad B_d = \left\{ \begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ \hline 1 \\ \hline 0 \\ \vdots \\ 0 \end{array} \right\}$$

$$C_d = (a_1, \dots, a_n, \bar{b}_1, \dots, \bar{b}_n) \qquad D_d = 0$$

Poles and zeroes

$$
\begin{aligned}
H(z) &= \frac{b_0 + b_1 z^{-1} + \ \dots \ + b_n z^{-n}}{1 + a_1 z^{-1} + \ \dots \ + a_n z^{-n}} = \frac{b_0 z^n + b_1 z^{n-1} + \ \dots \ + b_n}{z^n + a_1 z^{n-1} + \ \dots \ + a_n} \\
&= \prod \frac{(z - z_k)(z^2 + 2\mu_i \omega_i z + \omega_i^2)}{(z - p_l)(z^2 + 2\mu_j \omega_j z + \omega_j^2)}
\end{aligned}
$$

Link to Internal description/State Space

$$\mathcal{C}(z) = \mathbf{A}_p(z) = Det(zI - A) = z^n + a_1 z^{n-1} + \dots + a_n$$

$$\mathbf{B}(z) = C \ \ adj\{zI - A\}B + \mathbf{A}_p(z)D = b_0 z^n + \dots + b_n$$

Same story in C-time.

Transformation and sampling

$\lambda_d = e^{\lambda_c T_s}$ *goes for poles, but not for zeroes (only a assyptotic result)*

$$A_c = \left\{ \begin{array}{cccc} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{array} \right\}$$

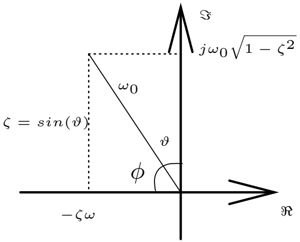
$$\begin{aligned} A_d &= e^{A_c T_s} = I + A_c T_s + \frac{1}{2} A_c^2 T^2 + \ldots \\ &= \left\{ \begin{array}{cccc} 1 + \lambda_1 T_s + \frac{1}{2}\lambda_1^2 T_s^2 + \ldots & 0 & \ldots & 0 \\ 0 & x & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x \end{array} \right\} = \left\{ \begin{array}{cccc} e^{\lambda_1 T_s} & 0 & \ldots & 0 \\ 0 & e^{\lambda_2 T_s} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & e^{\lambda_n T_s} \end{array} \right\} \end{aligned}$$

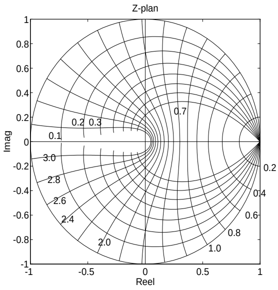
DTU

Transformation of Real Poles

$$\lambda_c = -\sigma \qquad \sigma \in \mathbb{R}$$

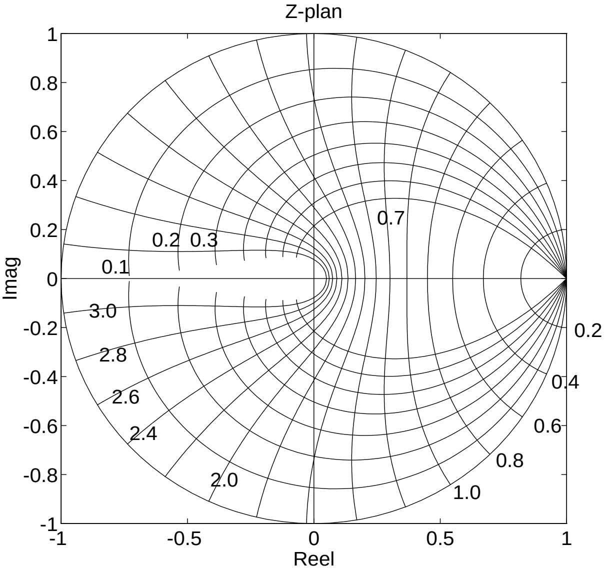
$$\lambda_d = e^{-\sigma T_s}$$





Where is $\sigma = 0$ and $\sigma \to \infty$?

Poles



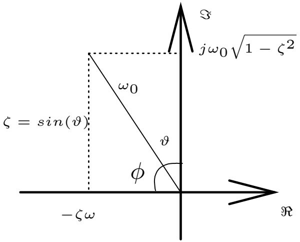
$\omega_0 T_s$ and ζ.

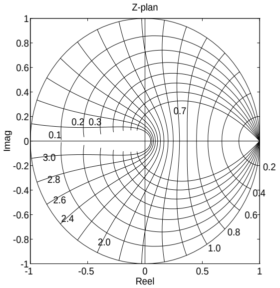
Complex poles

$$\lambda_c = -\zeta\omega_0 \pm j\omega_0\sqrt{1-\zeta^2} \qquad \mathcal{C}(s) = s^2 + 2\zeta\omega_0 s + \omega_0^2$$

$$\lambda_d = e^{\lambda_c T} = e^{-\zeta\omega_0 T \pm j\omega_0 T\sqrt{1-\zeta^2}} = e^{-\zeta\omega_0 T}\left(cos(\omega_0 T\sqrt{1-\zeta^2}) \pm j sin(\omega_0 T\sqrt{1-\zeta^2})\right)$$

$$\rho = e^{-\zeta\omega_0 T} \qquad \vartheta = \omega_0 T\sqrt{1-\zeta^2}$$





Prelude to Zeroes

For ZOH we have exactly that:

$$H_c(s) = \frac{1}{s^n} \qquad\qquad H(z) = \frac{T_s^n}{n} \frac{B_n^\star}{(z-1)^n}$$

$$B_n^\star(z) = b_1^n z^{n-1} + b_2^n z^{n-2} + \cdots + b_n^n$$

n	$B_n^\star(z)$	zeros in $B_n^\star(z)$ outside UD.
1	1	
2	$z+1$	-1
3	z^2+4z+1	-3.732
4	$z^3+11z^2+11z+1$	-1, -9.899
5	$z^4+26z^3+66z^2+26z+1$	-2.322, -23.2
6	$z^5+57z^4+302z^3+302z^2+57z+1$	-1, -4.5542,-51.22

DTU

Zeroes:

For the zeroes the problem is a bit more complex than for poles.

We have only asymptotic results ($T_s \to 0$).

- In continuous time we have $m < n$ zeroes, but in discrete time we have (for $D = 0$) $n - 1$ zeroes. The $n - 1 - m$ zeroes are denoted as sampling zeroes.

- $m : z_d \to e^{z_c T_s}$
- $n - m - 1 : z_d \to$ zeros in B^*_{n-m}

- For $n_k - m_k \geq 2$ there exists T_s such that $|z_i| > 1$

(Open loop) LTI stability

$$H(q) = \frac{\mathbf{B}_p(q)}{\mathbf{A}_p(q)}$$

$$x_{t+1} = Ax_t + Bu_t$$

$$y_t = Cx_t + Du_t$$

$$\mathcal{C}(z) = \mathbf{A}_p(z) = Det\,(zI - A)$$

- All poles/eigenvalues inside unit disk $\rightarrow$ asymptotic stable
- One pole/eigenvalue outside unit disk $\rightarrow$ unstable
- Poles on the unit circle $\rightarrow$ multiplicity plays a role

LTI stability

In general:

$$m_a \geq m_g$$

If $m_a > m_g$ then system is unstable and if $m_a = m_g$ then system is stable (but not asymptotic stable).

Algebraic multiplicity

$$Det(zI - A) = (z - \lambda)^{m_a}(...)$$

Geometric multiplicity

m_g is the number of linearly independent eigenvectors to one eigenvalue (λ).

$$Av = \lambda v$$

Lesson learned

- Sampling
- Shift operator, Z-transform and Fourier
- Internal vs. External descriptions
- Similarity transformation
- Canonical forms
- Poles and zeros (sampling process)
- LTI stability

End of L2

Learning objectives in L3

- Where do we go, if we apply a certain input sequence?

- How shall we steer, if we want to go to a specific point?
  - Where do we want to go.
  - To the origin ?
  - To an arbitrary point in the state space ?
- Can we steer the system to any place? from any places?
- How much does it costs?
- What is the optimal input sequence?

- Do we know where we are?
- Can we (from output measurements) determine the state of the system?

DTU

Response of a State Space Sytem (where do we go, if we steer in a certain way)

$$x_{t+1} = Ax_t + Bu_t \qquad x_0$$

$$\begin{aligned}
x_1 &= Ax_0 + Bu_0 \\
x_2 &= A{\color{red}x_1} + Bu_1 = A\Big(Ax_0 + Bu_0\Big) + Bu_1 \\
&= A^2x_0 + ABu_0 + Bu_1 \\
x_3 &= A{\color{red}x_2} + Bu_2 \\
&= A^3x_0 + A^2Bu_0 + ABu_1 + Bu_2 \\
\vdots \quad & \quad \vdots \\
x_t &= A^tx_0 + A^{t-1}Bu_0 + A^{t-2}Bu_1 + \cdots + ABu_{t-2} + Bu_{t-1}
\end{aligned}$$

$$x_t = A^tx_o + \begin{bmatrix} A^{t-1}B & \cdots & AB & B \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_{t-2} \\ u_{t-1} \end{bmatrix}$$

DTU

$$U_{t-1} = (u_0,\ u_1,\ \ldots\ u_{t-2},\ u_{t-1})^\top$$

$$\mathcal{W}_c(t) = \left(A^{t-1}B,\ A^{t-2}B,\ \ldots\ AB,\ B\right) = \left\{\ \vdots \quad \vdots \quad \vdots\ \right\}$$

$$x_t = A^t x_0 + \mathcal{W}_c(t) U_{t-1}$$

Alternatively:

$$x_t = A^t x_0 + \sum_{s=0}^{t-1} A^{t-s-1} B u_s$$

DTU

The Output Equation

$$x_{t+1} = Ax_t + Bu_t$$
$$y_t = Cx_t + Du_t$$

$$x_t = A^t x_0 + \mathcal{W}_c(t) U_{t-1} = A^t x_o + \begin{bmatrix} A^{t-1}B & \cdots & AB & B \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_{t-2} \\ u_{t-1} \end{bmatrix}$$

(just a copy)

$$y_t = CA^t x_0 + \begin{bmatrix} CA^{t-1}B & \cdots & CAB & CB \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_{t-2} \\ u_{t-1} \end{bmatrix} + Du_t$$

Impulse response

$$Y_t = (y_0, \ldots, y_{t-1}, y_t)^\top$$

$$Y_t = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^t \end{bmatrix} x_0 + \begin{bmatrix} D & 0 & \ldots & 0 \\ CB & D & \ldots & 0 \\ CAB & CB & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{t-1}B & CA^{t-2}B & \ldots & D \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_t \end{bmatrix}$$

$$Y_t = \mathcal{W}_o(t)x_0 + \Pi_t U_t$$

$$U_t = (u_0, u_1, ..., u_{t-1}, u_t)^\top$$

$$Y_t = (y_0, ..., y_{t-1}, y_t)^\top$$

$$\mathcal{W}_o(t) = \left\{ \begin{array}{c} C \\ CA \\ \vdots \\ CA^t \end{array} \right\} = \left\{ \begin{array}{c} ... \\ ... \\ ... \end{array} \right\} \qquad \Pi_t = \left\{ \begin{array}{cccccc} h_0 & 0 & 0 & 0 & ... & 0 \\ h_1 & h_0 & 0 & 0 & ... & 0 \\ h_2 & h_1 & h_0 & 0 & ... & 0 \\ h_3 & h_2 & h_1 & h_0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_t & h_{t-1} & h_{t-2} & h_{t-3} & ... & h_0 \end{array} \right\}$$

DTU

Cayley-Hamiltons Theorem

$$\mathcal{C}(s) = Det(sI - A) = s^n + a_1 s^{n-1} + ... + a_n$$

$$\mathcal{C}(\lambda_i) = 0$$

$$A^n + a_1 A^{n-1} + ... + a_n I = 0$$

Results

$$A^n = -\sum_{i=0}^{n-1} a_{n-i} A^i$$

$$A^m = \sum_{i=0}^{n-1} \alpha_i A^i \quad m \geq n$$

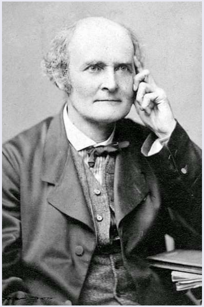
Draft paper

$$A^n = \left[\begin{array}{ccc} A^{n-1} & ... & I \end{array} \right] \otimes a$$

$$\begin{aligned} A^{n+1} &= AA^n = A \left[\begin{array}{ccc} A^{n-1} & ... & I \end{array} \right] \otimes a \\ &= \left[\begin{array}{ccc} A^n & ... & A \end{array} \right] \otimes a \\ &= \left[\begin{array}{ccc} A^{n-1} & ... & I \end{array} \right] \otimes \alpha \end{aligned}$$

$$A^m B = \sum_{i=0}^{n-1} \alpha_i A^i B$$

Arthur Cayley



- Born: 16 August 1821 Richmond, Surrey, UK
- Died: 26 January 1895 (aged 73) Cambridge, England
- Alma mater: Trinity College, Cambridge

William Rowan Hamilton

- Born: 4 August 1805 Dublin
- Died: 2 September 1865 (aged 60) Dublin
- Alma mater: Trinity College, Dublin

Controlability

$$x_t = A^t x_0 + \mathcal{W}_c U_{t-1}$$ just a copy

$$\mathcal{W}_c = \left(A^{t-1}B,\ A^{t-2}B,\ \ldots\ AB,\ B\right) = \left\{ \vdots \quad \vdots \quad \vdots \right\}$$

Controlability: The ability to take the system from any x_0 to the origin in a finite time. It might be such that $\exists k \leq n : A^k = 0$.

Reachability: from any x_0 to any x_t in finite time.
Reachable if

$$W_c(A, B) = (B, AB, \ldots, A^{n-1}B)$$

has full rank. If Reachable then controlable.

The non-reachable part of the state space is the null space for W_c^T.
$\mathcal{N} = \{x \mid W_c^T x = 0\}$. In Matlab: NC=null(Wc')).

Similarity transformation: $z = Tx$.

$$(A_z, B_z, C_z, D_z) = (\mathbf{T}A\mathbf{T}^{-1}, \mathbf{T}B, C\mathbf{T}^{-1}, D)$$

$$W_c(A_z, B_z) = \mathbf{T}W_c(A, B)$$

DTU

The Controlability matrix (W_c)

$\mathcal{A}_-^{-T}$	Controller cf.
I	Controlability cf.
$\bar{B}_p(A)\tilde{I}$	Observer cf
$M(1,n)$	Observability cf.

$$\mathcal{A}_- = \begin{Bmatrix} 1 & 0 & \dots & 0 & 0 \\ a_1 & 1 & \dots & 0 & 0 \\ a_2 & a_1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1} & a_{n-2} & \dots & a_1 & 1 \end{Bmatrix} \qquad \tilde{I} = \begin{Bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{Bmatrix}$$

$$M[1,n] = \begin{Bmatrix} h_1 & h_2 & \dots & h_{n-1} & h_n \\ h_2 & h_3 & \dots & h_n & h_{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ h_{n-1} & h_n & \dots & h_{2n-2} & h_{2n-1} \\ h_n & h_{n+1} & \dots & h_{2n-1} & h_{2n} \end{Bmatrix}$$

Stabilizability

A system is **reachable** iff the eigenvalues of $A - BL$ can be choosen abitrarely by a suitable value of L.

A system (A, B) is **stabilizable** if there exist a feedback, $u = -Lx$, such that $A - BL$ is stable. The uncontrollable part has to be asymptotic stable.

Take the system from any x_0 to any x_∞ during infinte time.

If the $rank(W_c) = r < n$ then there exists a similarity transformation which take the description into

$$A_z = \begin{bmatrix} A_c & A_{c\bar{c}} \\ 0 & A_{\bar{c}} \end{bmatrix} \quad B_z = \begin{bmatrix} B_c \\ 0 \end{bmatrix} \quad z_t = \begin{bmatrix} z_t^c \\ z_t^{\bar{c}} \end{bmatrix}$$

The uncontrollable part of the state space ($A_{\bar{c}}$) has to be asymptotic stable.

Controlability Gramian - the first glimse to MPC

Task: steer a system from x_0 to x_r during $[0;t]$.

$$x_t = A^t x_0 + \mathcal{W}_c(t) U_{t-1} = x_r$$

where:

$$U_{t-1} = (u_0,\ u_1,\ \dots u_{t-2},\ u_{t-1})^\top$$

$$\mathcal{W}_c(t) = \left(A^{t-1}B,\ A^{t-2}B,\ \dots AB,\ B\right) = \left\{ \vdots \quad \vdots \quad \vdots \right\} \qquad n \times t$$

For $t \geq n$:

$$\tilde{x}_t = x_t - A^t x_0 = \mathcal{W}_c(t) U_{t-1}$$

$$\Sigma_t^c = \mathcal{W}_c(t) \mathcal{W}_c^\top(t) \qquad n \times n$$

Reachable (controlable) iff Σ^c is non singular.

$$J = \frac{1}{2} U_{t-1}^\top U_{t-1} \qquad s.t. \qquad \mathcal{W}_c(\tau) U_{t-1} = \tilde{x}_t = x_r - x_0$$

$$U_{t-1}^\star = \mathcal{W}_c^\top(\tau) \left(\Sigma_\tau^c\right)^{-1} \tilde{x}_t \qquad = \qquad \mathcal{W}_c^\top(\tau) \left(\Sigma_\tau^c\right)^{-1} \left[x_r - A^t x_0\right]$$

$$J^\star = \tilde{x}_t^\top \left(\Sigma_\tau^c\right)^{-1} \tilde{x}_t$$

Controlability Gramian

$$\begin{aligned}\Sigma_\tau^c &= \mathcal{W}_c(\tau)\mathcal{W}_c^\top(\tau) \\ &= A^{\tau-1}BB^\top(A^{\tau-1})^\top + ... + ABB^\top A^\top + BB^\top \\ &= \sum_{i=0}^{\tau-1} A^iBB^\top(A^i)^\top\end{aligned}$$

or:

$$\Sigma_{\tau+1}^c = A\Sigma_\tau^c A^\top + BB^\top \qquad \Sigma_0^c = 0$$

Infinite horizon

$$\Sigma_\infty^c = A\Sigma_\infty^c A^\top + BB^\top$$ *Lyaponov equation*

Stability, interpretation

$$J^\star = \tilde{x}_t^\top(\Sigma_\tau^c)^{-1}\tilde{x}_t$$

Similarity transformation:

$$\Sigma_\tau^c(A_z, B_z) = T\Sigma_\tau^c(A,B)T^\top$$ special case: $T^T = T^{-1}$

DTU

Aleksandr Mikhailovich Lyapunov

Born: June 6, 1857 Yaroslavl, Imperial Russia.

Died: November 3, 1918 (aged 61).

Alma mater: Saint Petersburg State University.

$$\dot{x} = Ax + Bu$$

$$\Sigma_\tau^c = \int_0^\tau e^{A\tau} B B^\top (e^{A\tau})^\top d\tau \qquad \dot{\Sigma}_t^c = A\Sigma_t^c + \Sigma_t^c A^\top + BB^\top$$

Controlable iff Σ is non singular.

Infinite horizon

$$\Sigma^c = \int_0^\infty e^{A\tau} B B^\top (e^{A\tau})^\top d\tau \qquad 0 = A\Sigma^c + \Sigma^c A^\top + BB^\top$$

Observability

$$Y_t = \mathcal{W}_o(t)x_0 + \Pi U_t \qquad\qquad x_t = A^t x_0 + \mathcal{W}_c(t)U_{t-1}$$

$$\mathcal{W}_o(t) = \left\{ \begin{matrix} C \\ CA \\ \vdots \\ CA^t \end{matrix} \right\} = \left\{ \begin{matrix} \cdots \\ \cdots \\ \cdots \end{matrix} \right\}$$

The ability to determine x_t from observations of y_τ and u_τ ($0 \leq \tau \leq t$).

Observability: determine x_0 from Y_t.

Constructablility: determine x_t from Y_t. The problem is different if $\exists k : A^k = 0$.

$$W_o = \left\{ \begin{matrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{matrix} \right\}$$

Non-observable part of state space is the null space of W_o. (Matlab: null).

Observability and similarity transformations

$$W_o(C_z, A_z)\mathbf{T} = W_o(C, A)$$

$\tilde{I}\bar{B}_p(A)$	Controller cf.
$M(1,n)$	Controlability cf.
$\mathcal{A}_{-}^{-1}$	Observer cf
I	Observability cf.

Observability Gramian

Can we see the states in the output?
Focus on a quantitative measure of observability. How much energi is transmitted to output in period $[0, t]$ due to x_0. (For simplicity assume $U = 0$).

$$J = Y_t^\top Y_t = x_0^\top \mathcal{W}_o^\top \mathcal{W}_o x_0 \qquad \mathcal{W}_o = \mathcal{W}_o(t)$$

System observable iff

$$\Sigma_t^o = \mathcal{W}_o^\top \mathcal{W}_o$$

is non singular and

$$J^* = x_0^\top \Sigma_t^o x_0$$

Interpretation.

Since

$$Y_t = \mathcal{W}_o x_0 + \Pi U_t$$

a LS estimate of x_0 is

$$\hat{x}_0 = (\Sigma_t^o)^{-1} \mathcal{W}_o^\top \left[Y_t - \Pi U_t \right]$$

$$\Sigma_t^o = \mathcal{W}_o^\top \mathcal{W}_o$$

$$\Sigma_{t+1}^o = A^\top \Sigma_t^o A + C^T C \qquad \Sigma_0^o = 0$$

$$\Sigma_\infty^o = A^\top \Sigma_\infty^o A + C^T C$$

$$T^\top \Sigma_t^o(C_z, A_z) T = \Sigma_t^o(C, A)$$

$$\dot{x} = Ax + Bu$$

$$y = Cx$$

$$\Sigma^o_\tau = \int_0^\tau (e^{A\tau})^\top C^\top C e^{A\tau} d\tau$$

$$\dot{\Sigma}^o_\tau = A^\top \Sigma^o_\tau + \Sigma^o_\tau A + C^\top C$$

Resume

- Canceling poles and zeroes
- Minimal realizations
- Simultaneous Observable and Controlable

Stability

$$x_{t+1} = f(x_t) \qquad x_0 = \underline{x_0} \qquad\qquad \dot{x} = f_c(x)$$

Stable stationary point x_0

$$\forall \varepsilon \;\; \exists \delta, t_0 \;\; \|x_{t_0} - x_0\| < \delta \Rightarrow \|x_t - x_0\| < \varepsilon \;\; \forall t > t_0$$

Asymptotic stability if stable and

$$\|x_{t_0} - x_0\| < \delta \Rightarrow \|x_t - x_0\| \to 0 \;\; t \to \infty$$

Lyapunov method

$x_0 = 0$ for simplicity

If

- $V(x)$ is C^1 i Ω
- $V(0) = 0$
- $V(x) > 0$ for all $x \in \Omega \backslash \{0\}$

$\Delta V(x_t) = V(x_{t+1}) - V(x_t) \leq 0 \;\; x_0 = 0$ *stable*

$\Delta V(x_t) < 0 \;\; x_0 = 0$ *asymptotic stable*

Stability of LTI Systems

$$x_{t+1} = Ax_t \qquad \dot{x}(t) = A_c x(t)$$

Candidate $V(x) = x^T P x \;\; P > 0$

$$A^\top P A + Q = P \qquad A_c^\top P + P A_c = -Q$$

$x = 0$ asymptotic stable if $Q > 0$ and $P > 0$.

```
------------------------------------------------------------
 DLYAP  Discrete Lyapunov equation solver.

     X = DLYAP(A,Q) solves the discrete Lyapunov equation:

                A*X*A' - X + Q = 0

     See also  LYAP.
------------------------------------------------------------
```

Lyapunov eq has a solution $P > 0$ iff eigenvalues of A are strictly inside stability area.

Sensitivity

Why - parameter estimation ao.

$$x_{t+1} = f(x_t, \theta) \qquad x_0 = \underline{x}_0$$

$$S_t = \frac{dx_t}{d\theta} \qquad n \times p$$

$$S_{t+1} = \frac{\partial f}{\partial x} S_t + \frac{\partial f}{\partial \theta} \qquad S_0 = \frac{d\underline{x}_0}{d\theta}$$

$$\dot{x} = f(x, \theta) \qquad x_0 = \underline{x}_0$$

$$\dot{S} = \frac{\partial f}{\partial x} S + \frac{\partial f}{\partial \theta} \qquad S_0 = \frac{d\underline{x}_0}{d\theta}$$

Lesson learned in L3

- Solution to state space equations
- Controlability
  - Controlability matrix
  - Controlability Gramian, Lyaponov
  - Canonical forms
  - Similarity transformation
- Observability
- Stability <-> Lyaponov equation
- Sensitivity

Slut L3

Impulse response

$$y_t = CA^t x_0 + \begin{bmatrix} CA^{t-1}B & \cdots & CAB & CB \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_{t-2} \\ u_{t-1} \end{bmatrix} + Du_t$$

(just a copy)

$$y_t = CA^t x_0 + \sum_{s=0}^{t-1} CA^{t-s-1} B u_s + Du_t$$

▸ Output equation

$$h_t = \begin{cases} CA^{t-1}B & \text{for } t > 0 \\ D & \text{for } t = 0 \\ 0 & \text{for } t < 0 \end{cases}$$

$$y_t = CA^t x_0 + \sum_{s=0}^{t} h_{t-s} u_s$$

$$y_t = CA^t x_0 + h_t \star u_t$$

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

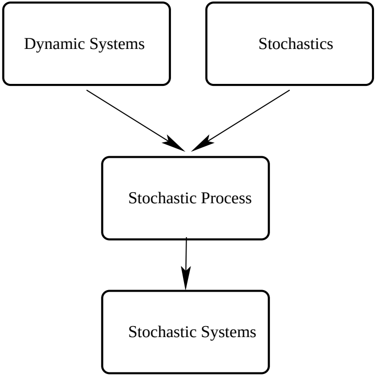
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Stochastics

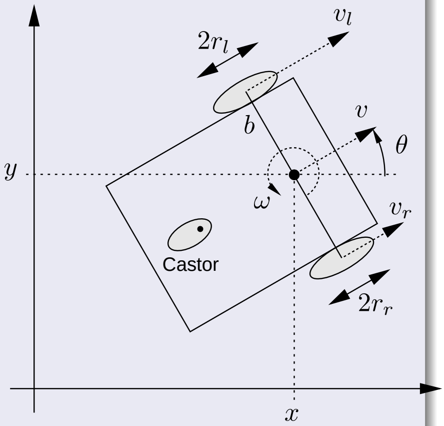
The local view (ie. the horizon until Stochastic Systems).



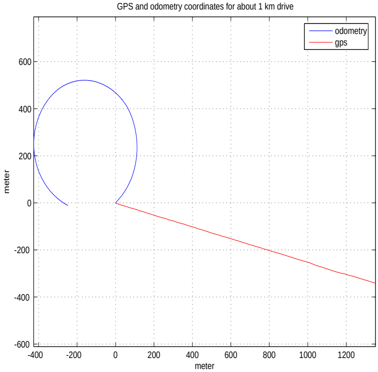
Robotics

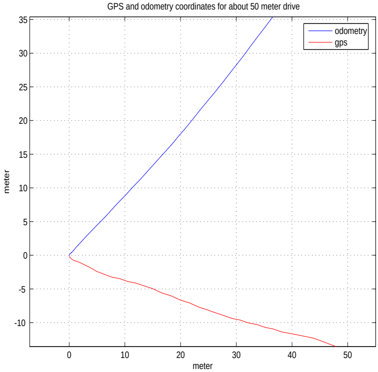
Jens Christian Andersen, DTU EE
Nils Axel Andersen, DTU EE
Ole Ravn, DTU EE

AGV set-up

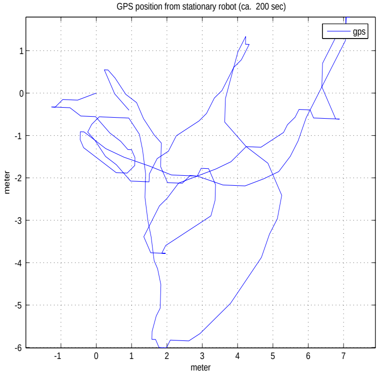


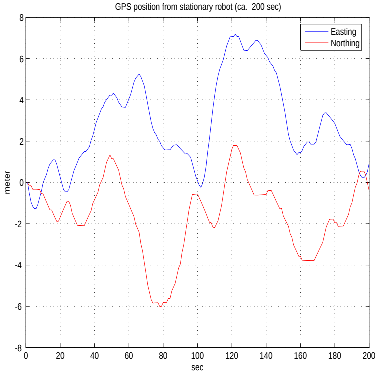
Sensors: Odemetri, GPS and vision based system.





Results from a short drive along the road from Erimitage slottet to Fortun Porten.





GPS results from a stationary robot.

Odometric dynamics

$$\frac{d}{dt_c}\begin{bmatrix} x_t \\ y_t \\ \theta_t \end{bmatrix} = \begin{bmatrix} v_t\ cos(\theta_t) \\ v_t\ sin(\theta_t) \\ \omega_t \end{bmatrix} + w_t \qquad v_r = \frac{\omega_r r_r + \omega_l r_l}{2} \qquad \omega_t = \frac{\omega_r r_r - \omega_l r_l}{b}$$

Vision based sensor

$$y_t = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ \theta_t \\ x_g \\ y_g \end{bmatrix} + e_t \qquad \begin{bmatrix} x_g \\ y_g \end{bmatrix} \in \mathbf{N}\left(m_g, P_g\right)$$

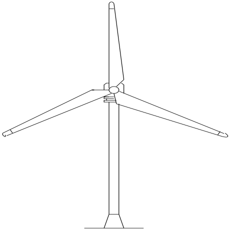
Encoder measurements

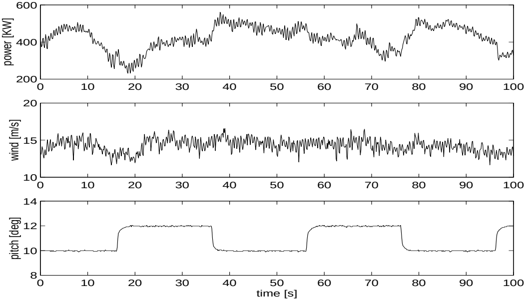
$$z_t = \begin{bmatrix} \omega_r \\ \omega_l \end{bmatrix} + \xi_t$$

GPS measurements

$$w = \begin{bmatrix} x_t \\ y_t \end{bmatrix} + \zeta_t$$

Example: Wind turbine





Stochastic Scalar Variable

Uncertainty. Random variable, Stochastic variable.

Definition (Scalar variable): A quantity ($X \in \mathbb{R}$) we can't predict precisely.

Description: probability (cdf [distribution function], pdf [density function]). Degree of belief.

$$F(x) = \mathbf{P}\Big\{X \le x\Big\} \qquad \mathbf{P}\Big\{a < X \le b\Big\} = F(b) - F(a) \qquad f(x) = \frac{\partial}{\partial x} F(x) \ge 0$$

Moments and expectation.

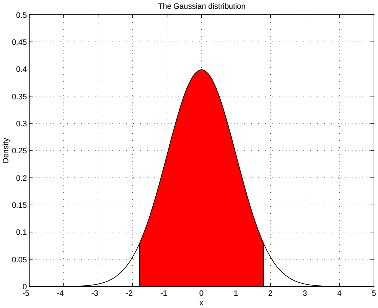
$$\mathbf{E}\Big\{g(X)\Big\} = \int_{\Omega_x} g(x)\, f(x) dx \qquad \mathbf{E}\Big\{a g_1(X) + g_2(X)\Big\} = a\mathbf{E}\Big\{g_1(X)\Big\} + \mathbf{E}\Big\{g_2(X)\Big\}$$

Description: Mean and variance.

$$m_x = \mathbf{E}\Big\{X\Big\} \qquad V_x = \mathbf{E}\Big\{(X - m_x)^2\Big\}$$

Description: Confidence interval

$$\mathbf{P}\Big\{a < X \leq b\Big\} = 0.95$$



2.5 % in each tails.

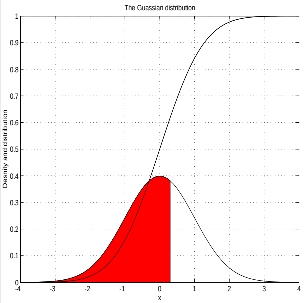
Carl Friedrich Gauss

- Born: 30 April 1777 Braunschweig, Holy Roman Empire
- Died: 23 February 1855 (aged 77) Gøttingen, Kingdom of Hanover.
- Institutions: University of Gøttingen
- Alma mater: University of Helmstedt

The Standard Gaussian distribution (scalar)

$$X \in \mathbf{N}(0,1) \qquad f(x) = \frac{1}{\sqrt{2\pi}}\ exp\left(-\frac{1}{2}x^2\right)$$

Distribution (cdf and pdf)



p-values

p	f_p
0.0005	-3.2905
0.001	-3.0902
0.005	-2.5758
0.01	-2.3263
0.05	-1.6449
0.5	0
0.95	1.6449
0.99	2.3263
0.995	2.5758
0.999	3.0902
0.9995	3.2905

Matlab: **normpdf, normcdf** and **norminv**.

The Gaussian distribution (scalar)

$$X \in \mathbf{N}(0,1) \qquad Y = aX + b \in \mathbf{N}\left(b, a^2\right)$$

$$Y \in N\left(m, \sigma^2\right) \qquad \frac{Y-m}{\sigma} \in N(0,1) \qquad Y = m + \sigma X$$

$$\begin{aligned} P\{Y \leq \#\} &= P\left\{\frac{Y-m}{\sigma} \leq \frac{\#-m}{\sigma}\right\} \\ &= P\left\{X \leq \frac{\#-m}{\sigma}\right\} \quad \text{from table} \end{aligned}$$

Assume:

$$X \in N\left(m_x, \sigma_x^2\right), \qquad V \in N\left(m_v, \sigma_v^2\right), \qquad \mathbf{C}ov\left\{X, V\right\} = r$$

Then (linear scalar operations):

$$X + m \in N\left(m_x + m, \sigma_x^2\right)$$

$$aX \in N\left(am_x, a^2\sigma_x^2\right)$$

$$X + V \in N\left(m_x + m_v, \sigma_x^2 + \sigma_v^2 + 2r\right)$$

The Gamma function

Definition and properties

For $x \in \mathbb{R}_+$:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad \text{for} \quad x > 0$$

$$\Gamma(x+1) = x\Gamma(x)$$

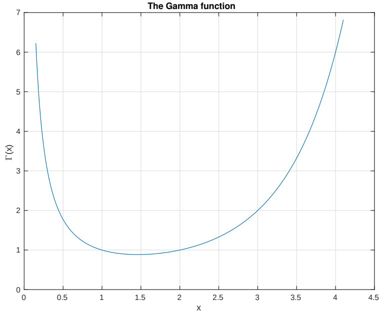
$$\Gamma(1) = 1 \qquad \Gamma(\frac{1}{2}) = \sqrt{\pi}$$

For $n \in \mathbb{N}_+$:

$$\Gamma(n) = (n-1)!$$

$$\Gamma(n + \frac{1}{2}) = \frac{(2n-1)!}{2^n}\sqrt{\pi}$$

Plot



The χ^2−distribution

Properties

$$X_i \in \mathbf{N}_{iid}\,(0,1) \qquad i = 1, ..., n$$

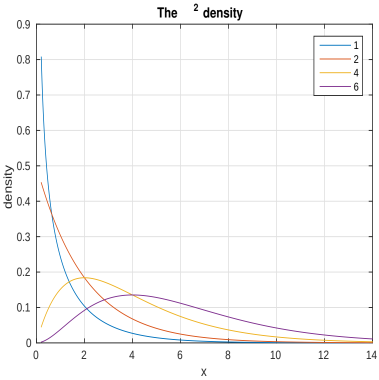
$$Y = \sum_{i=1}^{n} X_i^2 \in \chi^2(n)$$

$$f(y) = \frac{1}{\Gamma(\frac{1}{2})} \; \frac{1}{\sqrt{2y}} \; exp(-\frac{1}{2}y)$$

$$\mathbf{E}\Big\{y\Big\} = n \qquad \mathbf{V}ar\Big\{y\Big\} = 2n$$

$$X_i \in \mathbf{N}_{iid}\,(0,\sigma_i^2) \qquad \sum_{i=1}^{n} \frac{X_i^2}{\sigma_i^2} \in \chi^2(n)$$

Plot



The F-distribution

$$F(n,m) = \frac{\chi^2(n)}{\chi^2(m)} \frac{m}{n}$$

The Rayleigh distribution

$$X_i \in \mathbf{N}_{iid}\left(0,\sigma^2\right) \quad i = 1,2$$

$$R = \sqrt{X_1^2 + X_2^2} \in Ray(\sigma^2)$$

$$\theta = \arctan(\frac{X_2}{X_1}) \in U(0,2\pi)$$

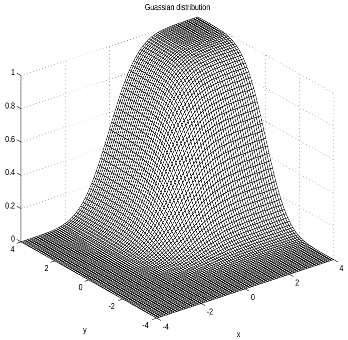
The student T distribution

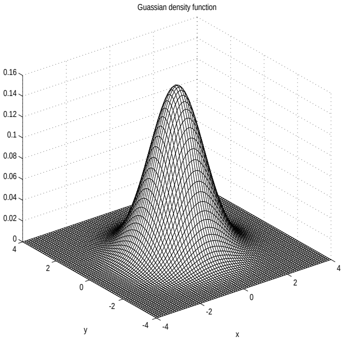
Marginal description

$$F_x(x) = P\{X \leq x\} \qquad F_y(y) = P\{Y \leq y\}$$

Simultaneous description

$$F_{xy}(x,y) = P\{X \leq x, Y \leq y\} \qquad f_{xy}(x,y) = \frac{d^2}{dxdy} F_{xy}(x,y)$$





$$F_x(x) = F_{xy}(x, \infty) \qquad f_x(x) = \int_{\Omega_y} f_{xy}(x,y)\, dy$$

Independency

Janis Joplin (Kris Kristofferson, Fred Foster): Freedom is just another word for nothing left to lose.

$$F(x,y) = F(x)F(y) \qquad\qquad f(x,y) = f(x)f(y)$$

Dependency, covariance, correlation:

$$Y = 3X + E \qquad\qquad E \perp X$$

Covariance:

$$\mathbf{Cov}\{X,Y\} = \mathbf{E}\Big\{(X-m_x)(Y-m_y)\Big\} \underset{Schwartz}{\leq} \sqrt{Var\{X\}\,Var\{Y\}}$$

Correlation (coefficient):

$$\mathbf{Cov}\{X,Y\} = \rho\,\sqrt{Var\{X\}\,Var\{Y\}} \qquad \textit{or} \qquad \rho = \frac{\mathbf{Cov}\{X,Y\}}{\sqrt{Var\{X\}\,Var\{Y\}}}$$

$$-1 \leq \rho \leq 1$$

- $\rho = 0$ uncorrelated
- $\rho = \pm 1$ perfect correlated ($Y = aX$ for some a).
- Independency $\Rightarrow\ \rho = 0$

DTU

Covariance and model

Assume:

$$X \in \mathbf{N}\left(m_x, \sigma_x^2\right)$$

$$Y \in \mathbf{N}\left(m_y, \sigma_y^2\right)$$

$$\mathbf{Cov}\left\{Y, X\right\} = r$$

Then (model)

$$Y = aX + E$$

where

$$E \in \mathbf{N}\left(m_e, \sigma_e^2\right) \qquad \mathbf{E} \perp \mathbf{X}$$

Analysis

Analysis of model results in:

$$m_y = am_x + m_e$$

$$\sigma_y^2 = a^2\sigma_x^2 + \sigma_e^2$$

$$r = a\sigma_x^2$$

Consequently:

$$a = \frac{r}{\sigma_x^2}$$

$$\sigma_e^2 = \sigma_y^2 - a^2\sigma_x^2 \qquad \geq 0$$

$$m_e = m_y - am_x$$

Stochastic vectors

Definition: $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$

Description:

- $F(x_i)$ marginal distributions (only partial information)
- $F(\mathbf{x}) = P\{X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n\}$ Simultanuos distribution.
- $$f(\mathbf{x}) = \frac{d^n}{dx_1 dx_2 ... dx_n} F(\mathbf{x})$$
- If independent elements in $\mathbf{X}$ then it is easy to determine

$$F(\mathbf{x}) = \prod_{i=1}^{n} F(x_i) \qquad f(\mathbf{x}) = \prod_{i=1}^{n} f(x_i)$$

Moments

$$\mathbf{E}\Big\{g(\mathbf{X})\Big\} = \int_{\Omega_x} g(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$$

$$\mathbf{E}\Big\{A g_1(\mathbf{X}) + g_2(\mathbf{X})\Big\} = A\mathbf{E}\Big\{g_1(\mathbf{X})\Big\} + \mathbf{E}\Big\{g_2(\mathbf{X})\Big\}$$

Mean

$$\mathbf{m_x} = \mathbf{E}\Big\{\mathbf{X}\Big\} = \begin{bmatrix} \mathbf{E}\Big\{X_1\Big\} \\ \mathbf{E}\Big\{X_2\Big\} \end{bmatrix} \qquad \text{(for } n = 2)$$

Variance

$$\mathbf{V}ar\Big\{\mathbf{X}\Big\} = \mathbf{E}\Big\{\,(\mathbf{X} - \mathbf{m}_x)\,(\mathbf{X} - \mathbf{m}_x)^\top\,\Big\}$$

Symmetric, Positive semi definite

Assume $\mathbf{m} = 0$ and $n = 2$.

$$\begin{aligned} \mathbf{V}ar\Big\{\mathbf{X}\Big\} &= \mathbf{E}\Big\{\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}[X_1\ X_2]\Big\} \\ &= \mathbf{E}\Big\{\begin{bmatrix} X_1X_1 & X_1X_2 \\ X_2X_1 & X_2X_2 \end{bmatrix}\Big\} \\ &= \begin{bmatrix} Var\{X_1\} & Cov\Big\{X_1, X_2\Big\} \\ Cov\Big\{X_1, X_2\Big\} & Var\{X_2\} \end{bmatrix} \end{aligned}$$

In general:

$$\mathbf{V}ar\Big\{\mathbf{X}\Big\} = Matr\,\Big\{Cov\Big\{X_r, X_c\Big\}\Big\}$$

Covariance II

$$X, V \in \mathbb{R}^n \qquad Y \in \mathbb{R}^p$$

$$\mathbf{Cov}\{X, Y\} = \mathbf{E}\Big\{ (X - m_x)(Y - m_y)^\top \Big\} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1,p} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_{np} \end{bmatrix} \qquad \sigma_{rc} = Cov\Big\{X_r, Y_c\Big\}$$

$$\mathbf{Cov}\{X, X\} = \mathbf{V}ar\Big\{X\Big\}$$

$$\mathbf{Cov}\{Y, X\} = \mathbf{Cov}\{X, Y\}^\top$$

$$\mathbf{Cov}\{AX, Y\} = A\mathbf{Cov}\{X, Y\}$$

$$\mathbf{Cov}\{X, AY\} = \mathbf{Cov}\{X, Y\} A^\top$$

$$\mathbf{Cov}\{X + V, Y\} = \mathbf{Cov}\{X, Y\} + \mathbf{Cov}\{V, Y\}$$

DTU

$$\mathbf{X} \in \mathbb{F}(m_x, P_x) \qquad \mathbf{V} \in \mathbb{F}(m_v, P_v) \qquad R_{vx} = \mathbf{Cov}\{V, X\}$$

Notation: $\mathbb{F}(m_x, P_x)$ means second order stoch. variable with mean m_x and variance P_x.

$$\mathbf{X} + m \in \mathbb{F}(m_x + m, P_x)$$

$$A\mathbf{X} \in \mathbb{F}\left(Am_x, AP_xA^T\right)$$

$$\mathbf{X} + \mathbf{V} \in \mathbb{F}\left(m_x + m_v, P_x + P_v + R_{vx} + R_{vx}^T\right)$$

Calculus

$$\mathbf{E}\left\{\mathbf{X}\mathbf{X}^\top\right\} = P_x + m_x m_x^\top$$

$$\begin{aligned} P_x &= \mathbf{E}\left\{(X - m_x)(X - m_x)^\top\right\} \\ &= \mathbf{E}\left\{XX^\top - Xm_x^\top - m_x X^\top + m_x m_x^\top\right\} \\ &= \mathbf{E}\left\{XX^\top\right\} - m_x m_x^\top \end{aligned}$$

$$\mathbf{E}\left\{X^\top SX\right\} = tr\left(SP_x\right) + m_x^\top S m_x$$

$$X \in \mathbb{F}(m_x, P_x) \qquad V \in \mathbb{F}(m_v, P_v)$$

$$\textbf{Cov}\{V, X\} = R$$

Y	$\mathbf{E}\{Y\}$	$Var\{Y\}$	$\textbf{Cov}\{Y, X\}$
AX	Am_x	AP_xA^T	AP_x
$X + m$	$m_x + m$	P_x	P_x
$X + V$	$m_x + m_v$	$P_x + P_v + R + R^T$	$P_x + R$

Example

$$Y = AX + V$$

$$X \in \mathbb{F}(m_x, P_x) \qquad V \in \mathbb{F}(0, R_1) \qquad \textbf{Cov}\{X, V\} = 0$$

assumptions

$$Y \in \mathbb{F}\left(Am_x, AP_xA^\top + R_1\right)$$

$$\textbf{Cov}\{Y, X\} = AP_x$$

Covariance and model

Assume:

$$X \in \mathbf{N}(m_x, P_x)$$

$$Y \in \mathbf{N}(m_y, P_y)$$

$$\mathbf{Cov}\{Y, X\} = R$$

Then (model)

$$Y = AX + E$$

where

$$E \in \mathbf{N}(m_e, P_e) \quad \text{and} \quad \mathbf{E} \perp \mathbf{X}$$

Analysis

Analysis of model results in:

$$m_y = Am_x + m_e$$

$$P_y = AP_xA^T + P_e$$

$$R = AP_x$$

Consequently:

$$A = RP_x^{-1}$$

$$P_e = P_y - AP_xA^T \quad \geq 0$$

$$m_e = m_y - Am_x$$

An example

A: a cow is sick (1) or healthy (0)
B: a test is positive (1) or negative (0)

Based on the test we would like to state that the cow is healthy or sick, but can only give $P\{A|B\}$

A test and its quality:

$P\{A,B\}$	$A=0$	$A=1$	$P\{B\}$
$B=0$	89	1	90
$B=1$	2	8	10
$P\{A\}$	91	9	100

All numbers in %

$$P\{A|B\} = \frac{P\{A,B\}}{P\{B\}}$$

Bayes Theorem

$P\{A|B\}$	$A=0$	$A=1$
$B=0$	99	1
$B=1$	20	80

All numbers in percent and rounded.

Thomas Bayes (c. 1702 - 17 April 1761) was a British mathematician and Presbyterian minister, known for having formulated a specific case of the theorem that bears his name: Bayes' theorem, which was published posthumously.

Born: c.1702 London, England.
Died: 17 April 1761 (aged 59). Tunbridge Wells, Kent, England

Dependency

$$f(x|y) = \frac{f(x,y)}{f(y)}$$

$$f(x,y) = f(x|y)f(y) = f(y|x)f(x)$$

$$f(x) = \int f(x,y)dy = \int f(x|y)f(y)dy$$

Independence

$$F(x,y) = F(x)F(y) \qquad f(x,y) = f(x)f(y) \qquad f(x|y) = f(x)$$

Mean

$$m_x = \mathbf{E}\Big\{X\Big\} = \int_{\Omega_x} x f(x)dx \qquad m_{x|y} = \mathbf{E}\Big\{X|Y\Big\} = \int_{\Omega_x} x f(x|y)dx$$

Variance

$$\mathbf{V}ar\Big\{X\Big\} = \mathbf{E}\Big\{(X - m_x)^2\Big\} \qquad \mathbf{V}ar\Big\{X|Y\Big\} = \mathbf{E}\Big\{(X - m_{x|y})^2|Y\Big\}$$

$$= \int_{\Omega_x} (x - m_{x|y})^2 f(x|y)dx$$

DTU

Vector Gaussian distribution

Definition:

$$X \in N_n \qquad \forall a: \; a^\top X \in N$$

$$X \in N(m_x, P_x)$$

$$Y = AX + m \in N(Am_x + m, AP_xA^\top)$$

$$X \in N(m_x, P_x)$$

$$X = AZ + m_x \qquad Z \in N(0, I) \qquad P_x = AA^\top$$

$$f_X(x) = \frac{1}{\sqrt{Det(P_x)}\left(\sqrt{2\pi}\right)^n} \times exp\left(-\frac{1}{2}\left[x - m_x\right]^\top P_x^{-1}\left[x - m_x\right]\right)$$

$$Z_i: \quad f(z_i) = \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2} z_i^2)$$

$$f(z) = f(z_1)f(z_2)...f(z_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n exp\left(-\frac{1}{2}\sum_{i=1}^{n} z_i^2\right) \qquad \sum_{i=1}^{n} z_i^2 = z^T z$$

$Z: \; f_z(z)$ $\qquad$ $X = g(Z)$ $\qquad$ g and g^{-1} do exists and are C^1

$$f_X(x) = f_z\left[g^{-1}(x)\right] \times Det\left(\frac{\partial g^{-1}(x)}{\partial x}\right)$$

$X = AZ + m \qquad P_x = AA^\top > 0$

$Z = A^{-1}\left[X - m_x\right]$ $\qquad$ A^{-1} do exists

$$\begin{aligned} z^\top z &= (x - m_x)^\top A^{-\top} A^{-1} (x - m_x) \\ &= (x - m_x)^\top P_x^{-1} (x - m_x) \end{aligned}$$

$$Det\left(\frac{\partial}{\partial x}\left[A^{-1}(X - m_x)\right]\right) = Det\left[A^{-1}\right] = \frac{1}{\sqrt{Det P_x}}$$

Confidence areas

Def.: smallest areas with a given probability.

$$f_X(x) = \frac{1}{\sqrt{Det(P_x)}\left(\sqrt{2\pi}\right)^n} \times \ exp\left(-\frac{1}{2}\left[x - m_x\right]^\top P_x^{-1}\left[x - m_x\right]\right) > level$$

or

$$g(x) = \left[x - m_x\right]^\top P_x^{-1}\left[x - m_x\right] < C \qquad = -2\log\left[\ level\ \sqrt{Det(P_x)}\left(\sqrt{2\pi}\right)^n\right]$$

By the way:

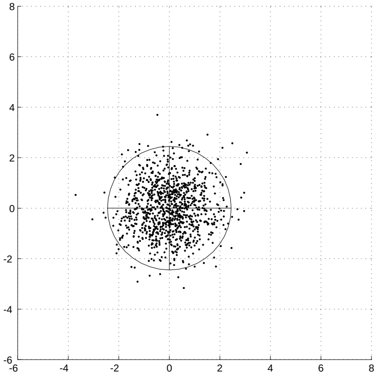
$$\left[X - m_x\right]^\top P_x^{-1}\left[X - m_x\right] \in \chi^2(n) \qquad C = \chi^2_{1-\alpha}(n)$$

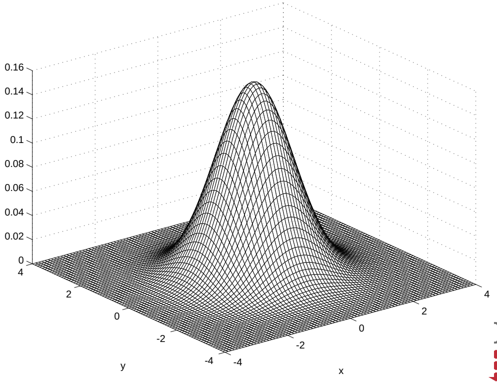
Example a

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

$$g(x) = [x - m_x]^\top P_x^{-1} [x - m_x] = [x_1^2 + x_2^2] \leq r^2$$

$$Z = [X_1^2 + X_2^2] \in \chi^2(2) \qquad \chi^2(2)_{0.95} = 5.99$$





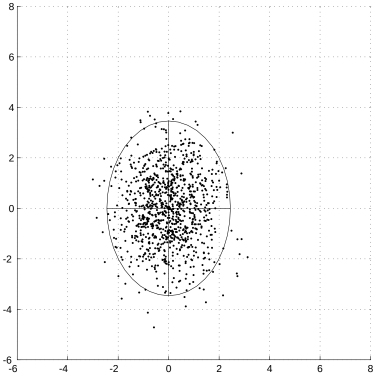
Example b

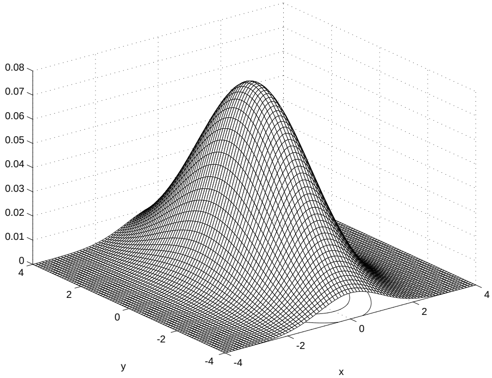
$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \in N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right)$$

Confidence area:

$$g(x) = [x - m_x]^\top P_x^{-1} [x - m_x] = \frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} \leq f_{1-\alpha}^{\chi^2}$$

Axis: $\sigma_i \sqrt{f_{1-\alpha}^{\chi^2(2)}}$





Example c:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \in N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.28 \\ 0.28 & 2 \end{bmatrix}\right)$$

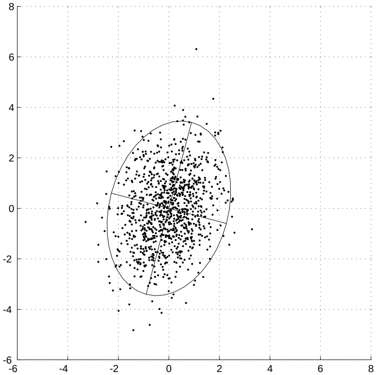
$$PV = VD \qquad V^T V = VV^T = I$$

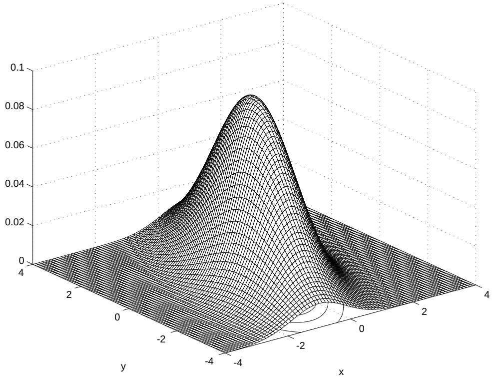
$$P = VDV^\top$$

$$V = \begin{bmatrix} -0.9671 & 0.2546 \\ 0.2546 & 0.9671 \end{bmatrix} \qquad D = \begin{bmatrix} 0.9255 & 0 \\ 0 & 2.0745 \end{bmatrix}$$

$$Z = V^\top X \in N(0, V^\top PV) = N(0, D)$$

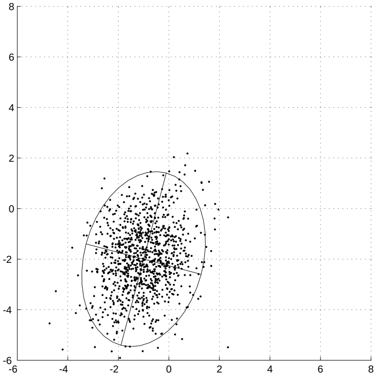
Change of basis $X = VZ$

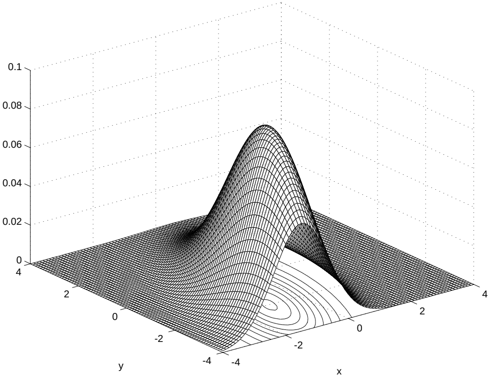




Example d:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \in N\left(\begin{bmatrix} -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 & 0.28 \\ 0.28 & 2 \end{bmatrix}\right)$$





Lesson learned in L4

- Stochastic variable
- Two scalar stochastic variable
- Stochastic vectors
- Vector Gaussian distribution

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

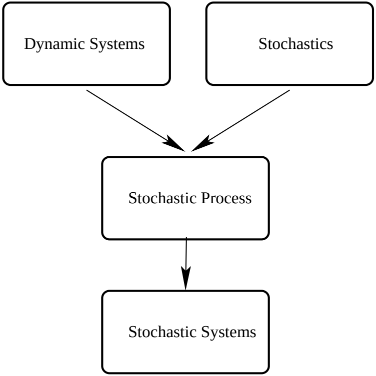
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

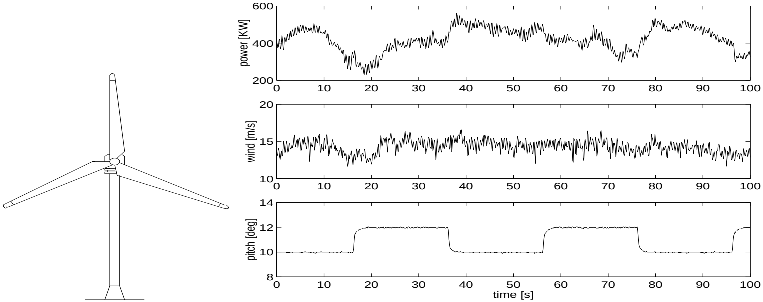
Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Stochastic Processes I - L5

The local view (ie. the horizon until Stochastic Systems).



Example: Wind turbine



This is not a stochastic process. It is a data sequence and a realization or an outcome of a stochastic process.

Definition

1. One stochastic variable X described by $F(x)$, $f(x)$, m, P. If mathematically correct $X(\omega)$
2. Two stochastic variables X, Y described by $F(x)$, $F(y)$, $F(x,y)$, ... , $\mathbf{C}ov\Big\{X,Y\Big\}$
3. Three stochastic variables X, Y and Z by $F(x,y,z)$
4. A stochastic vector $\mathbf{X}$ describes by $F(\mathbf{x})$.
5. A stochastic process (in D-time) is a sequence of stochastic variable and can (for a finite process) be represented as a vector.

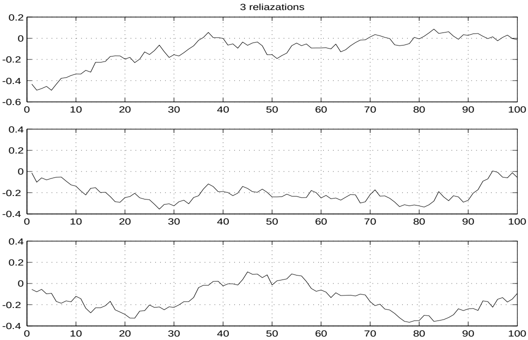
DTU

Definition

Discrete time: Stochastic process is a sequence of stochastic variables

$$\{X_t(\omega),\ t \in T,\ \omega \in \Omega\}$$

If ω is fixed we have a time function or a realization.
If t is fixed then we have a stochastic variable.



We need a lot of realisation if we will estimate properties of $X_t(\omega)$ unless the process has a nice property such as stationarity or/and ergodicity.

Definition

Continuous time: Stochastic process is a family of stochastic variables

$$\{X_t(\omega),\ t \in T,\ \omega \in \Omega\}$$

where the index set is continuous.

Description

How to describe Stochastic Processes. Analysis and model building (synthesis).

Distribution Consider a (Finite and in D-time) process: $X_i \in \mathbb{R}, \quad i = 1, \ldots, k$

$$F(x_1, x_2, \ldots, x_k)$$

$$f(x_1, x_2, \ldots, x_k)$$

Moments

$$M = \mathbf{E}\left\{\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}\right\} \quad k \times 1 \qquad \Sigma = \mathbf{V}ar\left\{\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}\right\} \quad k \times k$$

What if $X_i \in \mathbb{R}^n, \quad i = 1, \ldots, k$

Infinite processes (D and C-time):

$$T_{sub} = \{t_1, t_2, \ldots, t_k\} \qquad \text{(for any subset)}$$

$$F(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$$

$$f(x_{t_1}, x_{t_2}, \ldots, x_{t_k})$$

Moments

$$m_x(t) = \mathbf{E}\Big\{X_t\Big\}$$

$$P_x(t) = \mathbf{E}\Big\{\tilde{X}_t\tilde{X}_t^\top\Big\} \qquad \tilde{X}_t = X_t - m_x(t)$$

$$R_x(s,t) = \mathbf{E}\Big\{\tilde{X}_s\tilde{X}_t^\top\Big\}$$

Description

Dynamic function of white noise $\{e_t, v_t\}$.

State space (A, R_1, C and R_2).

$$x_{t+1} = Ax_t + v_t \qquad v_t \in \mathbb{F}(0, R_1)$$
$$y_t = Cx_t + e_t \qquad e_t \in \mathbb{F}(0, R_2)$$

Transfer function ($A_p(q^{-1})$, $C_p(q^{-1})$ and σ^2).

$$y_t = \frac{C_p(q^{-1})}{A_p(q^{-1})} e_t \qquad e_t \in \mathbb{F}\left(0, \sigma^2\right)$$

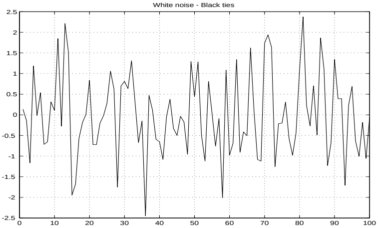
White noise

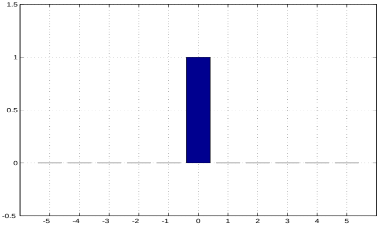
A sequence of stochastic independent stochastic variables, as eg.:

$$e_t \in \mathbf{N}_{iid}\left(\mu, \sigma^2\right)$$

has the property:

$$R(s,t) = 0 \text{ for } s \neq t$$





Can not be realized in continuous time - just a mathematical abstraction.

A very simple process (Finite discrete, two variable)

$i = 0,\ 1.$

$$x_1 = Ax_0 + v_0 \qquad x_0 \in \mathbf{N}(m_0, P_0) \qquad v_0 \in \mathbf{N}(0, R_1) \qquad v_0 \perp x_0$$

$$x_0 \in \mathbf{N}(m_0, P_0)$$

$$x_1 \in \mathbf{N}\left(Am_0, AP_0A^T + R_1\right) \qquad \mathbf{Cov}\{x_1, x_0\} = AP_0$$

$$\begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \in \mathbf{N}\left(\begin{bmatrix} m_0 \\ Am_0 \end{bmatrix}, \begin{bmatrix} P_0 & P_0A^T \\ AP_0 & AP_0A^T + R_1 \end{bmatrix}\right)$$

DTU

Yet another simple process

$i = 0,\ 1,\ 2.$

$$x_{i+1} = Ax_i + v_i \qquad x_0 \in \mathbf{N}(m_0, P_0) \qquad v_i \in \mathbf{N}_{iid}(0, R_1) \qquad v_i \perp x_0$$

$$x_i \in \mathbf{N}(m_i, P_i) \qquad m_1 = Am_0 \qquad m_2 = Am_1$$

$$P_1 = AP_0A^T + R_1 \qquad P_2 = AP_1A^T + R_1$$

$$R_{10} = \mathbf{Cov}\{x_1, x_0\} = AP_0 \qquad R_{21} = AP_1 \qquad R_{20} = A^2P_0$$

$$\begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} \in \mathbf{N}\left(\begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix}, \begin{bmatrix} P_0 & R_{10}^T & R_{20}^T \\ R_{10} & P_1 & R_{21}^T \\ R_{20} & R_{21} & P_2 \end{bmatrix} \right)$$

Consider a model:

$$x_{t+1} = Ax_t + v_t \qquad v_t \in \mathbb{F}(0, R_1) \qquad x_{t_0} \in \mathbb{F}(m_0, P_0)$$

where:

$\mathbf{Cov}\{v_t, v_s\} = 0$ $\qquad$ $\mathbf{Cov}\{v_t, x_s\} = 0$ *for* $s \leq t$

$\mathbb{F} = N \rightarrow \{x_t\}$ a Gaussian process

Then:

$x_t \in \mathbb{F}(m_t, P_t)$

where:

$$m_{t+1} = Am_t \qquad m_{t_0} = m_0$$

$$P_{t+1} = AP_tA^\top + R_1 \qquad P_{t_0} = P_0$$

$$R(\tau, t) = A^{\tau-t}P_t \qquad \tau \geq t$$

LTI process model

The proof:

$$x_\tau = A^{\tau-t} x_t + \mathcal{W}_c V_{t:\tau-1} \qquad\qquad \tau \geq t$$

$$x_\tau = A^{\tau-t} x_t + \left[A^{\tau-t-1} \; ... \; I \right] \begin{bmatrix} v_t \\ \vdots \\ v_{\tau-1} \end{bmatrix}$$

$$R(\tau, t) = \mathbf{E}\left\{ x_\tau x_t^T \right\} = A^{\tau-t} P_t \qquad\qquad \tau \geq t$$

Recap results

Consider a model:

$$x_{t+1} = Ax_t + v_t \qquad v_t \perp x_s \quad s \le t$$

$$v_t \in \mathbb{F}(0, R_1) \qquad x_{t_0} \in \mathbb{F}(m_0, P_0)$$

$$v_t \perp vs \text{ for} t \neq s$$

$$m_{t+1} = Am_t$$

$$P_{t+1} = AP_tA^T + R_1$$

$$R_x(k,t) = A^{k-t}P_t \qquad k > t$$

Finite sample

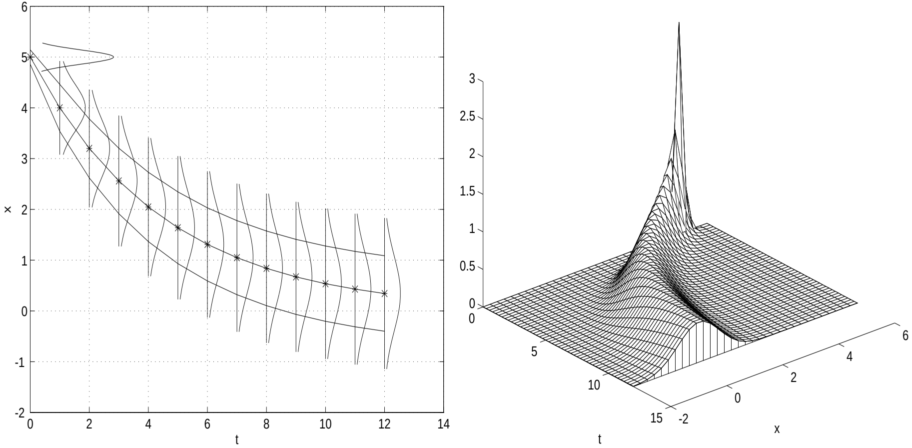
$$m = \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_t \end{bmatrix} \qquad X = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_t \end{bmatrix}$$

$$X \in \mathbf{N}(m, \Sigma)$$

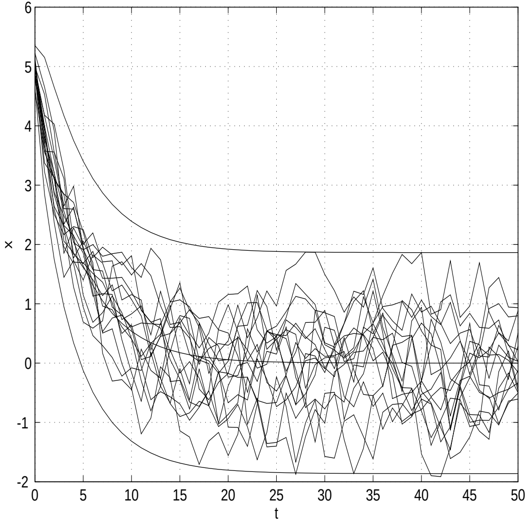
$$\Sigma = \begin{bmatrix} P_0 & \times & & & & \\ R_x(1,0) & P_1 & & & & \\ \vdots & \vdots & \ddots & & & \\ \dots & R_x(r-1,c-1) & \dots & P_{c-1} & & \\ & & & & \ddots & \\ & & & & & P_t \end{bmatrix}$$

Example I

$$x_{t+1} = 0.98 x_t + v_t \quad v_t \in \mathbf{N}_{iid}(0, 0.2) \quad x_{t_0} \in \mathbf{N}(5, 0.02)$$



Example I

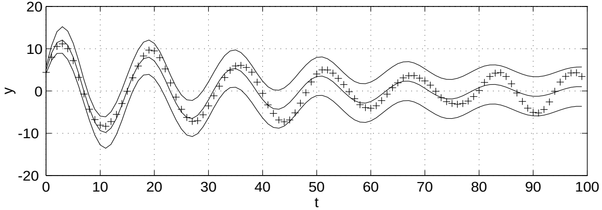


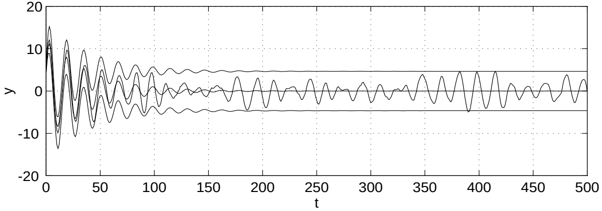
Example II

$$x_{t+1} = \begin{bmatrix} 1.8 & 1 \\ -0.95 & 0 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} v_t \qquad x_t \text{ has two elements}$$

$$x_0 \in N\left(\begin{bmatrix} 5 \\ 0 \end{bmatrix}; \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}\right) \text{ and } v_t \in \mathbf{N}_{iid}(0, 0.05)$$

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} x_t$$





Process output

Now, consider the process output (measure, controlled):

$$y_t = Cx_t + e_t \qquad\qquad y_t = Cx_t$$

where:

$$e_t \in \mathbb{F}(0, R_2) \qquad x_t \in \mathbb{F}(m_x(t), P_x(t))$$

$$\mathbf{Cov}\{e_t, e_s\} = 0 \qquad \mathbf{Cov}\{e_t, x_s\} = 0 \qquad \text{for } s \le t$$

then:

$$y_t \in \mathbb{F}(m_y(t), P_y(t))$$

$$m_y(t) = Cm_x(t)$$

$$P_y(t) = CP_x(t)C^\top + R_2 \qquad\qquad P_y(t) = CP_x(t)C^\top$$

$$R_y(\tau, t) = CA^{\tau-t}P_x(t)C^\top \qquad \tau > t$$

$\mathbb{F} = \mathbf{N} \rightarrow \{y_t\}$ a Gaussian process

DTU

Stationarity - I

For the LTI-process (with standard assumptions):

$$m_{t+1} = Am_t \qquad m_{t_0} = m_0$$

$$P_{t+1} = AP_tA^\top + R_1 \qquad P_{t_0} = P_0$$

$$R(\tau, t) = A^{\tau-t}P_t \qquad \tau \geq t$$

---

which for $\forall|\lambda(A)| < 1$ results in:

$$m_t \to 0 \qquad \text{for } t_0 \to -\infty$$

$$P_t \to P_\infty \geq 0 \qquad \text{for } t_0 \to -\infty$$

$$R(\tau - t) = A^{\tau-t}P_\infty \qquad \tau \geq t \qquad \text{for } t_0 \to -\infty$$

The stationary variance can be found as a solution to the (Discrete) **Lyapunov equation**

$$P_\infty = AP_\infty A^\top + R_1$$

In Matlab solved by dlyap . What happens if $\exists|\lambda(A)| > 1$

For the process output we have (for $t_0 \to -\infty$ and $\forall \lambda\,(A) < 1$):

$$m_y(t) = Cm_x(t) \to 0$$

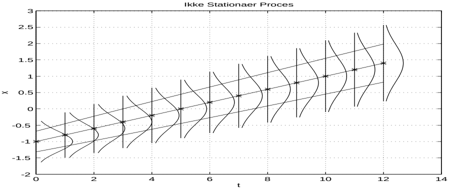
$$P_y(t) \to CP_\infty C^\top + R_2 \qquad\qquad P_y(t) \to CP_\infty C^\top$$

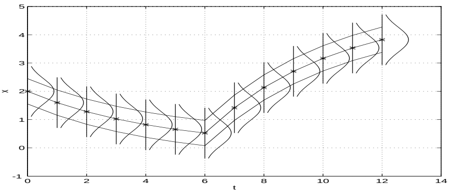
$$R_y(\tau - t) = CA^{\tau - t}P_\infty \qquad \tau \geq t$$

Stationarity - II

Definition: **The statistical properties of the process are time invariant.**

NOT stationary processes:





DTU

Stationarity

Order 1

$$f(x_t) = f(x_{t+\tau})$$

$$m_x(t) = m_x \qquad\qquad P_x(t) = P_x$$

Order 2

$$f(x_t, x_s) = f(x_{t+\tau}, x_{s+\tau})$$

$$r(s,t) = r(s-t)$$

Order n

$$f(x_{t_1}, ..., x_{t_k}) = f(x_{t_1+\tau}, ..., x_{t_k+\tau})$$

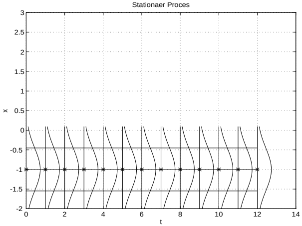
Order $n \Rightarrow$ Order $m < n$
Strong stationary iff strong stationary for all n.

Weak (or wide sense) stationarity
Second order process and:

$$m_x(t) = m_x \qquad P_x(t) = P_x$$

$$r(s,t) = r(s-t)$$

White noise



(in discrete time)

$$e_t \in \mathbf{N}_{iid}\left(\mu, \sigma^2\right)$$

is stationary (if μ and σ^2 are constants).
It is also a Markov process.

$$\dot{x}_t = Ax_t + v_t \qquad x_{t_0} \in N(m_0, P_0)$$

$$dx_t = Ax_t dt + dv_t$$

$$dv_t \in N(0, R_1 dt)$$

$$\Delta x_t = Ax_t \Delta t + \Delta v_t$$

$$\Delta v_t \in N(0, R_1 \Delta t)$$

LTI process

$$\dot{m}_t = Am_t \qquad m_{t_0} = m_0$$

$$\dot{P}_t = AP_t + P_t A^\top + R_1 \qquad P_{t_0} = P_0$$

$$R_x(s,t) = e^{A(s-t)} P_t$$

If asymptotic stable

$$x_t \in N(0, P_\infty)$$

C-time Lyapunov equation.

$$AP_\infty + P_\infty A^\top + R_1 = 0$$

Wind speed model

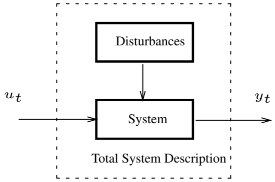
$$y = H(\frac{d}{dt})e_w \qquad H(s) = \frac{k}{(1+sp_1)(1+sp_2)}$$

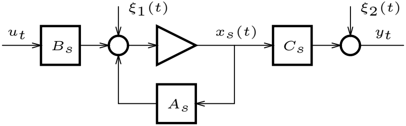
$$x_1 = v \qquad x_2 = \dot{v}$$

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{p_1 p_2} & -\frac{p_1+p_2}{p_1 p_2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k}{p_1 p_2} \end{bmatrix} e_w$$

$$v = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Systems and disturbances





$$x_s(t+1) = A_s x_s(t) + B_s u(t) + \xi_1(t)$$

$$y(t) = C_s x_s(t) + \xi_2(t)$$

$$x_s(t+1) = A_s x_s(t) + B_s u(t) + \xi_1(t)$$

$$y(t) = C_s x_s(t) + \xi_2(t)$$

$$x_1(t+1) = A_1 x_1(t) + v_1(t)$$

$$\xi_1(t) = C_1 x_1(t) + e_1(t)$$

$$x_2(t+1) = A_2 x_2(t) + v_2(t)$$

$$\xi_2(t) = C_2 x_2(t) + e_2(t)$$

$$\begin{bmatrix} x_s \\ x_1 \\ x_2 \end{bmatrix}_{t+1} = \begin{bmatrix} A_s & C_1 & 0 \\ 0 & A_1 & 0 \\ 0 & 0 & A_2 \end{bmatrix} \begin{bmatrix} x_s \\ x_1 \\ x_2 \end{bmatrix}_t + \begin{bmatrix} B_s \\ 0 \\ 0 \end{bmatrix} u_t + \begin{bmatrix} e_1 \\ v_1 \\ v_2 \end{bmatrix}_t$$

$$y_t = \begin{pmatrix} C_s & 0 & C_2 \end{pmatrix} \begin{bmatrix} x_s \\ x_1 \\ x_2 \end{bmatrix}_t + D_s u_t + e_2(t)$$

Disturbances, set point variation, dependencies ao.

Wind turbine

$$\begin{bmatrix} \dot{\theta}_\epsilon \\ \dot{\omega}_r \\ \dot{\omega}_g \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} 0 & 1 & -\frac{1}{n_{gear}} & 0 \\ -\frac{K_s}{J_r} & \frac{1}{J_r}\frac{\partial T_w}{\partial \omega_r} & 0 & \frac{1}{J_r}\frac{\partial T_w}{\partial \beta} \\ \frac{\eta_{gear} K_s}{n_{gear} J_g} & 0 & -\frac{D_g}{J_g} & 0 \\ 0 & 0 & 0 & -\frac{1}{\tau_\beta} \end{bmatrix} \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{K_\beta}{\tau_\beta} \end{bmatrix} \beta_{ref} + \begin{bmatrix} 0 \\ \frac{1}{J_r}\frac{\partial T_w}{\partial v} \\ 0 \\ 0 \end{bmatrix} v$$

$$P_e = \begin{bmatrix} 0 & 0 & \frac{\eta(1-S)\omega_0}{n_p} D_g & 0 \end{bmatrix} \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \end{bmatrix}$$

$$\dot{x} = Ax + Bu + B_v v$$

$$y_t = Cx$$

Wind model

$$\begin{bmatrix} \dot{v} \\ \ddot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{p_1 p_2} & -\frac{p_1+p_2}{p_1 p_2} \end{bmatrix} \begin{bmatrix} v \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k}{p_1 p_2} \end{bmatrix} e_w$$

$$v = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ \dot{v} \end{bmatrix}$$

$$\dot{x}_w = A_w x_w + B_w e_w$$

$$v = C_w x_w$$

$$\begin{bmatrix} \dot{x} \\ \dot{x}_w \end{bmatrix} = \begin{bmatrix} A & B_v C_w \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x \\ x_w \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ B_w \end{bmatrix} e_w$$

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

$$\begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \\ v \\ \dot{v} \end{bmatrix}_{t+1} = [6 \times 6] \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \\ v \\ \dot{v} \end{bmatrix}_t + [6 \times 1] \, u_t + v_t$$

$$R_1 = [6 \times 6]$$

Stochastic systems

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad x_{t_0} \in \mathbb{F}(\hat{x}_0, P_0) \qquad v_t \in \mathbb{F}(0, R_1)$$

$$y_t = Cx_t + e_t \qquad e_t \in \mathbb{F}(0, R_2)$$

- **Cov** $\{v_t, v_s\} = 0$ **Cov** $\{e_t, e_s\} = 0$ for $s \neq t$.
- **Cov** $\{v_t, x_s\} = 0$ **Cov** $\{e_t, x_s\} = 0$ for $s \leq t$

$$x_t \in \mathbb{F}(\hat{x}_t, P_t) \qquad y_t \in \mathbb{F}(m_t, \Sigma_t)$$

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t \qquad \hat{x}_{t_0} = \hat{x}_0$$

$$P_{t+1} = AP_tA^\top + R_1 \qquad P_{t_0} = P_0$$

$$m_t = C\hat{x}_t$$

$$\Sigma_t = CP_tC^\top + R_2$$

$$R_x(\tau, t) = A^{\tau-t}P_t$$

$$R_y(\tau, t) = CA^{\tau-t}P_tC^\top$$

Stochastic Systems in C-time

$\dot{x} = Ax + Bu + v \qquad x_{t_0} \in \mathbb{F}(\hat{x}_0, P_0) \qquad v_t \in \mathbb{F}(0, R_1)$

$\mathbf{Cov}\{v_t, v_s\} = 0$ *for* $s \neq t \qquad \mathbf{Cov}\{v_t, x_s\} = 0$ *for* $s \leq t$

$x_t \in \mathbb{F}(\hat{x}_t, P_t)$

$\dot{\hat{x}} = A\hat{x} + Bu \qquad \hat{x}_{t_0} = \hat{x}_0$

$\dot{P} = AP + PA^T + R_1 \qquad P_{t_0} = P_0$

DTU

Sampling II

Notice the local notation: $(t_c \in \mathbb{R})$ and $(i \in \mathbb{Z})$

$$\dot{x}(t_c) = A_c x(t_c) + B_c u(t_c) + v(t_c) \qquad v(t_c) \in N_{iid}(0, \Sigma_1)$$

$$y_i = Cx(iT) + e(iT) \qquad e(iT) \in N_{iid}(0, R_2)$$

$$x_{i+1} = Ax_i + Bu_i + v_i \qquad v_i \in N_{iid}(0, R_1)$$

$$y_i = Cx_i + e_i \qquad e_i \in N_{iid}(0, R_2)$$

$$A = e^{A_c T} \qquad B = \int_0^{\top} e^{A_c s} B_c ds$$

$$R_1 = \int_0^{\top} e^{A_c s} \Sigma_1 (e^{A_c^{\top} s}) ds$$

In Matlab: *cn2dn*.

Lesson learned (in L5)

- Stochastic process (definition)
- White noise as a building block
- Evolution of mean and variance for a LTI process (analysis)
-
- Building description of stochastic systems
- Analysis of LTI systems

Stochastic Adaptive Control (02421)

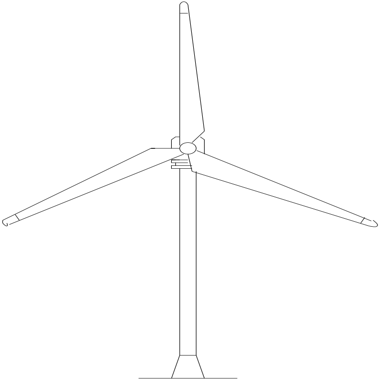
www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Filter Theory (L6-8)



L7 L8

Evolution of mean and variance

Process

$$x_{i+1} = Ax_i + v_i \qquad v_i \in \mathbf{N}(0, R_1)$$

$$x_0 \in \mathbf{N}(\underline{m}_0, \underline{P}_0)$$

Mean and variance function

$$m_{i+1} = Am_i \qquad m_0 = \underline{m}_0$$

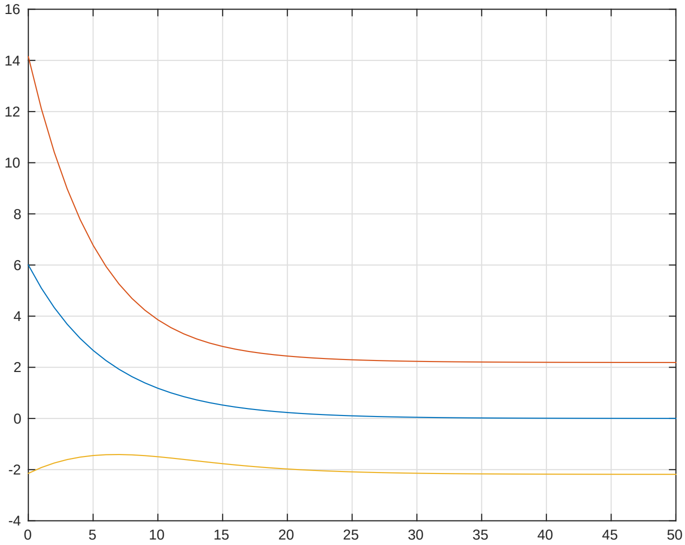
$$P_{i+1} = AP_iA^T + R_1 \qquad P_0 = \underline{P}_0$$

DTU

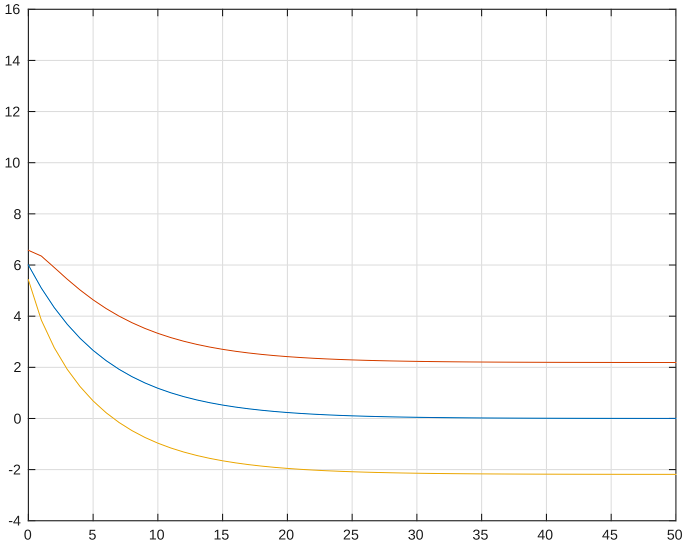
sima - ex5

```
function sima(a,r1,m0,p0,frac,nx,nstp)
%Usage: sima(a,r1,m0,p0,frac,nx,nstp)
%
%Plots the evolution of the mean and confidence
%interval (determined by the fractile frac) for state nx
%in the process
%
% x(t+1)=a*x+v      v ~ N(0,r1)   x_0 ~ N(m0,p0)
%
%The length of the plot is nstp.

t=0:nstp; t=t(:)-t(1);

p=p0; m=m0;
data=[m(nx) p(nx,nx)];
for i=1:nstp
 m=a*m;
 p=a*p*a'+r1;
 data=[data; m(nx) p(nx,nx)];
end

mu=data(:,1)+sqrt(data(:,2))*frac;
ml=data(:,1)-sqrt(data(:,2))*frac;

data=[data(:,1) mu ml];

plot(t,data); grid on;
```

sima - ex5



sima - ex5



Simulation

Process

$$x_{i+1} = Ax_i + v_i \qquad v_i \in \mathbf{N}\left(0, R_1\right)$$

$$x_0 \in \mathbf{N}\left(\underline{m}_0, \underline{P}_0\right)$$

Cholesky factorization

$$R = chol(P) \qquad R^T R = P$$

$$S = chol(P)^T \qquad SS^T = P$$

Generation

$$X \in \mathbf{N}\left(m, P\right) \qquad P = SS^T$$

$$Z \in \mathbf{N}\left(0, I\right) \qquad Z = randn(n, 1)$$

$$X = SZ + m$$

Eigenvalue decomposition

$$Pv_i = \lambda_i v_i \qquad PV = VD \qquad P = VDV^T$$

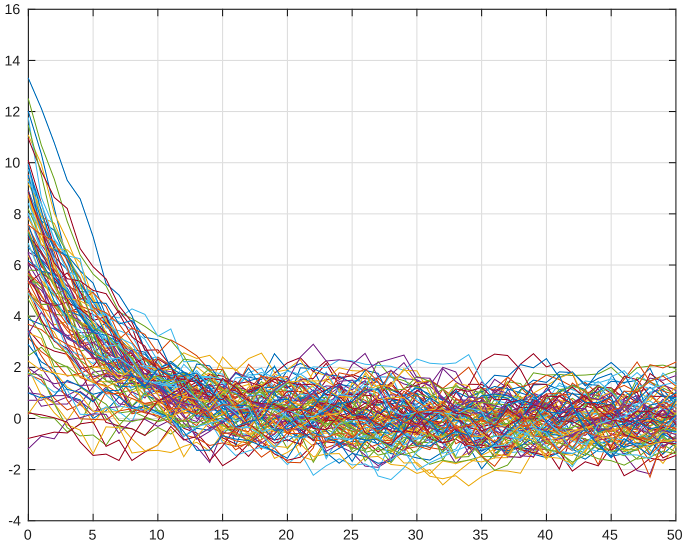
$$VV^T = V^T V = I$$

$$S = V\sqrt{D}$$

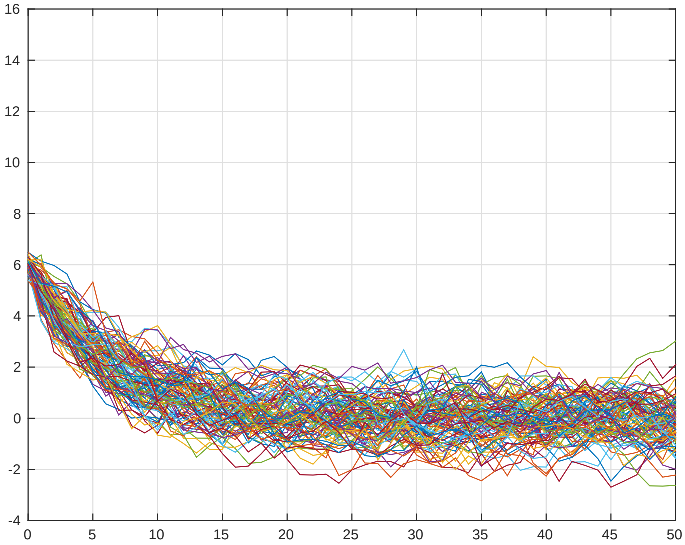
$$SS^T = V\sqrt{D}\sqrt{D}V^T = VDV^T = P$$

```
function [t,data]=simb(a,r1,m0,p0,nx,nstp,nk)
%Usage [t,data]=simb(a,r1,m0,p0,nx,nstp,nk)
%
%Simulation of nk realizations of
%
% x(t+1)=a*x+v      v ~ N(0,r1)  x_0 ~ N(m0,p0)
%
%each having the length nstp. The result (ie. state # nx) is plotted and
%stored in data. Each realization in each column. t is the time.

t=0:nstp; t=t(:)-t(1);
ns=length(a);
data=[];
[v,d]=eig(r1); s1=v*diag(sqrt(diag(d)));
[v,d]=eig(p0); s0=v*diag(sqrt(diag(d)));
for k=1:nk,
 x=s0*randn(ns,1)+m0;
 xt=x(nx);
 for i=1:nstp,
  x=a*x+s1*randn(ns,1);
  xt=[xt;x(nx)];
 end
 data=[data xt];
end
```

simb - ex5



simb - ex5



Background - WT

$$\begin{bmatrix} \dot{\theta}_\epsilon \\ \dot{\omega}_r \\ \dot{\omega}_g \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} 0 & 1 & -\frac{1}{n_{gear}} & 0 \\ -\frac{K_s}{J_r} & \frac{1}{J_r}\frac{\partial T_w}{\partial \omega_r} & 0 & \frac{1}{J_r}\frac{\partial T_w}{\partial \beta} \\ \frac{\eta_{gear} K_s}{n_{gear} J_g} & 0 & -\frac{D_g}{J_g} & 0 \\ 0 & 0 & 0 & -\frac{1}{\tau_\beta} \end{bmatrix} \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{K_\beta}{\tau_\beta} \end{bmatrix} \beta_{ref} + \begin{bmatrix} 0 \\ \frac{1}{J_r}\frac{\partial T_w}{\partial v} \\ 0 \\ 0 \end{bmatrix} v$$

$$\dot{x}_s = A_s x_s + B_s u + B_v v$$

$$P_e = \begin{bmatrix} 0 & 0 & \frac{\eta(1-S)\omega_0}{n_p} D_g & 0 \end{bmatrix} \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \end{bmatrix} + e$$

$$y = C_s x_s + e$$

$$\begin{bmatrix} \dot{v} \\ \ddot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{p_1 p_2} & -\frac{p_1+p_2}{p_1 p_2} \end{bmatrix} \begin{bmatrix} v \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k}{p_1 p_2} \end{bmatrix} e_w$$

$$\dot{x}_w = A_w x_w + B_w e_w$$

$$v = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ \dot{v} \end{bmatrix}$$

$$v = C_w x_w$$

$$\begin{bmatrix} \dot{x}_s \\ \dot{x}_w \end{bmatrix} = \begin{bmatrix} A_s & B_v C_w \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x \\ x_w \end{bmatrix} + \begin{bmatrix} B_s \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ B_w \end{bmatrix} e_w$$

$$y = \begin{bmatrix} C_s & 0 \end{bmatrix} \begin{bmatrix} x_s \\ x_w \end{bmatrix} + e$$

Task

Given measurement equation:

$$y_t = g\left[s_t, n_t\right]$$

and some data or measurement in the interval from t_0 to t_1:

$$Y_{t_0:t_1}$$

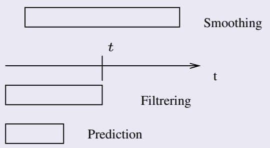
Here is Measurement(y_t), signal (s_t) and noise (n_t).

Task: find the estimate ($\hat{s}_t$) of the signal (or the quantity) s_t based on the available information in $Y_{t_0:t_1}$.

Filter disciplines

Information available from the interval $[t_0;\ t_1]$. Quantity at interest is s_t.

- $t < t_1$ Smoothing, Interpolation
- $t = t_1$ Filtering
- $t > t_1$ Prediction



State estimation (Filtering)

$$x_{t+1} = Ax_t + Bu_t + v_t$$

$$y_t = Cx_t + e_t$$

- Estimation: of x_t by observing y_t
- Dynamics: x_t is changing in time (and stochastic).

- An estimate $\hat{x}_{t|s}$ - of x_t based on Y_s -
- An estimate of its uncertainty $P_{t|s}$
- The likelihood for free
- A residual

Kalman filter

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad x_{t_0} \in \mathbf{N}(\hat{x}_0, P_0) \qquad v_t \in \mathbf{N}(0, R_1)$$

$$y_t = Cx_t + e_t \qquad e_t \in \mathbf{N}(0, R_2)$$

Obey the standard assumptions.

Data updating (inference)

$$\begin{aligned} \hat{x}_{t|t} &= \hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1} \right] \\ \kappa_t &= P_{t|t-1} C^\top [C P_{t|t-1} C^\top + R_2]^{-1} \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1} C^\top [C P_{t|t-1} C^\top + R_2]^{-1} C P_{t|t-1} \end{aligned}$$

▸ data update

Time update (prediction)

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t \qquad \hat{x}_{0|0} = \hat{x}_0$$

$$P_{t+1|t} = AP_{t|t}A^\top + R_1 \qquad P_{0|0} = P_0$$

▸ time evolution

Rudolf Emil Kalman

Kalman was born in Budapest, Hungary, on May 19, 1930.

He received the bachelor's degree (S.B.) and the master's degree (S.M.) in electrical engineering, from the Massachusetts Institute of Technology in 1953 and 1954 respectively.

Problem definition

1. Characteristics of signal and noise
2. Observation model (relation between measurements, signal and noise)
3. Criterium (What is a good estimate)
4. Restrictions (which information can be used)

Focus on 3 and 4.

Restrictions

$$\hat{x} = func(Y)$$

Criterium:

$$J = \mathbf{E}\left\{ \|x - \hat{x}\|^2 \right\}$$

$$\hat{x} = \mathbf{E}\left\{ x|Y \right\}$$

First a little theorem

$$\mathbf{E}\Big\{g(x)\Big\} = \mathbf{E}_Y\Big\{\ \mathbf{E}\Big\{g(x)|Y\Big\}\ \Big\}$$

$$\begin{aligned} J &= \mathbf{E}\Big\{(x-\hat{x})^T(x-\hat{x})\Big\} \\ &= \mathbf{E}_Y\Big\{\ \mathbf{E}\Big\{(x-\hat{x})^T(x-\hat{x})|Y\Big\}\ \Big\} \end{aligned}$$

$$\hat{x} = func(Y)$$

$$\begin{aligned} J_{in} &= \mathbf{E}\Big\{x^Tx - x^T\hat{x} - \hat{x}^Tx + \hat{x}^T\hat{x}|Y\Big\} \\ &= \mathbf{E}\Big\{x^Tx|Y\Big\} - 2\hat{x}^T\mathbf{E}\Big\{x|Y\Big\} + \hat{x}^T\hat{x} \end{aligned}$$

$$\frac{\partial}{\partial\hat{x}}J_{in} = -2\mathbf{E}\Big\{x|Y\Big\} + 2\hat{x}$$

DTU

Evolution of estimate and uncertainty

$$x_{t+1} = Ax_t + Bu_t + v_t$$

$$y_t = Cx_t + e_t$$

$$Y_s = \begin{bmatrix} y_0 \\ \vdots \\ y_s \end{bmatrix} \qquad s \leq t$$

$$\hat{x}_{t+1|s} = A\hat{x}_{t|s} + Bu_t$$

$$P_{t+1|s} = AP_{t|s}A^\top + R_1$$

▸ Kalman filter

DTU

The Projection theorem

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \in N\left(\begin{bmatrix} m_x \\ m_y \end{bmatrix}, \begin{bmatrix} P_x & P_{xy} \\ P_{xy}^\top & P_y \end{bmatrix}\right)$$

$$Y = CX + E$$

$$X|Y \in N(\hat{x}, P_1)$$

$$f(x|y) = \frac{f(x,y)}{f(y)}$$

$$\hat{x} = m_x + P_{xy}P_y^{-1}(y - m_y)$$

$$P_1 = P_x - P_{xy}P_y^{-1}P_{xy}^\top \qquad \leq P_x$$

$$X - \hat{x} \perp Y$$

State estimation

System model:

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \in \mathbf{N}(0, R_1) \qquad x_{t_0} \in \mathbf{N}(\hat{x}_0, P_0)$$

$$y_t = Cx_t + e_t \qquad e_t \in \mathbf{N}(0, R_2) \qquad e_t, v_t \; white \perp x_s \;\; s \leq t$$

Given: $x_t | Y_{t-1} \in \mathbf{N}\left(\hat{x}_{t|t-1}, P_{t|t-1}\right)$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} | Y_{t-1} \in \mathbf{N}\left(\begin{bmatrix} \times \\ \times \end{bmatrix}, \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix} \right) \qquad Y_t = \begin{bmatrix} Y_{t-1} \\ y_t \end{bmatrix}$$

The projection theorem gives then

$$x_t | y_t, Y_{t-1} \quad = \quad x_t | Y_t \quad \in \quad \mathbf{N}\left(\hat{x}_{t|t}, P_{t|t}\right)$$

Let's see how we can find the $\times$'s, then we have a recursion for $\hat{x}$ and P.

$$x_t | Y_{t-1} \in \mathbf{N}\left(\hat{x}_{t|t-1}, P_{t|t-1}\right)$$

This first is for free.

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} | Y_{t-1} \in \mathbf{N}\left(\begin{bmatrix} \hat{x}_{t|t-1} \\ \times \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & \times \\ \times & \times \end{bmatrix}\right)$$

State estimation

$$x_t|Y_{t-1} \in \mathbf{N}\left(\hat{x}_{t|t-1}, P_{t|t-1}\right)$$

Just a copy from last page

$$y_t = Cx_t + e_t \qquad e_t \in \mathbf{N}(0, R_2) \qquad e_t, v_t \; white \; \perp x_s \;\; s \leq t$$

$$y_t|Y_{t-1} \in \mathbf{N}\left(C\hat{x}_{t|t-1}, \;\; CP_{t|t-1}C^\top + R_2\right)$$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} |Y_{t-1} \in \mathbf{N}\left(\begin{bmatrix} \hat{x}_{t|t-1} \\ C\hat{x}_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & \times \\ \times & CP_{t|t-1}C^\top + R_2 \end{bmatrix}\right)$$

State estimation

$$x_t|Y_{t-1} \in \mathbf{N}\left(\hat{x}_{t|t-1}, P_{t|t-1}\right)$$

Just a copy from last page

$$y_t = Cx_t + e_t \qquad e_t \in \mathbf{N}(0, R_2) \qquad e_t, v_t \; white \; \perp x_s \;\; s \leq t$$

$$\mathbf{Cov}\{y_t, x_t|Y_{t-1}\} = CP_{t|t-1}$$

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} |Y_{t-1} \in \mathbf{N}\left(\begin{bmatrix} \hat{x}_{t|t-1} \\ C\hat{x}_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}C^\top \\ CP_{t|t-1} & CP_{t|t-1}C^\top + R_2 \end{bmatrix}\right)$$

State estimation

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} | Y_{t-1} \in \mathbf{N}\left(\begin{bmatrix} \hat{x}_{t|t-1} \\ C\hat{x}_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}C^\top \\ CP_{t|t-1} & CP_{t|t-1}C^\top + R_2 \end{bmatrix}\right)$$

$$x_t | y_t, Y_{t-1} \quad = \quad x_t | Y_t \quad \in \quad \mathbf{N}\left(\hat{x}_{t|t}, P_{t|t}\right)$$

$$\hat{x} = m_x + P_{xy}P_y^{-1}(y - m_y)$$

$$P_1 = P_x - P_{xy}P_y^{-1}P_{xy}^\top$$

Kalman filter

$$X - \hat{x} \perp Y$$

$$\begin{aligned} \hat{x}_{t|t} &= \hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1}\right] \\ \kappa_t &= P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} CP_{t|t-1} \end{aligned}$$

Data updating

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1} \right]$$
$$\kappa_t = P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1}$$
$$P_{t|t} = P_{t|t-1} - P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} CP_{t|t-1}$$

$$x_t - \hat{x}_{t|t} \perp Y_t$$

Time update

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t$$

$$P_{t+1|t} = AP_{t|t}A^\top + R_1$$

Notice: Matrices might be time varying (but in a deterministic way).

Model linearized around $v_m = 7\ m/s$ $\beta_0 = 1^o$ $\omega_{r,0} = 3.68\ rad/s$ and sampled $f_s = 10\ Hz$.
Figure for estimation $v_m = 16\ m/s$.

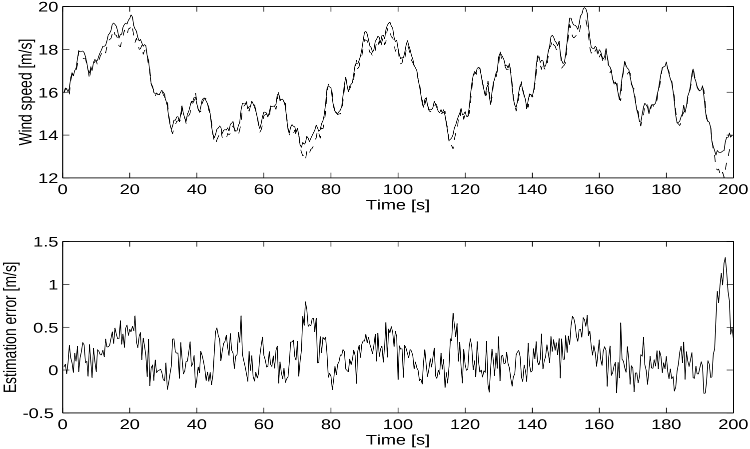
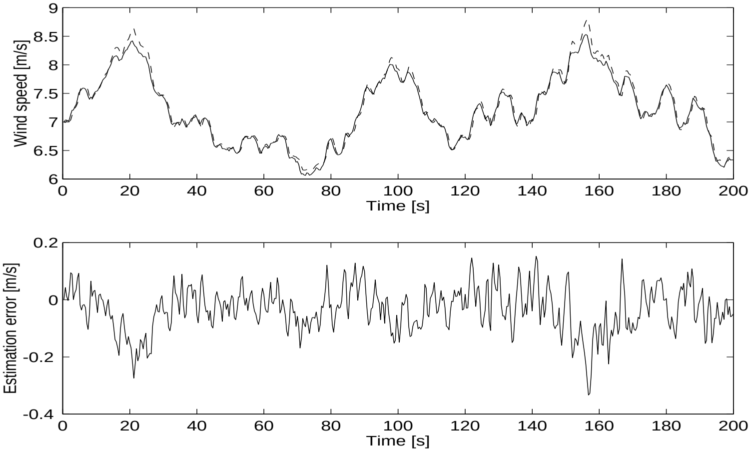


Figure for estimation $v_m = 7\ m/s$.



Mechanization (of data update)

$$\begin{aligned}
\hat{x}_{t|t} &= \hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1}\right] \\
\kappa_t &= P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} CP_{t|t-1}
\end{aligned}$$

$$\begin{aligned}
P_{t|t} &= [I - \kappa_t C]P_{t|t-1} \\
P_{t|t} &= P_{t|t-1} - \kappa_t [CP_{t|t-1}C^\top + R_2]\kappa_t^T
\end{aligned}$$

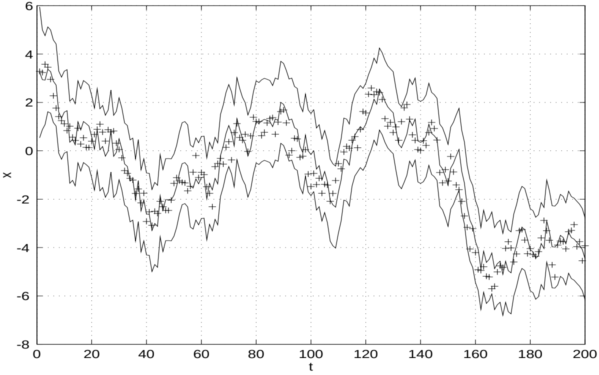
$$\begin{aligned}
\varepsilon_t &= y_t - C\hat{x}_{t|t-1} \\
s_t &= CP_{t|t-1}C^\top + R_2 \\
\kappa_t &= P_{t|t-1}C^\top s_t^{-1} \\
\hat{x}_{t|t} &= \hat{x}_{t|t-1} + \kappa_t \varepsilon_t \\
P_{t|t} &= P_{t|t-1} - \kappa_t s_t \kappa_t^T
\end{aligned}$$

```
%-----------------------------------------------------------------------
% Initil values for the Kalman filter goes here
P=100;
xh=5;
%-----------------------------------------------------------------------

measinit;               % Initilialise the measurement system
for it=1:length(wt),
 w=wt(it);
 [y,t]=meas;                        % Measure levels
 %-----------------------------------------------------------------------
 % Data update goes here
 %-----------------------------------------------------------------------
 err=y-C*xh;
 S=C*P*C'+R2;
 K=P*C'*inv(S);  % K is kappa
 xh=xh+K*err;
 P=P-K*S*K';

 u=0;                               % Uncontrolled situation
 act(u);                            % Actuate control

 data=[data; t w Xs(1) u err];                    % Simulation data
 edata=[edata; t err/sqrt(S) xh' Xs' diag(P)' K']; % Estimation data

 %-----------------------------------------------------------------------
 % Time update goes here
 %-----------------------------------------------------------------------
 xh=A*xh+B*u;
 P=A*P*A'+R1;
end
```

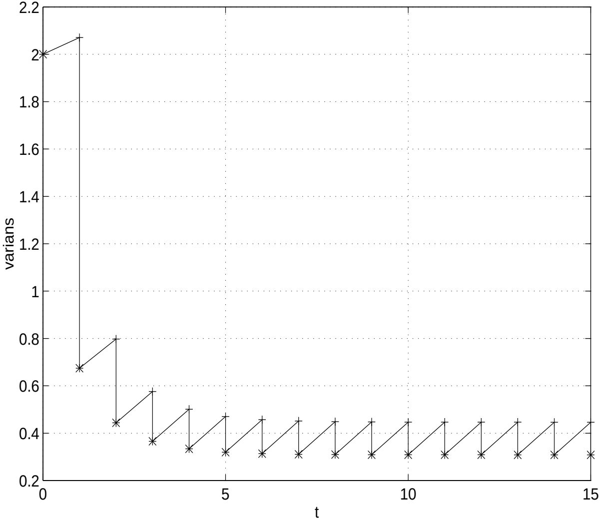
Example

$$x_{t+1} = a x_t + v_t \quad a = 0.98 \;\; r_1 = 0.25$$

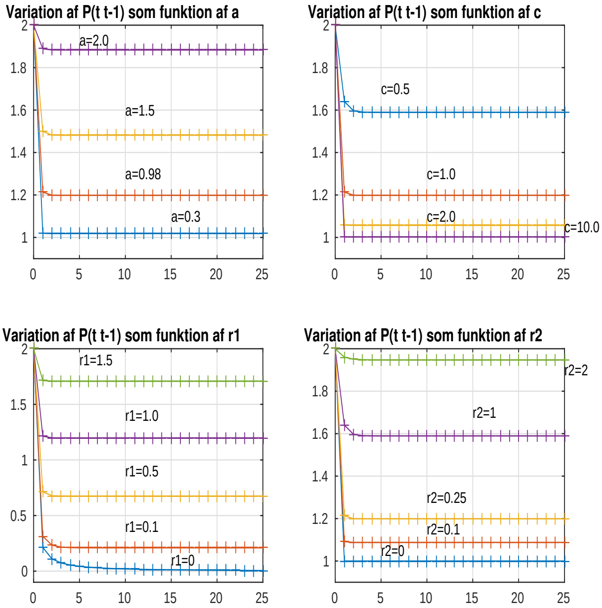
$$y_t = c x_t + e_t \qquad c = 1 \qquad r_2 = 1.25$$



Example ($P_{t|t}$ and $P_{t|t-1}$)



Example (Nominal values: $a = 0.98, c = 1, r_1 = 1, r_2 = 0.25$)



Kalman-Bucy filter (C-time Kalman filter)

$$\frac{d}{dt_c}x = Ax + Bu + v \qquad t_c \in \mathbb{R}$$

$$y = Cx + e$$

Continuous time observations

$$Int\{v_t\} = R_1 \qquad Int\{e_t\} = R_2$$

$$\frac{d}{dt_c}\hat{x} = A\hat{x} + Bu + K[y - C\hat{x}]$$

$$K = PC^T R_2^{-1}$$

$$\frac{d}{dt_c}P = AP + PA^T + R_1 - PC^T R_2^{-1} CP$$

DTU

CD Kalman filter

$$\frac{d}{dt_c}x = Ax + Bu + v$$

$$y_k = Cx_k + e_k \qquad\qquad t_c = kT_s$$

Discrete time observations

Time update $\quad ((k-1)T_s < t_c \leq kT_s)$

$$\frac{d}{dt_c}\hat{x} = A\hat{x} + Bu$$

$$\frac{d}{dt_c}P = AP + PA^T + R_1$$

Data updating

$$\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + \kappa_k \left[y_k - C\hat{x}_{k|k-1}\right] \qquad = \qquad [I - \kappa_k C]\hat{x}_{k|k-1} + \kappa_k y_k \\
\kappa_k &= P_{k|k-1}C^\top[CP_{k|k-1}C^\top + R_2]^{-1} \\
P_{k|k} &= [I - \kappa_t C]P_{k|k-1}
\end{aligned}$$

Lesson learned in L6

- The Kalman filter
  - Time evolution (as for state space descriptions)
  - Projection theorem (Estimation)
  - State estimation (Combined time evolution and estimation)
  - Mechanization (Matlab)
  - Evolution of estimate and uncertainties
  - C-time versions

End L6

Kalman filter versions

Learning objective

- Time and data update version
- The Predictive Kalman filter
- The Ordinary Kalman filter
- The Stationary Kalman tilter
- Corellation bewteen process and measurement noise
- Filter and Corrector version
- Closed filter implementation

System model (Standard Problem):

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \in \mathbf{N}(0, R_1) \qquad x_0 \in \mathbf{N}(\hat{x}_0, P_0)$$

$$y_t = Cx_t + e_t \qquad e_t \in \mathbf{N}(0, R_2)$$

Standard assumptions:

e_t, v_t white and uncorrelated with $x_s \quad s \leq t$. Mutual uncorrelated (for the time being).

State estimation (in copy)

Data updating:

$$\begin{aligned}
\hat{x}_{t|t} &= \overbrace{\hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1} \right]}^{\text{corrector version}} = \overbrace{(I - \kappa_t C)\hat{x}_{t|t-1} + \kappa_t y_t}^{\text{Filter version}} \\
\kappa_t &= P_{t|t-1} C^\top [C P_{t|t-1} C^\top + R_2]^{-1} \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1} C^\top [C P_{t|t-1} C^\top + R_2]^{-1} C P_{t|t-1}
\end{aligned}$$

Time updating:

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t$$

$$P_{t+1|t} = A P_{t|t} A^\top + R_1$$

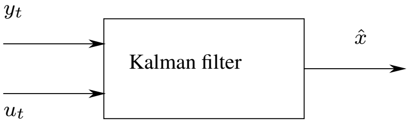
$$x_t - \hat{x}_{t|t} \perp Y_t$$

Notice: Matrices might be time varying (but in a known way).

Telescopic problem

$$\begin{bmatrix} \hat{x}_{t|t-1} \\ P_{t|t-1} \end{bmatrix} \underbrace{\rightarrow}_{\text{data update}} \begin{bmatrix} \hat{x}_{t|t} \\ P_{t|t} \end{bmatrix} \underbrace{\rightarrow}_{\text{time update}} \begin{bmatrix} \hat{x}_{t+1|t} \\ P_{t+1|t} \end{bmatrix} \underbrace{\rightarrow}_{\text{data update}} \begin{bmatrix} \hat{x}_{t+1|t+1} \\ P_{t+1|t+1} \end{bmatrix}$$

or a (closed) filter version



$$\begin{aligned} z_{i+1} &= A_f z_i + B_f u_i^f \\ \hat{x}_i &= C_f z_i + D_f u_i^f \end{aligned}$$

$$u_i^f = \begin{bmatrix} y_i \\ u_i \end{bmatrix}$$

$\hat{x}_i = \hat{x}_{i|i}$ or $\hat{x}_i = \hat{x}_{i|i-1}$

The Predictive Kalman filter

$$\begin{bmatrix} \hat{x}_{t-1|t-1} \\ P_{t-1|t-1} \end{bmatrix} \rightarrow \underbrace{\begin{bmatrix} \hat{x}_{t|t-1} \\ P_{t|t-1} \end{bmatrix} \rightarrow \begin{bmatrix} \hat{x}_{t|t} \\ P_{t|t} \end{bmatrix} \rightarrow \begin{bmatrix} \hat{x}_{t+1|t} \\ P_{t+1|t} \end{bmatrix}}_{The\ \ Predictive\ \ Kalman\ \ filter} \rightarrow \begin{bmatrix} \hat{x}_{t+1|t+1} \\ P_{t+1|t+1} \end{bmatrix}$$

For short $\quad \bar{x}_t = \hat{x}_{t|t-1} \qquad \bar{P}_t = P_{t|t-1}$

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t \qquad \hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1} \right]$$

$$\kappa_t = P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} \qquad K_t = A\kappa_t$$

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K_t[y_t - C\bar{x}_t]$$

$$P_{t+1|t} = AP_{t|t}A^\top + R_1 \qquad P_{t|t} = P_{t|t-1} - P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} CP_{t|t-1}$$

$$\bar{P}_{t+1} = A\bar{P}_t A^\top + R_1 - A\bar{P}_t C^T [C\bar{P}_t C^\top + R_2]^{-1} C\bar{P}_t A^T$$

The Predictive Kalman filter

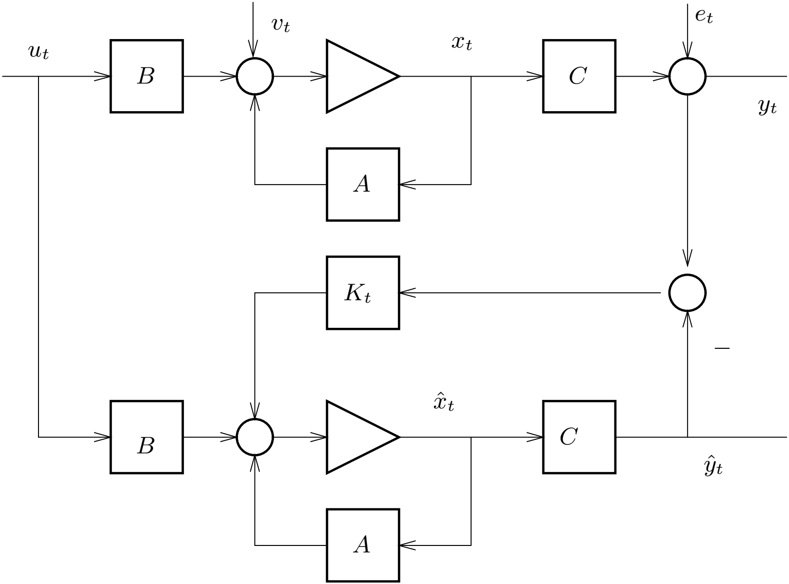
$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K_t[y_t - C\bar{x}_t] \qquad = \qquad [A - K_tC]\bar{x}_t + Bu_t + K_ty_t$$

$$K_t = A\bar{P}_tC^\top[C\bar{P}_tC^\top + R_2]^{-1}$$

$$\bar{P}_{t+1} = A\bar{P}_tA^\top + R_1 - A\bar{P}_tC^T[C\bar{P}_tC^\top + R_2]^{-1}C\bar{P}_tA^T$$

$$\bar{x}_{t_0} = m_0 \quad \bar{P}_{t_0} = \bar{P}_0$$

Block diagram



Kalman-Bucy filter (CC-time Kalman filter)

Purpose of slide: same structure as Predictive Kalman filter.

$$\frac{d}{dt_c}x = Ax + Bu + v \qquad t_c \in \mathbb{R}$$

$$y = Cx + e$$ *Continuous time observations*

$$Int\{v_t\} = R_1 \qquad Int\{e_t\} = R_2$$

$$\frac{d}{dt_c}\hat{x} = A\hat{x} + Bu + K[y - C\hat{x}] \qquad = (A - KC)\hat{x} + Bu + Ky$$

$$K = PC^T R_2^{-1}$$

$$\frac{d}{dt_c}P = \underbrace{AP + PA^T + R_1}_{Free\ \ evolution} - PC^T R_2^{-1} CP$$

More Recursions (KP-recursions)

$$\bar{P}_{t+1} = A\bar{P}_t A^\top + R_1 - A\bar{P}_t C^\top [C\bar{P}_t C^\top + R_2]^{-1} C\bar{P}_t A^\top$$ Riccati version

$$\bar{P}_{t+1} = [A - K_t C]\bar{P}_t A^\top + R_1$$ The short version

$$\bar{P}_{t+1} = A\bar{P}_t A^\top + R_1 - K_t[C\bar{P}_t C^\top + R_2]K_t^T$$ The symmetric version

$$\bar{P}_{t+1} = (A - K_t C)\bar{P}_t (A - K_t C)^\top + R_1 + K_t R_2 K_t^\top$$ Josephson stabilized form

Numerics: Use a UDU ($P = UDU^T$) or a square root method ($P = LL^T$).

Tricks:

$$K = A\bar{P}C^T[C\bar{P}C^T + R_2]^{-1} \qquad K[C\bar{P}C^T + R_2]K^T = KC\bar{P}A^T = A\bar{P}C^T K^T$$

Jacopo Francesco Riccati

Riccati

- Born: 28 May 1676 Venice, Venetian Republic (now Italy)
- Died: 15 April 1754 Treviso, Venetian Republic (now Italy)
- Alma mater: University of Padua

First order system

System

$$x_{i+1} = ax_i + bu_i + v_i \qquad r_1$$
$$y_i = cx_i + e_i \qquad r_2$$

Standard assumptions

Predictive Kalman filter

$$\bar{x}_{i+1} = a\bar{x}_i + bu_i + k_i[y_i - c\bar{x}_i]$$
$$= [a - k_t c]\bar{x}_t + bu_t + k_t y_t$$

$$k_i = a\frac{\bar{p}_i c}{\bar{p}_i c^2 + r_2}$$

$$\bar{p}_{i+1} = a^2\bar{p}_i + r_1 - \frac{(a\bar{p}_i c)^2}{c^2\bar{p}_i + r_2}$$

Kalman filter - the gain

$$k(\bar{p}) = a\frac{\bar{p}c}{\bar{p}c^2 + r_2} = a\frac{1}{c + \frac{r_2}{\bar{p}c}}$$

$$k(0) = 0$$

$$k(\bar{p}) \rightarrow \frac{a}{c} \quad \text{for} \quad \frac{r_2}{\bar{p}c} \rightarrow 0$$

DTU

A first order system - The Riccati equation

For a first order system:

$$\begin{aligned}\bar{p}_{i+1} &= a^2\bar{p}_i + r_1 - \frac{(a\bar{p}_i c)^2}{c^2\bar{p}_i + r_2} \\ &= f(\bar{p}_i)\end{aligned}$$

Firstly :

$$f(0) = r_1 \qquad\qquad r_1 \le f(\bar{p}) \le r_1 + \frac{a^2 r_2}{c^2}$$ (see below for proof)

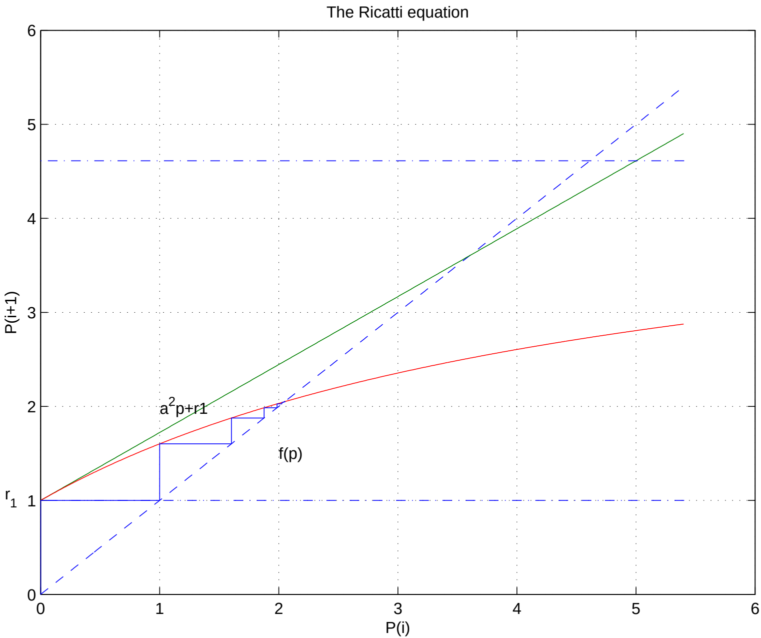
Furthermore

$$f(\bar{p}) \le a^2\bar{p} + r_1$$

Finally

$$\begin{aligned}f(\bar{p}) &= a^2\bar{p}\left(\frac{c^2\bar{p} + r_2 - \bar{p}c^2}{c^2\bar{p} + r_2}\right) + r_1 \\ &= \frac{a^2\bar{p}r_2}{c^2\bar{p} + r_2} + r_1 = \frac{a^2 r_2}{c^2}\left[\frac{1}{1 + \frac{r_2}{c^2\bar{p}}}\right] + r_1 \rightarrow \frac{a^2 r_2}{c^2} + r_1 \quad \text{for} \quad \bar{p} \rightarrow \infty\end{aligned}$$

Riccati equation



Stationarity

If $\bar{P}_t$ converge to a bounded P then P is a solution to the ARE:

$$P = APA^\top + R_1 - APC^T[CPC^T + R_2]^{-1}CPA^\top$$

Let (A, C) detectable (observable). Then for every P_0 there is a bounded limiting solution to the ARE. Furthermore P is positive semi definite.

$$K = APC^T[CPC^T + R_2]^{-1}$$

Let (A, R_1) be reachable (controlable) and $R_2 > 0$. Then (A, C) is detectable iff

- There is a unique positive limiting solution to the ARE which is independent of P_0. P is the unique positive definite solution to the ARE.
- The characteristic system matrix

  $$(A - KC)$$

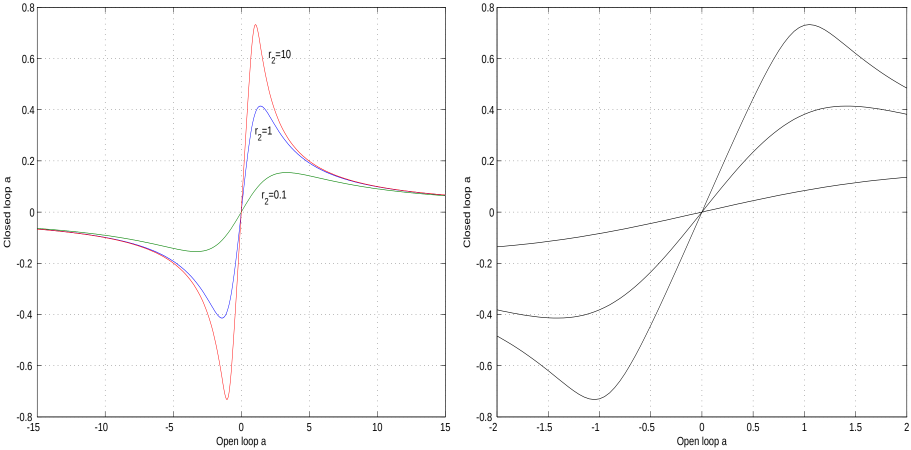
  is asymptotic stable.

DTU

Example

$$x_{t+1} = ax_t + v_t \qquad r_1 = 1 \qquad a_{cl} = a - kc$$

$$y_t = cx_t + e_t \qquad c = 1 \qquad r_2 = 1$$



Notice the assymptotic $1/a$

Correlation between process and measurement I

Process equation:

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad x_{t_0} \in \mathbb{F}(\hat{x}_0, P_0) \qquad v_t \in \mathbb{F}(0, R_1)$$

Measurement equation:

$$y_t = Cx_t + e_t \qquad e_t \in \mathbb{F}(0, R_2)$$

where the disturbances obey the standard assumptions:

$$e_t, v_t \; white \;\; \perp x_t \;\; s \leq t \qquad \mathbf{C}ov\left\{v_t, e_t\right\} = R_{12}$$

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K_t[y_t - C\bar{x}_t]$$

$$K_t = [A\bar{P}_tC^T + R_{12}][C\bar{P}_tC^T + R_2]^{-1}$$

$$\bar{P}_{t+1} = A\bar{P}_tA^T + R_1 - [A\bar{P}_tC^T + R_{12}][C\bar{P}_tC^T + R_2]^{-1}[C\bar{P}_tA^T + R_{12}^T]$$

Correlation between process and measurement I

$$K_t = [A\bar{P}_t C^T + R_{12}][C\bar{P}_t C^T + R_2]^{-1}$$

$$\bar{P}_{t+1} = [A - K_t C]\bar{P}_t[A - K_t C]^T + (K_t - R_{12}R_2^{-1})R_2(K_t - R_{12}R_2^{-1})^T + R_1 - R_{12}R_2^{-1}R_{12}^T$$

In case of perfect correlation:

$$v_t = Ge_t$$

$$R_1 = GR_2G^T \qquad\qquad R_{12} = GR_2$$

The point:

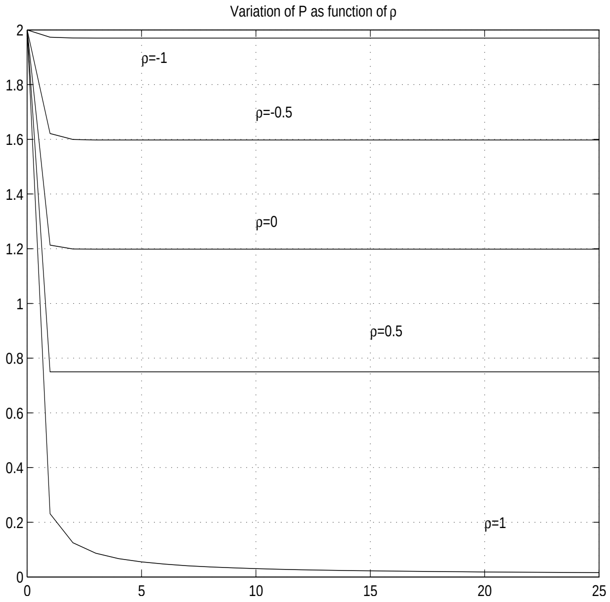
$$P_\infty = 0 \qquad\qquad (K = G)$$

is a stationary point. It is a stable stationary point iff

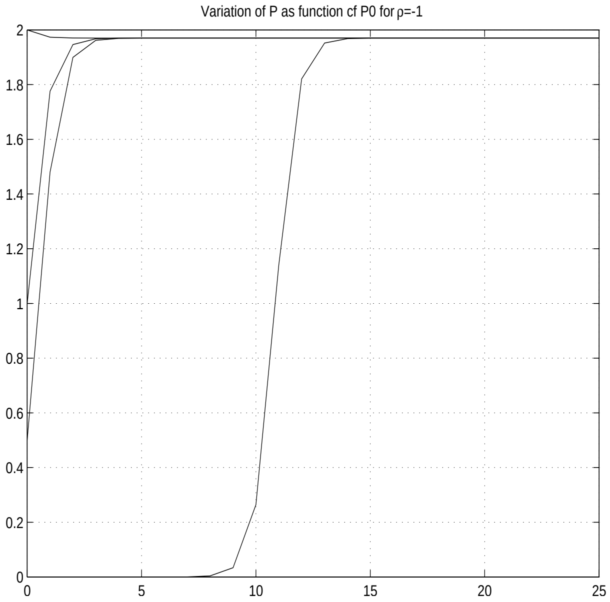
$$A - GC$$

is asymptotically stable.

Example (Nominal values: $a = 0.98, c = 1, r_1 = 1, r_2 = 0.25$)



Example



Variation of P as function cf P0 for ρ=-1
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25

Copy of state estimation

Time update

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t$$

$$P_{t+1|t} = AP_{t|t}A^\top + R_1$$

Data updating

$$\begin{aligned}
\hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + \kappa_{t+1}\left[y_{t+1} - C\hat{x}_{t+1|t}\right] \\
\kappa_{t+1} &= P_{t+1|t}C^\top[CP_{t+1|t}C^\top + R_2]^{-1} \\
P_{t+1|t+1} &= P_{t+1|t} - P_{t+1|t}C^\top[CP_{t+1|t}C^\top + R_2]^{-1}CP_{t+1|t} = [I - \kappa_t C]P_{t|t-1}
\end{aligned}$$

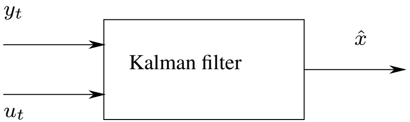
$$x_t - \hat{x}_{t|t} \perp Y_t$$

Notice: Matrices might be time varying.

Telescopic problem

$$\begin{bmatrix} \hat{x}_{t|t-1} \\ P_{t|t-1} \end{bmatrix} \rightarrow \underbrace{\begin{bmatrix} \hat{x}_{t|t} \\ P_{t|t} \end{bmatrix} \rightarrow \begin{bmatrix} \hat{x}_{t+1|t} \\ P_{t+1|t} \end{bmatrix} \rightarrow \begin{bmatrix} \hat{x}_{t+1|t+1} \\ P_{t+1|t+1} \end{bmatrix}}_{\text{Ordinary Kalman filter}}$$

or a (closed) filter version



$$\begin{aligned} z_{t+1} &= A_f z_t + B_f u_t^f \\ y_t^f &= C_f z_t + D_f u_t^f \end{aligned}$$

The Ordinary Kalman filter

For short: $\hat{x}_t = \hat{x}_{t|t}$ $\quad P_t = P_{t|t}$

$\bar{x}_t = \hat{x}_{t|t-1}$ $\quad \bar{P}_t = P_{t|t-1}$

$$\begin{aligned}\hat{x}_{t+1} &= \hat{x}_{t+1|t} + \kappa_{t+1}\big(y_{t+1} - C\hat{x}_{t+1|t}\big) \\ &= A\hat{x}_t + Bu_t + \kappa_{t+1}\big(y_{t+1} - C\big[A\hat{x}_t + Bu_t\big]\big) \\ &= \big(I - \kappa_{t+1}C\big)\big[A\hat{x}_t + Bu_t\big] + \kappa_{t+1}y_{t+1}\end{aligned}$$

$$\hat{x}_{t+1|t} = A\hat{x}_t + Bu_t$$

$$P_{t+1} = P_{t+1|t+1} = [I - \kappa_{t+1}C]P_{t+1|t}$$

$$P_{t+1|t} = AP_tA^T + R_1$$

$$P_{t+1} = [I - \kappa_{t+1}C](AP_tA^T + R_1)$$

$$\hat{x}_t = \big(I - \kappa_tC\big)\big[A\hat{x}_{t-1} + Bu_{t-1}\big] + \kappa_ty_t$$

$$P_t = [I - \kappa_tC](AP_{t-1}A^T + R_1)$$

Corellation beween process and measurement II

System model:

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad x_0 \in \mathbf{N}\left(\hat{x}_0, P_0\right)$$

$$y_t = Cx_t + e_t \qquad e_t, v_t \ white \ \perp x_s \ \ s \leq t$$

$$v_t \in \mathbf{N}\left(0, R_1\right) \qquad\qquad e_t \in \mathbf{N}\left(0, R_2\right)$$

$$R_{12} = \mathbf{C}ov\left\{v_t, e_t\right\}$$

Correlation between process and measurement II

$$x_{t+1} = Ax_t + Bu_t + v_t + \underbrace{M(y_t - y_t)}_{trick} \qquad y_t = Cx_t + e_t$$

$$x_{t+1} = Ax_t + Bu_t + My_t - MCx_t + \underbrace{v_t - Me_t}_{\tilde{v}_t} \qquad \tilde{A} = A - MC$$

$$\tilde{R}_{12} = \mathbf{E}\left\{(v_t - Me_t)e_t^T\right\} = R_{12} - MR_2 \quad = \quad 0 \quad \textit{for} \quad R_{12} = MR_2$$

$$\tilde{R}_1 = R_1 + MR_2M^T - R_{12}M^T - MR_{12}^T \quad = \quad R_1 - R_{12}R_2^{-1}R_{12} \quad \text{for} \quad R_{12} = MR_2$$

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t + M(y_t - C\hat{x}_{t|t}) = (A - MC)\hat{x}_{t|t} + Bu_t + My_t$$

$$P_{t+1|t} = \tilde{A}P_{t|t}\tilde{A}^T + \tilde{R}_1$$

Data update as usual.

DTU

The Ordinary Kalman filter (with correlation)

Time update:

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t + M(y_t - C\hat{x}_{t|t}) \quad = \quad (A - MC)\hat{x}_{t|t} + Bu_t + My_t$$

$$P_{t+1|t} = \tilde{A}P_{t|t}\tilde{A}^T + \tilde{R}_1 \qquad \tilde{A} = A - MC \qquad \tilde{R}_1 = R_1 - R_{12}R_2^{-1}R_{12}$$

Data updating

$$\begin{aligned}
\hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + \kappa_{t+1}\left[y_{t+1} - C\hat{x}_{t+1|t}\right] \\
\kappa_{t+1} &= P_{t+1|t}C^\top[CP_{t+1|t}C^\top + R_2]^{-1} \\
P_{t+1|t+1} &= P_{t+1|t} - P_{t+1|t}C^\top[CP_{t+1|t}C^\top + R_2]^{-1}CP_{t+1|t} = [I - \kappa_{t+1}C]P_{t+1|t}
\end{aligned}$$

$$x_t - \hat{x}_{t|t} \perp Y_t$$

Notice: Matrices might be time varying. L8a

Close filter forms

The Predictive Kalman filter

For short: $\hat{x}_t = \hat{x}_{t|t}$ $\quad P_t = P_{t|t}$

$\bar{x}_t = \hat{x}_{t|t-1}$ $\quad \bar{P}_t = P_{t|t-1}$

$$\bar{x}_{t+1} = (A - K_t C)\bar{x}_t + Bu_t + K_t y_t = A\bar{x}_t + Bu_t + K_t(y_t - C\bar{x}_t)$$

$$\hat{x}_{t|t-1} = I \quad \bar{x}_t$$

$$K_t = A\bar{P}_t C^\top (C\bar{P}_t C^\top + R_2)^{-1} = A\kappa_t$$

$$\bar{P}_{t+1} = (A - K_t C)\bar{P}_t A^\top + R_1$$

$$\tilde{x}_t = x_t - \hat{x}_{t|t-1} \perp Y_{t-1}$$

$$\in \mathbf{N}(0, \bar{P}_t)$$

DTU

The Ordinary Kalman filter parametrized as a predictive KLM

For short:
$$\hat{x}_t = \hat{x}_{t|t} \qquad P_t = P_{t|t}$$
$$\bar{x}_t = \hat{x}_{t|t-1} \qquad \bar{P}_t = P_{t|t-1}$$

$$\bar{x}_{t+1} = A(I - \kappa_t C)\bar{x}_t + Bu_t + A\kappa_t y_t \qquad = A\bar{x}_t + Bu_t + A\kappa_t(y_t - C\bar{x}_t)$$
$$\hat{x}_t = (I - \kappa_t C)\bar{x}_t + \kappa_t y_t \qquad = \bar{x}_t + \kappa_t(y_t - C\bar{x}_t)$$

$$\bar{P}_{t+1} = AP_tA^\top + R_1$$
$$= A(I - \kappa_t C)\bar{P}_tA^\top + R_1$$

State variance

$$\kappa_t = \bar{P}_tC^\top(C\bar{P}_tC^\top + R_2)^{-1}$$
$$P_t = (I - \kappa_t C)\bar{P}_t \qquad \leq \bar{P}_t$$

Output variance

$$\tilde{x}_t = x_t - \hat{x}_t \perp Y_t$$
$$\in \mathbf{N}(0, P_t)$$

The ordinare Kalman filter

$$\bar{x}_{t+1} = A\big(I - \kappa C\big)\bar{x}_t + Bu_t + A\kappa y_t \quad = A\bar{x}_t + Bu_t + A\kappa_t(y_t - C\bar{x}_t) \qquad K_t = A\kappa_t$$

$$\hat{x}_t = \big(I - \kappa C\big)\bar{x}_t + \kappa y_t \quad = \bar{x}_t + \kappa_t(y_t - C\bar{x}_t)$$

The predictive Kalman filter

$$\bar{x}_{t+1} = \big(A - KC\big)\bar{x}_t + Bu_t + Ky_t \quad = A\bar{x}_t + Bu_t + K(y_t - C\bar{x}_t)$$

$$\hat{x}_{t|t-1} = I \quad \bar{x}_t$$

The prediction error

$$\varepsilon_t = y_t - C\bar{x}_t$$

The Kalman filters

$$\bar{x}_{t+1} = A\big(I - \kappa C\big)\bar{x}_t + Bu_t + A\kappa_t y_t = A\bar{x}_t + Bu_t + A\kappa_t\left(y_t - C\bar{x}_t\right)$$

$$\begin{bmatrix} \hat{x} \\ \bar{x} \\ \varepsilon \end{bmatrix}_t = \begin{pmatrix} I - \kappa C \\ I \\ -C \end{pmatrix} \bar{x}_t + \begin{pmatrix} \kappa \\ 0 \\ I \end{pmatrix} y_t$$

End L7

L8 - Learning objectives

- Standard form and assumptions
- The estimation error
- The prediction error
- The innovation form
- Prediction I and II
- Non-linear systems

System model - The standard form

Process equation:

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad x_{t_0} \in \mathbb{F}(0, P_0) \qquad v_t \in \mathbb{F}(0, R_1)$$

Measurement equation:

$$y_t = Cx_t + e_t \qquad e_t \in \mathbb{F}(0, R_2)$$

where the disturbances obey the standard assumptions:

$$e_t, v_t \; white \;\; \perp x_s \;\; s \le t \qquad\qquad \mathbf{C}ov\left\{v_t, e_t\right\} = R_{12}$$

▸ Oklm

The (discrete time) Kalman filter

Data updating (Inference, estimation)

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1} \right] = [I - \kappa_t C]\hat{x}_{t|t-1} + \kappa_t y_t$$

$$\kappa_t = P_{t|t-1} C^\top [C P_{t|t-1} C^\top + R_2]^{-1}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1} C^\top [C P_{t|t-1} C^\top + R_2]^{-1} C P_{t|t-1} = [I - \kappa_t C] P_{t|t-1}$$

Time update (Prediction, evolution, dynamics)

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t + M \left[y_t - C\hat{x}_{t|t} \right] = \tilde{A}\hat{x}_{t|t} + Bu_t + My_t$$

$$P_{t+1|t} = \tilde{A} P_{t|t} \tilde{A}^T + \tilde{R}_1 \qquad \tilde{R}_1 = R_1 - R_{12} R_2^{-1} R_{12}$$

where

$$R_{12} = MR_2 \qquad \tilde{A} = A - MC$$

$$x_t - \hat{x}_{t|t} \perp Y_t$$

Notice: Matrices can be time dependent

The Predictive Kalman filter

For short $\quad \bar{x}_t = \hat{x}_{t|t-1} \qquad \bar{P}_t = P_{t|t-1}$

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K_t[y_t - C\bar{x}_t]$$

$$K_t = [A\bar{P}_t C^T + R_{12}][C\bar{P}_t C^T + R_2]^{-1}$$

$$\bar{P}_{t+1} = A\bar{P}_t A^T + R_1 - [A\bar{P}_t C^T + R_{12}][C\bar{P}_t C^T + R_2]^{-1}[C\bar{P}_t A^T + R_{12}{}^T]$$

- Same structure as C-time filter.
- Same structure as an observer.
- No computational time delay
- Less expression as the ordinary Kalman filter

Time update:

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + Bu_{t-1} + M(y_{t-1} - C\hat{x}_{t-1|t-1}) \quad = \quad (A - MC)\hat{x}_{t-1|t-1} + Bu_{t-1} + My_{t-1}$$

$$P_{t|t-1} = \tilde{A}P_{t-1|t-1}\tilde{A}^T + \tilde{R}_1 \qquad \tilde{A} = A - MC \qquad \tilde{R}_1 = R_1 - R_{12}R_2^{-1}R_{12}$$

Data updating

$$\begin{aligned}
\hat{x}_{t|t} &= \hat{x}_{t|t-1} + \kappa_t \left[y_t - C\hat{x}_{t|t-1}\right] \\
\kappa_t &= P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1}C^\top [CP_{t|t-1}C^\top + R_2]^{-1} CP_{t|t-1} = [I - \kappa_t C]P_{t|t-1}
\end{aligned}$$

$$x_t - \hat{x}_{t|t} \perp Y_t$$

Notice: Matrices might be time varying (but in a deterministic way).

Actual system

Consider the system

$$\begin{aligned} x_{i+1} &= Ax_i + Bu_i + Gw_i \\ y_i &= Cx_i + e_i \end{aligned}$$

where there is a correlation between w_i and the state vector x_i. For convenience assume the means are zero.

Then there exist an $v_i \in \mathbf{N}(0, R_1)$ and a H such that:

$$w_i = Hx_i + v_i \qquad\qquad v_i \perp x_i$$

The system can then in the standard form be given as:

$$\begin{aligned} x_{i+1} &= (A + GH)x_i + Bu_i + Gv_i \\ y_i &= Cx_i + e_i \end{aligned}$$

Recap

Assume

$$X \in \mathbf{N}(m_x, P_x) \qquad Y \in \mathbf{N}(m_y, P_y)$$

and $\mathbf{C}ov\left\{Y, X\right\} = R_{yx}$.

Then there exists an H and a

$$V \in \mathbf{N}(m_v, P_v) \perp X$$

such that:

$$Y = HX + V$$

where

$$R_{yx} = HP_x$$

$$m_y = Hm_x + m_v \qquad P_y = HP_xH^T + P_v$$

Coloured measurement noise

Consider the system

$$\begin{aligned} x_{i+1} &= Ax_i + Bu_i + v_i \\ y_i &= Cx_i + w_i \end{aligned}$$

where

$$\begin{aligned} z_{i+1} &= A_w z_i + \xi_i \\ w_i &= C_w z_i + \zeta_i \end{aligned}$$

This can also be expressed in the standard form as

$$\begin{aligned} \begin{bmatrix} x \\ z \end{bmatrix}_{i+1} &= \begin{bmatrix} A & 0 \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_i + \begin{bmatrix} B \\ 0 \end{bmatrix} u_i + \begin{bmatrix} v_i \\ \xi_i \end{bmatrix} \\ y_i &= \begin{bmatrix} C & C_w \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_i + \zeta_i \end{aligned}$$

Coloured process noise

Consider the system

$$\begin{aligned} x_{i+1} &= Ax_i + Bu_i + w_i \\ y_i &= Cx_i + e_i \end{aligned}$$

where

$$\begin{aligned} z_{i+1} &= A_w z_i + \xi_i \\ w_i &= C_w z_i + \zeta_i \end{aligned}$$

This can also be expressed in the standard form as

$$\begin{aligned} \begin{bmatrix} x \\ z \end{bmatrix}_{i+1} &= \begin{bmatrix} A & C_w \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_i + \begin{bmatrix} B \\ 0 \end{bmatrix} u_i + \begin{bmatrix} \zeta_i \\ \xi_i \end{bmatrix} \\ y_i &= \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_i + e_i \end{aligned}$$

Offset in the output

Consider the system

$$\begin{aligned} x_{i+1} &= Ax_i + Bu_i + v_i \\ y_i &= Cx_i + Hd_i + e_i \end{aligned}$$

Let us model this as:

$$d_{i+1} = d_i + \zeta_i$$

This can also be expressed in the standard form as

$$\begin{aligned} \begin{bmatrix} x \\ d \end{bmatrix}_{i+1} &= \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_i + \begin{bmatrix} B \\ 0 \end{bmatrix} u_i + \begin{bmatrix} v_i \\ \zeta \end{bmatrix}_i \\ y_i &= \begin{bmatrix} C & H \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_i + e_i \end{aligned}$$

Offset in the process equation

Consider the system

$$\begin{aligned} x_{i+1} &= Ax_i + Bu_i + Gd_i + v_i \\ y_i &= Cx_i + e_i \end{aligned}$$

Let us model this as:

$$d_{i+1} = d_i + \zeta_i$$

This can also be expressed in the standard form as

$$\begin{aligned} \begin{bmatrix} x \\ d \end{bmatrix}_{i+1} &= \begin{bmatrix} A & G \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_i + \begin{bmatrix} B \\ 0 \end{bmatrix} u_i + \begin{bmatrix} v_i \\ \zeta \end{bmatrix}_i \\ y_i &= \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_i + e_i \end{aligned}$$

The Estimation and its error

The results of the kalman filter:

$$x_t | Y_t \in \mathbf{N}\left(\hat{x}_{t|t}, P_{t|t}\right) = \mathbf{N}\left(\hat{x}_t, P_t\right)$$

$$x_t | Y_{t-1} \in \mathbf{N}\left(\hat{x}_{t|t-1}, P_{t|t-1}\right) = \mathbf{N}\left(\bar{x}_t, \bar{P}_t\right)$$

The (estimation) errors:

$$\begin{aligned} \tilde{x}_t &\equiv \tilde{x}_{t|t} = x_t - \hat{x}_{t|t} \perp Y_t \\ \underline{\tilde{x}}_t &\equiv \tilde{x}_{t|t-1} = x_t - \hat{x}_{t|t-1} \perp Y_{t-1} \end{aligned}$$

and their distributions (notice unconditional):

$$\begin{aligned} \tilde{x}_t &= \tilde{x}_{t|t} \in N(0, P_t) \\ \underline{\tilde{x}}_t &= \tilde{x}_{t|t-1} \in N(0, \bar{P}_t) \end{aligned}$$

Can also be found by (analysis of) its recursion:

$$\begin{aligned} \underline{\tilde{x}}_{t+1} &= (A - K_t C)\underline{\tilde{x}}_t - K_t e_t + v_t \\ \tilde{x}_t &= (I - \kappa_t C)\underline{\tilde{x}}_t - \kappa_t e_t \end{aligned} \qquad K = A\kappa$$

$$\tilde{x}_{t+1} = (I - \kappa_t C)(A\tilde{x}_t + v_t) + v_t - \kappa_t e_t$$

Alternative recursion

DTU

The prediction error, The residual, The innovation

$$\varepsilon_t = y_t - C\hat{x}_{t|t-1}$$

$$\varepsilon_t = C\underline{\tilde{x}}_t + e_t \;\perp\; Y_{t-1}$$

$$\varepsilon_t \in \mathbf{N}\left(0, C\bar{P}_t C^\top + R_2\right)$$ *and white (for correct model)*

---

1. Model validation
2. System representation
3. Diagnostic (fault detection)

System representation - Innovation form

The standard form:

$$x_{t+1} = Ax_t + Bu_t + v_t$$

$$y_t = Cx_t + e_t$$

The innovation form:

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K_t\varepsilon_t$$

$$y_t = C\bar{x}_t + \varepsilon_t$$

DTU

Stochastic S-SS to DS-trf

$$x_{t+1} = Ax_t + Bu_t + v_t$$

$$y_t = Cx_t + e_t$$

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K\varepsilon_t$$

$$y_t = C\bar{x}_t + \varepsilon_t$$

$$y_t = H_s(q)u_t + H_n(q)\epsilon_t$$

$$H_s(q) = C[qI - A]^{-1}B$$

$$H_n(q) = C[qI - A]^{-1}K + 1$$

dlqe, ss2tf

$$\frac{d}{dt_c} x(t_c) = A_c x(t_c) + B_c u(t_c) + v(t_c) \qquad v(t_c) \in N(0, \Sigma_1)$$

$$y_i = C^\top x(iT) + e(iT) \qquad e(iT) \in N(0, \Sigma_2)$$

$$x_{i+1} = Ax_i + Bu_i + v_i \qquad v_i \in N(0, R_1)$$

c2d, cn2dn.

$$y_i = Cx_i + e_i \qquad e_i \in N(0, R_2)$$

$$x_{i+1} = Ax_i + Bu_i + Ke_i$$

$$y_i = Cx_i + e_i$$

$$y_i = C[qI - A]^{-1} Bu_i + \left(C[qI - A]^{-1} K + 1\right) e_i$$

Jump estimation of a constant

Example: Estimation of a constant

$$x_{t+1} = x_t$$

$$y_t = x_t + e_t \qquad e_t \in \mathbf{N}(0, r_2)$$

$$\hat{x}_t = \frac{1}{t}\sum_{i=1}^{t} y_i \qquad \hat{x}_{t+1} = \hat{x}_t + \frac{1}{t+1}\left[y_{t+1} - \hat{x}_t\right] \qquad \frac{1}{t+1} = \frac{1}{t}\left[1 - \frac{1}{t+1}\right]$$

Time update:

$$\hat{x}_{t+1|t} = \hat{x}_{t|t} = \hat{x}_t \qquad p_{t+1|t} = p_{t|t} = p_t$$

Data update:

$$\kappa_{t+1} = \frac{p_t}{p_t + r_2}$$

$$\hat{x}_{t+1} = \hat{x}_t + \frac{p_t}{p_t + r_2}\left[y_{t+1} - \hat{x}_t\right]$$

$$p_{t+1} = \left[1 - \frac{p_t}{p_t + r_2}\right]p_t = \frac{p_t r_2}{p_t + r_2}$$

DTU

Defining $q_t = 1/p_t$ we have

$$q_{t+1} = q_t + \frac{1}{r_2} \qquad\qquad q_t = q_0 + t\frac{1}{r_2}$$

and

$$k_{t+1} = \frac{1}{1 + r_2 q_t} = \frac{1}{1+t} \qquad \text{for} \qquad q_0 = 0$$

Kalman filter coinside with LS/ML estimate if $q_0 \rightarrow 0$ ($p_0 \rightarrow \infty$).

Kalman filter is merging (fusion) a priori information with data.

Hemes Inversion lemma (Frobinius-Schur, Sherman-Morrison, Woodbury)

$$[\underline{A} + \underline{BC}^{-1}\underline{D}]^{-1} = \underline{A}^{-1} - \underline{A}^{-1}\underline{B}[\underline{C} + \underline{D}^T\underline{A}^{-1}\underline{B}]^{-1}\underline{D}^T\underline{A}^{-1}$$

If compared to (the data update of the P matrix)

$$P = \bar{P} - \bar{P}C^T\left[C\bar{P}^T + R_2\right]^{-1}C\bar{P}$$

$\underline{A}^{-1} = \bar{P}$ $\underline{B} = C^T$ $\underline{C} = R_2$ $\underline{D}^T = C$

we see that

$$P^{-1} = \bar{P}^{-1} + C^T R_2^{-1} C$$

$$\kappa = PC^T R_2^{-1}$$

The PK recursions

For short: $P = P_{t+1|t+1}$, $\bar{P} = P_{t+1|t}$, $\kappa = \kappa_{t+1}$

The update of the P-matrix can be done according to:

$$P = \bar{P} - \bar{P}C^T \left[C\bar{P}^T + R_2\right]^{-1} C\bar{P}$$

or

$$P^{-1} = \bar{P}^{-1} + C^T R_2^{-1} C$$

and then the gain according to

$$\kappa = PC^T R_2^{-1}$$

Kalman filter for scalar observation

Time update (Prediction, evolution, dynamic)

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t$$

$$P_{t+1|t} = AP_{t|t}A^T + R_1$$

Data updating (Inference, estimation)

$$\kappa_{t+1} = \frac{P_{t+1|t}C^T}{CP_{t+1|t}C^T + R_2}$$

$$\varepsilon_{t+1} = y_{t+1} - C\hat{x}_{t+1|t}$$

$$\hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + \kappa_{t+1}\varepsilon_{t+1}$$

$$P_{t+1|t+1} = [I - \kappa_{t+1}C]P_{t+1|t}$$

$$x_t - \hat{x}_{t|t} \perp Y_t$$

Processing one channel at a time

Assume the observations has a special structure:

$$R_2 = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \qquad C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}$$

The the prcission update

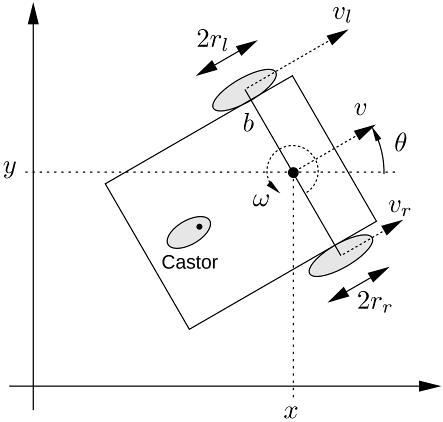
$$P^{-1} = \bar{P}^{-1} + C^T R_2^{-1} C$$

for this case results in

$$P^{-1} = \bar{P}^{-1} + C_1^T \Sigma_1^{-1} C_1 + C_2^T \Sigma_2^{-1} C_2$$

This means we can process one block at a time (also for the other alternative data updates).

What happens if we process one channel at a time even if R_2 do not have this structure.


v_l
$2r_l$
b
v
θ
y
ω
v_r
Castor
$2r_r$
x

Odometric dynamics

$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix}_{t+1} = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix}_t + \begin{bmatrix} v_t T_s \, cos(\theta_t + \frac{1}{2}\omega_t T_s) \\ v_t T_s \, sin(\theta_t + \frac{1}{2}\omega_t T_s) \\ \omega_t \end{bmatrix} + \zeta_t$$

$$\omega_r = \frac{2v_t + b\omega_t}{2r_r} \qquad \omega_l = \frac{2v_t - b\omega_t}{2r_l}$$

Encoder measurements

$$z_t = \begin{bmatrix} \omega_r \\ \omega_l \end{bmatrix} + \xi_t$$

Vision based sensor

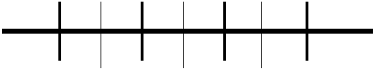
$$y_t = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ \theta_t \\ x_g \\ y_g \end{bmatrix} + e_t \qquad \begin{bmatrix} x_g \\ y_g \end{bmatrix} \in \mathbf{N}\left(m_g, P_g\right)$$

The quantities to be estimated:

$$x = \begin{bmatrix} x \\ y \\ \theta \\ v \\ \omega \end{bmatrix}$$

Different sampling rate

Different sampling rate but with $h_1 = kh_2$ (and k being a integer).



What if not a rational ratio ?

Prediction - marginal

From the Kalman filter we have

$$x_t | Y_t \in \mathbf{N}\left(\hat{x}_{t|t}, P_{t|t}\right)$$

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + v_t & R_1 \\ y_t &= Cx_t + e_t & R_2 \\ z_s &= C_z x_s + w_s & R_w \end{aligned}$$

$$\begin{aligned} \hat{x}_{s+1|t} &= A\hat{x}_{s|t} + Bu_s & s \geq t \\ P_{s+1|t} &= AP_{s|t}A^T + R_1 \end{aligned}$$

$$z_s | Y_t \in \mathbf{N}\left(C\hat{x}_{s|t}, CP_{s|t}C^T + R_w\right)$$

$$y_{t+1} | Y_t \in \mathbf{N}\left(C\hat{x}_{t+1|t}, CP_{t+1|t}C^T + R_2\right)$$

Obs - system identification

Prediction - Simultaneous

$$Z_{t:s} = \begin{bmatrix} z_t \\ \vdots \\ z_s \end{bmatrix}$$

$$Z_{t:s} = \mathcal{W}_o x_t + \Pi_u U_{t:s} + \Pi_v V_{t:s} + W_{t:s}$$

$$Z_{t:s}|Y_t \in \mathbf{N}\left(\hat{Z}_{t:s}, \Sigma_{t:s}\right)$$

$$\hat{Z}_{t:s} = \mathcal{W}_o \hat{x}_{t|t} + \Pi U_{t:s}$$

$$\Sigma_{t:s} = \mathcal{W}_o P_{t|t} \mathcal{W}_o^T + \Pi_v \mathbb{R}_1 \Pi_v^T + \mathbb{R}_2 \qquad + \Pi_v \mathbb{R}_{12} + \mathbb{R}_{12} \Pi_v^T$$

Non linear systems

$$
\begin{aligned}
x_{t+1} &= f(x_t, u_t, v_t) \simeq d + Ax_t + Bu_t + Gv_t + \ldots \\
y_t &= g(x_t, u_t, e_t) \simeq \delta + Cx_t + Du_t + He_t + \ldots
\end{aligned}
$$

$$
A = \frac{\partial}{\partial x} f \qquad B = \frac{\partial}{\partial u} f \qquad G = \frac{\partial}{\partial v} f \qquad R_1(x) = GR_vG^\top
$$

$$
C = \frac{\partial}{\partial x} g \qquad D = \frac{\partial}{\partial u} g \qquad H = \frac{\partial}{\partial e} g \qquad R_2(x) = HR_eH^\top
$$

Matrices depend on point of linearisation (trajectory).

$$\mathbf{E}\Big\{f(x_t, u_t, v_t)|Y_t\Big\} \equiv \hat{f}(\hat{x}_t, u_t) \simeq \hat{A}\hat{x}_{t|t} + \hat{B}u_t + d$$

$$\hat{y}_{t|t-1} = \mathbf{E}\Big\{g(x_t, u_t, e_t)|Y_t\Big\} \equiv \hat{g}(\hat{x}_t, u_t) \simeq C\hat{x}_{t|t-1} + \hat{D}u_t + \delta$$

Time update

$$\hat{x}_{t+1|t} = \hat{f}(\hat{x}_t, u_t)$$

$$P_{t+1|t} = \hat{A}P_{t|t}\hat{A}^\top + \hat{R}_1$$

Data update

$$\hat{x}_{t+1|t+1} = \hat{x}_{t+1|t} + \kappa_t\left[y_{t+1} - \hat{g}(\hat{x}_t, u_t)\right]$$

$$\kappa_{t+1} = P_{t+1|t}\hat{C}^\top[\hat{C}P_{t+1|t}\hat{C}^\top + \hat{R}_2]^{-1}$$

$$P_{t+1|t+1} = [I - \kappa_{t+1}\hat{C}]P_{t+1|t}$$

And similarily in Continuous time.

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

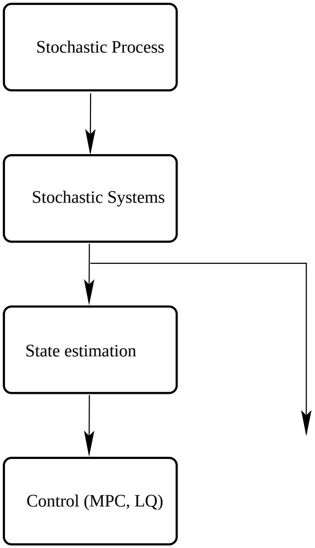
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

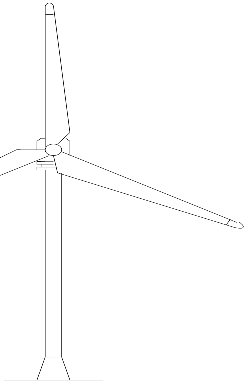
Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

State Space Control (L9-L10)

State space control



State space control



▸ L10

Wind turbine model

$$\begin{bmatrix} \dot{\theta}_\epsilon \\ \dot{\omega}_r \\ \dot{\omega}_g \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} 0 & 1 & -\frac{1}{n_{gear}} & 0 \\ -\frac{K_s}{J_r} & \frac{1}{J_r}\frac{\partial T_w}{\partial \omega_r} & 0 & \frac{1}{J_r}\frac{\partial T_w}{\partial \beta} \\ \frac{\eta_{gear} K_s}{n_{gear} J_g} & 0 & -\frac{D_g}{J_g} & 0 \\ 0 & 0 & 0 & -\frac{1}{\tau_\beta} \end{bmatrix} \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{K_\beta}{\tau_\beta} \end{bmatrix} \beta_{ref} + \begin{bmatrix} 0 \\ \frac{1}{J_r}\frac{\partial T_w}{\partial v} \\ 0 \\ 0 \end{bmatrix} v$$

$$P_e = \begin{bmatrix} 0 & 0 & \frac{\eta(1-S)\omega_0}{n_p} D_g & 0 \end{bmatrix} \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \end{bmatrix}$$

Wind model

$$\begin{bmatrix} \dot{v} \\ \ddot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{p_1 p_2} & -\frac{p_1+p_2}{p_1 p_2} \end{bmatrix} \begin{bmatrix} v \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k}{p_1 p_2} \end{bmatrix} e_w$$

$$v = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ \dot{v} \end{bmatrix}$$

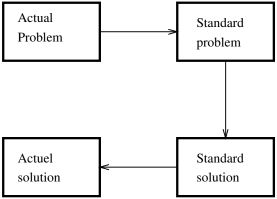
Total model

$$\begin{bmatrix} \dot{x} \\ \dot{x}_w \end{bmatrix} = \begin{bmatrix} A & B_v C_w \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x \\ x_w \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ B_w \end{bmatrix} e_w$$

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

Problem definition:

1. System dynamics (model)
2. Environment (set point, disturbances)
3. Objective, cost function (another model)
4. Constraints (available information)



Introducing the standard problem

Stochastic system:

$$x_{t+1} = Ax_t + Bu_t + v_t$$

Objective: to keep x_t as close as possible to origin

$$J = \sum_{i=t}^{N} \|x_i\|^2 \qquad = \sum_{i=t}^{N} x_i^T x_i$$

Maybe not all direction are equally important:

$$J = \sum_{i=t}^{N} x_i^T Q_1 x_i \qquad Q_1(n \times n) \geq 0, \text{SYMMETRIC}$$

But that have some costs (no free meals) $\rightarrow$ control cost in objective

$$J = \sum_{i=t}^{N} x_i^T Q_1 x_i + u_i^T Q_2 u_i$$

DTU

Introducing the standard problem

$$J = \sum_{i=t}^{N} x_i^T Q_1 x_i + u_i^T Q_2 u_i$$

Since the system (and x_t) is stochastic J is a stochastic variable.

How do we rank different strategies ie. how do we define $J_1 > J_2$.

One (and the most common) option is to use the expected or mean value:

$$J = \mathbf{E}\Big\{ \sum_{i=t}^{N} x_i^T Q_1 x_i + u_i^T Q_2 u_i \Big\}$$

This choice reflects some kind of average thinking.

The stochastic system

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \perp v_s,\ x_s \text{ for } s \leq t$$

The objective function

$$J_t = \mathbf{E}\Big\{ \sum_{i=t}^{N} x_i^\top Q_1 x_i + u_i^\top Q_2 u_i \Big\}$$

The constraints

$$u_t = func(x_t), \qquad u_t = func(Y_t), \qquad u_t = func(Y_{t-1})$$

Notice: disturbances, references, frequency weights ao. might be build into the model.

The Standard Problem I - Output control

Consider the problem of controlling the system

$$x_{i+1} = Ax_i + Bu_i + v_i$$

$$z_i = Cx_i$$

Notice: measurement and output might be different. The case $C = I$ gives all state.

$$J_t = E\left\{\sum_{i=t}^{N} z_i^T W z_i + u_i^T Q_2 u_i\right\}$$

This is equivalent to minimizing

$$J_t = E\left\{\sum_{i=t}^{N} x_i^T Q_1 x_i + u_i^T Q_2 u_i\right\}$$

iff

$$\boxed{Q_1 = C^T W C}$$

Consider the problem of controlling the system

$$\begin{aligned} x_{i+1} &= Ax_i + Bu_i + v_i \\ z_i &= Cx_i + Du_i = \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} x_i \\ u_i \end{bmatrix} \end{aligned}$$

The cost function:

$$J_t = E\Big\{\sum_{i=t}^{N} z_i^T W z_i\Big\} = E\Big\{\sum_{i=t}^{N} \begin{bmatrix} x_i^T & u_i^T \end{bmatrix} \begin{bmatrix} C^T \\ D^T \end{bmatrix} W \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} x_i \\ u_i \end{bmatrix}\Big\}$$

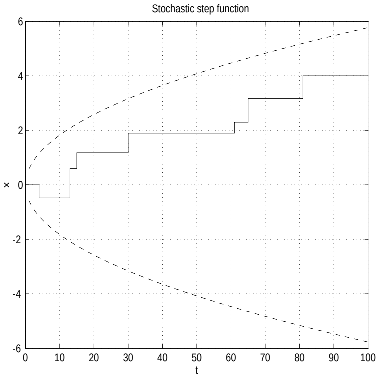
is equivalent to

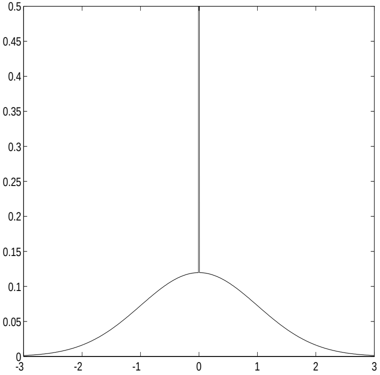
$$J_t = E\Big\{\sum_{i=t}^{N} x_i^T Q_1 x_i + u_i^T Q_2 u_i + 2x_i^T Q_{12} u_i\Big\} = E\Big\{\sum_{i=t}^{N} \begin{bmatrix} x_i^T & u_i^T \end{bmatrix} \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix} \begin{bmatrix} x_i \\ u_i \end{bmatrix}\Big\}$$

iff

$$\begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix} = \begin{bmatrix} C^T \\ D^T \end{bmatrix} W \begin{bmatrix} C & D \end{bmatrix}$$

Example - Set point (piece wise constant reference)





$$r_{t+1} = r_t + \eta_t \qquad \eta_t \in C(\ N(0, \sigma^2), (1-p)\delta,\ p)$$

Example - Set point

System:

$$x_{t+1} = ax_t + bu_t + v_t$$

Set point model:

$$r_{t+1} = r_t + \eta_t$$

Total state vector:

$$\bar{x}_t = \begin{bmatrix} x_t \\ r_t \end{bmatrix}$$

and description:

$$\begin{bmatrix} x \\ r \end{bmatrix}_{t+1} = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ r \end{bmatrix}_t + \begin{bmatrix} b \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ \eta_t \end{bmatrix}$$

Example - Set point

The objective is to minimize (in some sense)

$$\bar{J}_t = E\left\{\sum_{i=t}^{N}(x_i - r_i)^2 + \rho u_i^2\right\}$$

Let

$$x_i - r_i = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ r \end{bmatrix}_i = C_z \begin{bmatrix} x \\ r \end{bmatrix}_i$$

and the objective is the same as minimizing

$$\bar{J}_t = E\left\{\sum_{i=t}^{N}(x_i - r_i)^2 + \rho u_i^2\right\} = E\left\{\sum_{i=t}^{N} \begin{bmatrix} x_i & r_i \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} x_i \\ r_i \end{bmatrix} + \rho u_i^2\right\}$$

In the standard form this is equivalent to:

$$Q_1 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \qquad Q_2 = \rho$$

Example - Set point and integral action

System:

$$x_{t+1} = ax_t + bu_t + v_t$$

Set point model:

$$r_{t+1} = r_t + \eta_t$$

Integral state:

$$i_{t+1} = i_t + x_t - r_t$$

Total state vector:

$$\bar{x}_t = \begin{bmatrix} x \\ r \\ i \end{bmatrix}_t$$

and description:

$$\begin{bmatrix} x \\ r \\ i \end{bmatrix}_{t+1} = \begin{bmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ r \\ i \end{bmatrix}_t + \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ \eta_t \\ 0 \end{bmatrix}$$

The objective is to minimize (in some sense)

$$\bar{J}_{t_0} = E\left\{\sum_{t=t_0}^{N}(x_t - r_t)^2 + \rho_i i_t^2 + \rho_u u_t^2\right\}$$

In the standard form this is equivalent to:

$$Q_1 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \rho_i \end{bmatrix} \qquad Q_2 = \rho_u$$

Pause

The Standard Problem I

The stochastic system

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \perp v_s,\ x_s \text{ for } s \leq t$$

The objective function

$$J_t = \mathbf{E}\Big\{ \sum_{i=t}^{N} x_i^\top Q_1 x_i + u_i^\top Q_2 u_i \Big\}$$

The constraints

$$u_t = func(x_t),\ \ u_t = func(Y_t),\ \ u_t = func(Y_{t-1}),$$

Problems:

- Optimization (Quadratic)
- Stochastic (Gaussian)
- Dynamics (LTI)

DTU

$$J^* = \min_{u(x)} \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$

$$= x^\top S x$$

$$u^* = -h_{22}^{-1} h_{12}^\top x$$

$$S = h_{11} - h_{12} h_{22}^{-1} h_{12}^\top$$

Derivatives

Definition

Let

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

and

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} \in \mathbb{R}^m$$

Then

$$\frac{d}{dx} f(x) = \begin{bmatrix} \frac{df_1(x)}{dx_1} & \cdots & \frac{df_1(x)}{dx_n} \\ \vdots & \ddots & \vdots \\ \frac{df_m(x)}{dx_1} & \cdots & \frac{df_m(x)}{dx_n} \end{bmatrix}$$

Calculus

$$\frac{d}{dx}\left(y^T x\right) = y^T$$

$$\frac{d}{dx}\left(Ax\right) = A$$

$$\frac{d}{dx}\left(x^T Q x\right) = 2x^T Q$$

DTU

$$J = x^T h_{11} x + 2x^T h_{12} u + u^T h_{22} u$$

$$\frac{d}{du} J = 2x^T h_{12} + 2u^T h_{22} = 0$$

$$h_{12}^T x + h_{22} u = 0$$

$$u^* = -h_{22}^{-1} h_{12}^T x$$

$$\begin{aligned} J &= x^T h_{11} x + 2x^T h_{12} u + u^T h_{22} u && \textit{(Just a copy)} \\ &= x^T h_{11} x + (2x^T h_{12} + 2u^T h_{22}) u - u^T h_{22} u && \textit{(completion of the squares)} \end{aligned}$$

$$\begin{aligned} J^* &= x^T h_{11} x - u^T h_{22} u \\ &= x^T h_{11} x - x^T h_{12} h_{22}^{-1} h_{22} h_{22}^{-1} h_{12}^T x \\ &= x^T \left(h_{11} - h_{12} h_{22}^{-1} h_{12}^T \right) x \end{aligned}$$

Complete state information:

$$\min_{u(x)} \mathbf{E}\Big\{ I(x,u) \Big\} = \mathbf{E}\Big\{ \min_{u(x)} I(x,u) \Big\}$$

Incomplete state information:

$$\min_{u(y)} \mathbf{E}\Big\{ I(x,u) \Big\} = \mathbf{E}\Big\{ \min_{u(y)} \mathbf{E}\Big\{ \; I(x,u) | \; y \Big\} \Big\}$$

Theorem 2

$$x \,|y \; \in \mathbb{F}\left(\hat{x}, P\right)$$

$$\mathbf{E}\Big\{ x^\top S x \,|y \; \Big\} = \hat{x}^\top S \hat{x} + tr\left(SP\right)$$

Stochastic Q - optimization

$$J^* = \min_{u(y)} \mathbf{E}\left\{ \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \right\}$$

$$= \mathbf{E}\left\{ \min_{u(y)} \mathbf{E}\left\{ \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \middle| y \right\} \right\}$$

$x|y \in \mathbb{F}(\hat{x}, P)$

$$f(u,y) = \mathbf{E}\left\{ \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \middle| y \right\}$$

$$= \begin{bmatrix} \hat{x}^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} \hat{x} \\ u \end{bmatrix} + tr\left(h_{11} P\right)$$

Stochastic Q - optimization

$$f(u,y) \quad = \quad \begin{bmatrix} \hat{x}^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} \hat{x} \\ u \end{bmatrix} + tr\left(h_{11}P\right)$$

$$I^* = \min_{u(y)} f(u,y) = \hat{x}^\top S\hat{x} + tr(h_{11}P)$$

$$u^* = -h_{22}^{-1}h_{12}^\top\hat{x} \qquad\qquad u^* = -h_{22}^{-1}h_{12}^\top x$$

$$S = h_{11} - h_{12}h_{22}^{-1}h_{12}^\top$$

Stochastic optimization

- Optimization
- Stochasticity
- Dynamics $\rightarrow$ MPC, LQ(G)

Pause

Consider the discrete time system

$$x_{i+1} = Ax_i + Bu_i \qquad x_0 = \underline{x}_0 \tag{8}$$

$$y_i = Cx_i + Du_i \tag{9}$$

where the direct term (D) is include for the sake of generality.
A future state vector is a combination of the free response and the forced response, i.e.

$$x_i = A^i \underline{x}_0 + \begin{bmatrix} A^{i-1}B & \dots & AB & B \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_{i-2} \\ u_{i-1} \end{bmatrix}$$

or

$$x_i = A^i \underline{x}_0 + \mathcal{C}_{i-1} U_{i-1}$$

where

$$\mathcal{C}_i = \begin{bmatrix} A^i B & \dots & AB & B \end{bmatrix} \qquad U_i = \begin{bmatrix} u_0 \\ \vdots \\ u_{i-1} \\ u_i \end{bmatrix}$$

Furthermore, if we arrange the outputs into

$$Y_i = \begin{bmatrix} y_0 \\ \vdots \\ y_i \end{bmatrix}$$

and apply (9) then

$$\boxed{Y_i = \mathcal{O}_i x_0 + \mathbb{G}_i U_i}$$

where

$$\mathcal{O}_i = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^i \end{bmatrix} \qquad \mathbb{G}_i = \begin{bmatrix} D & 0 & 0 & \dots & 0 \\ CB & D & 0 & \dots & 0 \\ CAB & CB & D & \dots & 0 \\ CA^2B & CAB & CB & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ CA^{i-1}B & CA^{i-2}B & CA^{i-3}B & \dots & D \end{bmatrix}$$

Prediction (MPC)

$$\mathbb{G}_i = \begin{bmatrix} h_0 & 0 & 0 & \dots & 0 \\ h_1 & h_0 & 0 & \dots & 0 \\ h_2 & h_1 & h_0 & \dots & 0 \\ h_3 & h_2 & h_1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ h_i & h_{i-1} & h_{i-2} & \dots & h_0 \end{bmatrix}$$

$$h_i = \left\{ \begin{array}{lll} 0 & \text{for} & i < 0 \\ D & \text{for} & i = 0 \\ CA^{i-1}B & \text{for} & i > 0 \end{array} \right\}$$

DTU

```
function [G,O]=sspred(A,B,C,D,N)
% Determine the prediction matrices for the Model
% x_{t+1}=A*x+B*u
% y      =C*x+D*u
%Where the prediction is given by:
%
%        Y=O*x0 + G*U
%
% Y=[y_0, ... , y_N]^T
% U=[u_0, ... , u_N]^T
[n,m]=size(B);
p=size(C,1);
h=D; G=h;
nuls=zeros(p,m);
AnB=B;
for i=1:N,
 h=[C*AnB h];
 G=[G nuls; h];
 AnB=A*AnB;
 nuls=[nuls; zeros(p,m)];
end
O=C; CAn=C;
for i=1:N,
 CAn=CAn*A;
 O=[O; CAn];
end
```

Stochastic system

$$x_{i+1} = Ax_i + Bu_i + v_i$$
$$y_i = Cx_i + Du_i$$

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^k \end{bmatrix} x_0 + \begin{bmatrix} h_0 & 0 & \dots & 0 \\ h_1 & h_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h_k & h_{k-1} & \dots & h_0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_k \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ \tilde{h}_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{h}_k & \tilde{h}_{k-1} & \dots & 0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_k \end{bmatrix}$$

$$Y_k = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix} = \mathcal{O}x_0 + \mathbb{G}U_k + \tilde{\mathbb{G}}V_k$$

$$\hat{Y}_k = \mathcal{O}x_0 + \mathbb{G}U_k$$

MPC control

Scalar case

$$J = \mathbf{E}\Big\{ \sum_{i=0}^{N} y_i^2 + \rho u_i^2 \Big\}$$

Multivariable case

$$J = \mathbf{E}\Big\{ \sum_{i=0}^{N} y_i^T \mathcal{Q}_y y_i + u_i^T \mathcal{Q}_u u_i \Big\} \qquad Q = I_{N+1} \otimes \mathcal{Q}$$

$$J = \mathbf{E}\Big\{ \ \|Y\|_{Q_y}^2 + \|U\|_{Q_u}^2 \Big\} \qquad \|U\|_Q^2 = U^\top Q U$$

MPC 0

$$J = \mathbf{E}\left\{ \|Y\|^2_{Q_y} + \|U\|^2_{Q_u} \right\} \qquad \|U\|^2_Q = U^\top Q U$$

$$J^*_c = \min_{U(Y_i)} \mathbf{E}\left\{ \|Y\|^2_{Q_y} + \|U\|^2_{Q_u} \mid Y_i \right\} = \min_{U(Y_i)} \left\{ \underbrace{\|\hat{Y}\|^2_{Q_y} + \mathbf{tr}\left(P_y Q_y\right)}_{\text{From Theorem 2}} + \|U\|^2_{Q_u} \right\}$$

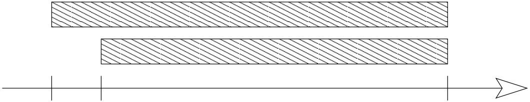
$$I = \left(\mathcal{O}\hat{x}_i + \mathbb{G}U_{i:k}\right)^\top Q_y \left(\mathcal{O}\hat{x}_i + \mathbb{G}U_{i:k}\right) + U^\top_{i:k} Q_u U_{i:k}$$

$$U_{i:k} = -\left[\ \mathbb{G}^\top Q_y \mathbb{G} + Q_u\ \right]^{-1} \mathbb{G}^\top Q_y \mathcal{O}\hat{x}_i$$

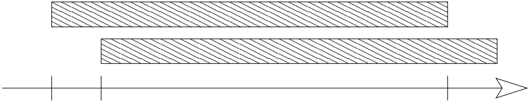
$$u_i = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} U_{i:k} \qquad = -L\hat{x}_i$$

DTU

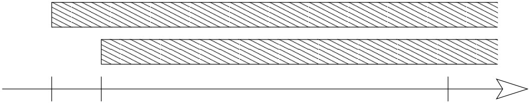
Shrinking horizon (Fixed end point)



Receding horizon



Inifinite horizon



Tricks of the trade

- Target points
- Feed forward
- Reference in cost function
- IMP
- Integral action
- control moves

Target point(set point)

$$x_{i+1} = Ax_i + Bu_i + d + v_i$$

$$y_i = Cx_i + Du_i + e_i = w$$

$$\begin{bmatrix} A - I & B \\ C & D \end{bmatrix} \begin{bmatrix} x_0 \\ u_0 \end{bmatrix} = \begin{bmatrix} -d \\ w \end{bmatrix}$$

$$u_i = u_0 - L(x - x_0)$$

Feedforward

$$u_i = Mw_i - Lx_i$$

Reference in cost function

$$J = \mathbf{E}\Big\{ \|Y - R\|^2_{Q_y} + \|U\|^2_{Q_u} \Big\}$$

$$U_{i:k} = \left[\mathbb{G}^\top Q_y \mathbb{G} + Q_u\right]^{-1} \mathbb{G}^\top Q_y \left[\hat{R}_{i:k} - \mathcal{O}\hat{x}_i\right]$$

$$u_i = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} U_{i:k}$$

$$u_i = M\hat{R}_{i:k} - L\hat{x}_i$$

DTU

Variation of the cost function

$$J = \mathbf{E}\Big\{ \|Y - R\|^2_{Q_y} + \|U\|^2_{Q_u} \Big\}$$

$$J = \mathbf{E}\Big\{ \sum_{k=i}^{N} (y_k - r_k)^2 + \rho u_k^2 \Big\}$$

$$J = \mathbf{E}\Big\{ \sum_{k=i}^{N} (y_k - r_k)^2 + \rho (u_k - \bar{u})^2 \Big\}$$

$$J = \mathbf{E}\Big\{ \sum_{k=i}^{N} (y_k - r_k)^2 + \rho \dot{u}_k^2 \Big\} \qquad \dot{u}_k \overset{\triangle}{=} u_k - u_{k-1}$$

$$J = \mathbf{E}\Big\{ \sum_{k=N_1}^{N_2} (y_k - r_k)^2 + \sum_{k=i}^{N_3} \rho u_k^2 \Big\}$$

$$u_k - u_{k-1} = 0 \quad \text{for} \quad k > N_3$$

DTU

End L9

The stochastic system

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \perp v_s,\ x_s \text{ for } s \leq t$$

The objective function

$$J_t = \mathbf{E}\Big\{ \sum_{i=t}^{N} x_i^\top Q_1 x_i + u_i^\top Q_2 u_i \Big\}$$

The constraints

$$u_t = func(x_t),\ \ u_t = func(Y_t),\ \ u_t = func(Y_{t-1}),$$

Problems:

- Optimization
- Stochastics
- Dynamics

Q - Optimization

$$\begin{aligned} J^* &= \min_{u(x)} \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\ &= \min_{u(x)} x^T h_{11} x + 2x^T h_{12} u + u^T h_{22} u \\ &= x^\top S x \end{aligned}$$

$$u^* = -h_{22}^{-1} h_{12}^\top x$$

$$S = h_{11} - h_{12} h_{22}^{-1} h_{12}^\top$$

DTU

Complete state information:

$$\min_{u(x)} \mathbf{E}\Big\{I(x,u)\Big\} = \mathbf{E}\Big\{\min_{u(x)} I(x,u)\Big\}$$

Incomplete state information:

$$\min_{u(y)} \mathbf{E}\Big\{I(x,u)\Big\} = \mathbf{E}\Big\{\min_{u(y)} \mathbf{E}\Big\{I(x,u)\ \Big|\ y\Big\}\Big\}$$

Theorem 2

$$x|y \in \mathbb{F}\left(\hat{x}, P\right)$$

$$\mathbf{E}\Big\{x^\top S x|y\Big\} = \hat{x}^\top S\hat{x} + tr\left(SP\right)$$

$$J^* = \min_{u(y)} \mathbf{E}\left\{ \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \right\}$$

$$x|y \in \mathbb{F}(\hat{x}, P)$$

$$f(u,y) = \begin{bmatrix} \hat{x}^\top & u^\top \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^\top & h_{22} \end{bmatrix} \begin{bmatrix} \hat{x} \\ u \end{bmatrix} + tr\left(h_{11}P\right)$$

$$I^* = \min_{u(y)} f(u,y) = \hat{x}^\top S \hat{x} + tr(h_{11}P)$$

$$u^* = -h_{22}^{-1} h_{12}^\top \hat{x} \qquad \left(u^* = -h_{22}^{-1} h_{12}^\top x \right)$$

$$S = h_{11} - h_{12} h_{22}^{-1} h_{12}^\top$$

$$x_{t+1} = f(x_t, u_t, v_t) = Ax_t + Bu_t + v_t$$

$$J_t = \min_{\{u_i\}_t^N} \mathbf{E}\left\{ \sum_{i=t}^{N} I_i(x_i, u_i) \right\} \qquad I_i(x_i, u_i) = x_i^T Q_1 x_i + u_i^T Q_2 u_i$$

$$u_t = func(\mathcal{F}_t) \qquad \mathcal{F}_t = x_t, \quad Y_t, \quad Y_{t-1}$$

$$V_t(\mathcal{F}_t) = \min_{\{u_i\}_t^N} \mathbf{E}\left\{ \sum_{i=t}^{N} I_i \;\middle|\; \mathcal{F}_t \right\}$$

Bellman function or optimal cost to go

$$V_t(\mathcal{F}_t) = \min_{u_t} \mathbf{E}\left\{ I_t + V_{t+1}(\mathcal{F}_{t+1}) \mid \mathcal{F}_t \right\} \qquad V_N(\mathcal{F}_N) = \min_{u_N} \mathbf{E}\left\{ I_N \mid \mathcal{F}_N \right\}$$

The Bellman equation - Proof

$$J_t = \min_{\{u_i\}_t^N} \mathbf{E}\left\{ \sum_{i=t}^N I_i \right\} = \mathbf{E}\left\{ \min_{\{u_i\}_t^N} \mathbf{E}\left\{ \sum_{i=t}^N I_i \middle| \mathcal{F}_t \right\} \right\} \qquad \{u_i\}_t^N = \{u_i(\mathcal{F}_i)\}_t^N$$

$$V_t(\mathcal{F}_t) = \min_{\{u_i\}_t^N} \mathbf{E}\left\{ \sum_{i=t}^N I_i \middle| \mathcal{F}_t \right\} = \min_{\{u_i\}_t^N} \mathbf{E}\left\{ I_t + \sum_{i=t+1}^N I_i \middle| \mathcal{F}_t \right\}$$

$$= \min_{u_t} \left[\mathbf{E}\left\{ I_t | \mathcal{F}_t \right\} + \min_{\{u_i\}_{t+1}^N} \mathbf{E}\left\{ \sum_{i=t+1}^N I_i \middle| \mathcal{F}_t \right\} \right]$$

$$= \min_{u_t} \left[\mathbf{E}\left\{ I_t + \min_{\{u_i\}_{t+1}^N} \mathbf{E}\left\{ \sum_{i=t+1}^N I_i \middle| \mathcal{F}_{t+1} \right\} \middle| \mathcal{F}_t \right\} \right]$$

$$V_t(\mathcal{F}_t) = \min_{u_t} \mathbf{E}\left\{ I_t + V_{t+1}(\mathcal{F}_{t+1}) \mid \mathcal{F}_t \right\} \qquad V_N(\mathcal{F}_N) = \min_{u_N} \mathbf{E}\left\{ I_N \mid \mathcal{F}_N \right\}$$

The Standard Problem - LQ control

The stochastic system

System, disturbances, reference (and frequency weights)

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \perp v_s,\ x_s \text{ for } s \leq t$$

The objective function

$$J_t = \mathbf{E}\left\{ \sum_{i=t}^{N} x_i^\top Q_1 x_i + u_i^\top Q_2 u_i \right\}$$

The constraints

$u_t = func(x_t)$ *Complete state information*

Notice: A, B, Q_1 and Q_2 might be time varying.

Let us start in the easy end - The end point at $t = N$

$$V_N(\mathcal{F}_N) = \min_{u_N} \mathbf{E}\left\{ \begin{bmatrix} x_N^\top & u_N^\top \end{bmatrix} \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \begin{bmatrix} x_N \\ u_N \end{bmatrix} \mid x_N \right\}$$

$$u_N = 0$$

$$V_N(x_N) = x_N^\top S_N x_N \quad S_N = Q_1$$

Notice, V_N is quadratic in x_N . We will use that as an inspiration for choosing a candidate function.

Now assume the optimization has been solved from N, $N-1$ and all the way to $t+1$.

Consider the following candidate function:

$$V_{t+1}(x_{t+1}) = x_{t+1}^\top S_{t+1} x_{t+1} + \beta_{t+1}$$

for the optimal cost to go.

Let us focus on $t < N$.

The Bellmann equation gives:

$$V_t(x_t) = \min_{u_t} \mathbf{E}\left\{ \begin{bmatrix} x_t^\top & u_t^\top \end{bmatrix} \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} + V_{t+1}(x_{t+1}) \middle| x_t \right\}$$

Let us first focus on $V_{t+1}(x_{t+1})$:

$$V_{t+1}(x_{t+1}) = x_{t+1}^\top S_{t+1} x_{t+1} + \beta_{t+1}$$

where

$$x_{t+1} = Ax_t + Bu_t + v_t = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} + v_t \qquad v_t \in \mathbf{N}(0, R_1)$$

$$V_{t+1}(x_{t+1}) = \begin{bmatrix} x_t^\top & u_t^\top \end{bmatrix} \begin{bmatrix} A^\top S_{t+1} A & A^\top S_{t+1} B \\ B^\top S_{t+1} A & B^\top S_{t+1} B \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} + v_t^\top S_{t+1} v_t$$

$$+ (Ax_t + Bu_t)^T S_{t+1} v_t + v_t^\top S_{t+1} (Ax_t + Bu_t) + \beta_{t+1}$$

$$V_t(x_t) = \min_{u_t} \mathbf{E}\left\{ \begin{bmatrix} x_t^\top & u_t^\top \end{bmatrix} \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} + V_{t+1}(x_{t+1}) \Big| x_t \right\}$$

$$V_t(x_t) = \min_{u_t} \left[\begin{bmatrix} x_t^\top & u_t^\top \end{bmatrix} \begin{bmatrix} A^\top S_{t+1} A + Q_1 & A^\top S_{t+1} B \\ B^\top S_{t+1} A & B^\top S_{t+1} B + Q_2 \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} \right.$$

$$\left. + \, tr\left(S_{t+1} R_1\right) + \beta_{t+1} \right]$$

$$\begin{aligned} u_t &= -\left[B^\top S_{t+1} B + Q_2\right]^{-1} B^\top S_{t+1} A x_t \\ &= -L_t \, x_t \end{aligned}$$

$$V_t(x_t) = x_t^\top S_t x_t + \beta_t$$

Hurra - it is quadratic

$$S_t = A^\top S_{t+1} A + Q_1 - A^\top S_{t+1} B \left[B^\top S_{t+1} B + Q_2\right]^{-1} B^\top S_{t+1} A$$

DTU

Pause

$$u_t = -L_t \, x_t$$

$$L_t = \left[B^\top S_{t+1} B + Q_2\right]^{-1} B^\top S_{t+1} A$$

$$S_t = A^\top S_{t+1} A + Q_1 - \; A^\top S_{t+1} B \, [B^\top S_{t+1} B + Q_2]^{-1} \, B^\top S_{t+1} A$$

$$S_{N+1} = 0$$

Riccati
Numerics
Deterministic sequence

Stationarity

Fixed horizon
Receding horizon
Stationary control

Closed loop 0

System:

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad z_t = Cx_t$$

Control:

$$u_t = -Lx_t$$

Closed loop:

$$x_{t+1} = \Big(A - BL\Big)x_t + v_t$$

Evolution of mean

$$m_{t+1} = \Big(A - BL\Big)m_t \qquad m_0 = \underline{m}_0$$

and variance:

$$\Sigma_{t+1} = \Big(A - BL\Big)\Sigma_t\Big(A - BL\Big)^T + R_1 \qquad \Sigma_0 = \underline{\Sigma}_0$$

$$u_t \in \mathbb{F}\Big(-L_t m_t, L_t \Sigma_t L_t^T\Big) \qquad z_t \in \mathbb{F}\Big(Cm_t, C\Sigma_t C^T\Big)$$

More Riccati

$$S_t = A^\top S_{t+1} A + Q_1 - L_t^T [B^\top S_{t+1} B + Q_2] L_t$$

Compact form:

$$S_t = (A - BL_t)^\top S_{t+1} + Q_1$$

Riccati equation:

$$S_t = A^\top S_{t+1} A + Q_1 - A^\top S_{t+1} B \left(B^\top S_{t+1} B + Q_2 \right)^{-1} B^\top S_{t+1} A$$

Any controller (Joseph stabilized version):

$$S_t = (A - BL_t)^T S_{t+1} (A - BL_t) + L_t^T Q_2 L_t + Q_1$$

$J_t = x_t^\top S_t x_t$ *in a deterministic setting*

Loss function

$$x_0 \in \mathbf{N}\left(m_0, P_0\right)$$

$$\begin{aligned} J \quad = \quad & m_0^\top S_0 m_0 \\ & +\mathbf{t}r\left(S_0 P_0\right) \\ & +\sum_{i=0}^{N} \mathbf{t}r\left(S_{t+1} R_1\right) \end{aligned}$$

Interpretation of S

If (A, B) is stabilizable (controlable) then $S_t \rightarrow S$ for $N \rightarrow \infty$.

Albegraic Riccati Equation (ARE):

$$S = A^\top SA + Q_1 - A^\top SB \left[B^\top SB + Q_2 \right] B^\top SA$$

$$L = \left[B^\top SB + Q_2 \right]^{-1} B^T SA$$

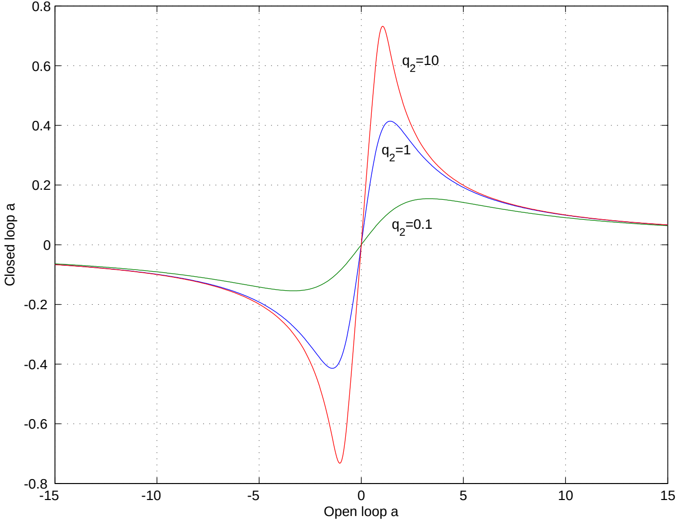
If additionally (A, Q_1) is observable then ARE have a unique, positive semidefinite solution and

$$(A - BL)$$

is asymptotic stable.

Stationarity $a = 0.98,\, b = 1,\, q_1 = 1$

$$x_{i+1} = ax_i + bu_i + v_i \qquad J = \mathbf{E}\Big\{ \sum_{i=0}^{\infty} q_1 x_i^2 + q_2 u_i^2 \Big\}$$



Pause

LQG Control

System and environment

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad\qquad v_t \perp v_s,\ x_s \text{ for } s \leq t$$

Measurements

$$y_t = Cx_t + e_t \qquad\qquad e_t \perp e_s,\ x_s \text{ for } s \leq t$$

Additional states might be needed in order to cope with non white measurements noise.

Cost:

$$J_t = \mathbf{E}\Big\{\sum_{i=t}^{N} x_i^\top Q_1 x_i + u_i^\top Q_2 u_i\Big\}$$

Restrictions:

$$u_t = func(Y_{t-1}), \qquad\qquad u_t = func(Y_t)$$

$$u_t = -L_t \hat{x}_{t|t-1} \qquad \hat{x}_{t|t-1} = \mathbf{E}\Big\{x_t | Y_{t-1}\Big\} \qquad\qquad u_t = -L_t \hat{x}_{t|t} \qquad \hat{x}_{t|t} = \mathbf{E}\Big\{x_t | Y_t\Big\}$$

LQG-control - Predictive version

$$u_t = -L_t \bar{x}_t \qquad \bar{x}_t = \hat{x}_{t|t-1} = \mathbf{E}\Big\{x_t | Y_{t-1}\Big\}$$

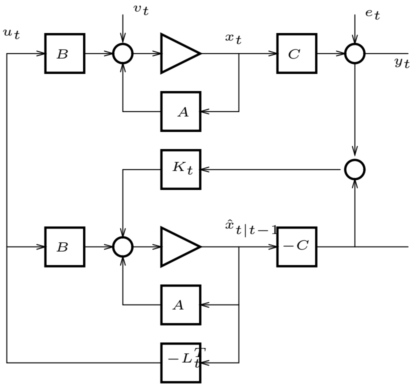
$$L_t = [B^\top S_{t+1} B + Q_2]^{-1} B^T S_{t+1} A$$

$$S_t = (A - BL_t)^\top S_{t+1} + Q_1$$

$$\bar{x}_{t+1} = A\bar{x} + Bu_t + K_t \left(y_t - C\bar{x}_t\right) \qquad K_t = A\kappa_t$$

---

1 Certainty equivalence: $\hat{x}_t$ is used instead of x_t as if it was correct.
2 Separation principle:
- L is given by the control problem and is the same as if x_t is known correctly.
- K is given by the estimation problem.



$$x_0 \in \mathbf{N}(m_0, P_0)$$

$$J \quad = \quad m_0^\top S_0 m_0 + \mathbf{t}r\,(S_0 P_0) + \sum_{i=0}^{N} \mathbf{t}r\,(S_{i+1} R_1) + \sum_{i=0}^{N} \mathbf{t}r\left(P_i L_i B^\top S_{i+1} A\right)$$

Duality

System and environment

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \perp v_s,\ x_s \text{ for } s \le t \qquad R_1$$
$$y_t = Cx_t + e_t \qquad e_t \perp e_s,\ x_s \text{ for } s \le t \qquad R_2$$

Cost function

$$J_t = \mathbf{E}\Big\{ \sum_t^N \begin{bmatrix} x_i^\top & u_i^\top \end{bmatrix} \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^\top & Q_2 \end{bmatrix} \begin{bmatrix} x_i \\ u_i \end{bmatrix} \Big\}$$

Constraints

$u_t = func(x_t) \qquad u_t = func(Y_t) \qquad u_t = func(Y_{t-1})$

Solution

$$u_t = -L_t x_t \qquad u_t = -L_t \hat{x}_{t|t} \qquad u_t = -L_t \hat{x}_{t|t-1}$$

$$L_t^T = [A^T S_{t+1} B + Q_{12}][B^\top S_{t+1} B + Q_2]^{-1}$$
$$K_t = [A P_t C^\top + R_{12}][C P_t C^\top + R_2]^{-1}$$

$$S_t = A^\top S_{t+1} A + Q_1 - L_t^T [B^\top S_{t+1} B + Q_2] L_t$$
$$P_{t+1} = A P_t A^\top + R_1 - K_t [C P_t C^\top + R_2] K_t^\top$$

$$x_{t+1} = Ax_t + Bu_t + v_t$$

$$y_t = C_y x_t + e_t \qquad\qquad z_t = C_z x_t$$

$$u_t = -L_t \bar{x}_t = -L_t x_t + L_t \tilde{x}_t = -\begin{bmatrix} L_t & -L_t \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{x}_t \end{bmatrix}$$

Closed loop:

$$x_{t+1} = (A - BL_t)\, x_t + BL_t \tilde{x}_t + v_t$$

Predictive Kalman filter

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K_t \left[y_t - C_y \bar{x} \right]$$

$$\tilde{x}_{t+1} = (A - K_t C_y)\, \tilde{x}_t + v_t - K_t e_t$$

$$\begin{bmatrix} x \\ \tilde{x} \end{bmatrix}_{t+1} = \begin{bmatrix} A - BL_t & BL_t \\ 0 & A - K_t C_y \end{bmatrix} \begin{bmatrix} x \\ \tilde{x} \end{bmatrix}_t + \begin{bmatrix} I & 0 \\ I & -K_t \end{bmatrix} \begin{bmatrix} v_t \\ e_t \end{bmatrix}$$

Closed loop I

$$\begin{bmatrix} x \\ \tilde{x} \end{bmatrix}_{t+1} = \underbrace{\begin{bmatrix} A - BL_t & BL_t \\ 0 & A - K_t C_y \end{bmatrix}}_{A_{cl}} \begin{bmatrix} x \\ \tilde{x} \end{bmatrix}_t + \underbrace{\begin{bmatrix} I & 0 \\ I & -K_t \end{bmatrix}}_{G} \begin{bmatrix} v_t \\ e_t \end{bmatrix}$$

$$m_{t+1} = A_{cl} m_t \to 0$$

iff asymtotic stable

$$\Sigma_{t+1} = A_{cl} \Sigma_t A_{cl}^\top + G \underline{R}_1 G^T \to \begin{bmatrix} P_x & P \\ P & P \end{bmatrix}$$

iff asymtotic stable

$$z_t = C_z x_t \in \mathbf{N}\left([C_z \ 0] m_t, \ [C_z \ 0] \Sigma_t \begin{bmatrix} C_z \\ 0 \end{bmatrix} \right) \qquad = \mathbf{N}\left(C_z m_x(t), \ C_z P_x(t) C_z \right)$$

$$u_t = -L_t \bar{x}_t = -L_t x_t + L_t \tilde{x}_t = -[L_t \ \ -L_t] \begin{bmatrix} x_t \\ \tilde{x}_t \end{bmatrix}$$

$$\in \mathbf{N}\left([L_t \ \ -L_t] m_t, \ [L_t \ \ -L_t] \Sigma_t \begin{bmatrix} L_t \\ -L_t \end{bmatrix} \right)$$

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \perp v_s,\ x_s \text{ for } s \le t$$
$$y_t = Cx_t + e_t \qquad e_t \perp e_s,\ x_s \text{ for } s \le t$$
$$z_t = C_z x_t$$

$$u_t = -L\hat{x}_t = -Lx_t + L\tilde{x}_t = -\begin{bmatrix} L & -L \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{x}_t \end{bmatrix}$$

Closed loop:

$$x_{t+1} = (A - BL)\, x_t + BL\tilde{x}_t + v_t$$

Ordinary Kalman filter

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t[y_t - C\hat{x}_{t|t-1}] \qquad \hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t$$

ie.

$$\tilde{x}_{t+1} = (I - \kappa_{t+1}C)(A\tilde{x}_t + v_t) - \kappa e_{t+1}$$

$$\begin{bmatrix} x \\ \tilde{x} \end{bmatrix}_{t+1} = \begin{bmatrix} A - BL_t & BL_t \\ 0 & A - \kappa_t CA \end{bmatrix} \begin{bmatrix} x \\ \tilde{x} \end{bmatrix}_t + \begin{bmatrix} I & 0 \\ I - \kappa_t C & -\kappa_t \end{bmatrix} \begin{bmatrix} v_t \\ e_{t+1} \end{bmatrix}$$

Closed loop II

$$\left[\begin{array}{c} x \\ \tilde{x} \end{array}\right]_{t+1} = \left[\begin{array}{cc} A - BL_t & BL_t \\ 0 & A - \kappa_t CA \end{array}\right] \left[\begin{array}{c} x \\ \tilde{x} \end{array}\right]_t + \left[\begin{array}{cc} I & 0 \\ I - \kappa_t C & -\kappa_t \end{array}\right] \left[\begin{array}{c} v_t \\ e_{t+1} \end{array}\right]$$

$$m_{t+1} = A_{cl} m_t \qquad \Sigma_{t+1} = A_{cl} \Sigma_t A_{cl}^\top + G^\top \underline{R}_1 G^\top$$

$$u_t = -L\hat{x}_t = -Lx_t + L\tilde{x}_t = -\left[\begin{array}{cc} L & -L \end{array}\right] \left[\begin{array}{c} x_t \\ \tilde{x}_t \end{array}\right]$$

$$z_t = C_z x_t \in \mathbf{N}\left([C_z\ 0] m_t,\ [C_z\ 0] \Sigma_t \left[\begin{array}{c} C_z \\ 0 \end{array}\right] \right) \qquad = \mathbf{N}\left(C_z m_x(t),\ C_z P_x(t) C_z \right)$$

DTU

Wind turbine control

System and wind disturbance:

$$\begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \\ v \\ \dot{v} \end{bmatrix}_{t+1} = A \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \\ v \\ \dot{v} \end{bmatrix}_t + Bu_t + v_t \quad R_1$$

$$z_t = C_z x_t \qquad (P_e, \textit{ an output, not a measurement}) \qquad y_t = C_y x_t + e_t$$

Set point model:

$$r_{t+1} = r_t + \eta_t$$

Integral action:

$$i_{t+1} = i_t + (r_t - C_z x_t)$$

Total model:

$$\begin{bmatrix} x \\ r \\ i \end{bmatrix}_{t+1} = \begin{bmatrix} A & 0 & 0 \\ 0 & 1 & 0 \\ -C_z & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ r \\ i \end{bmatrix}_t + \begin{bmatrix} B \\ 0 \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ \eta_t \\ 0 \end{bmatrix}$$

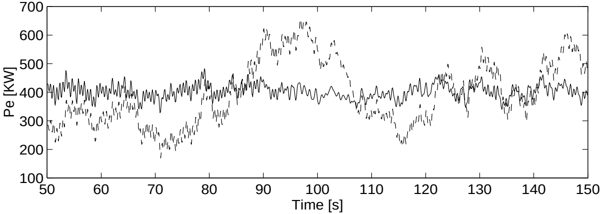
Objective:

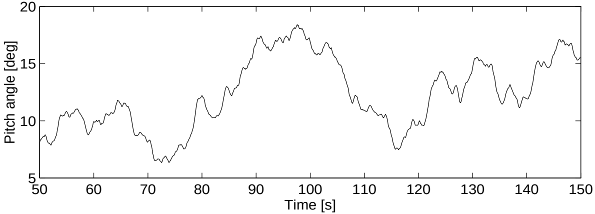
$$r_t - y_t \simeq 0 \quad i_t \simeq 0 \qquad u_t \simeq 0$$

$$J = \mathbf{E}\left\{\sum_{i=0}^{\infty}(r_i - y_i)^2 + \rho_z z_i^2 + \rho_u u^2\right\}$$

$$J = \mathbf{E}\left\{\sum_{i=0}^{\infty}\underline{x}_i^\top Q_1 \underline{x}_i + u_i^\top Q_2 u_i\right\}$$

$$u_t = -L\begin{bmatrix}\hat{x}_t \\ r_t \\ z_t\end{bmatrix}$$





DTU

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

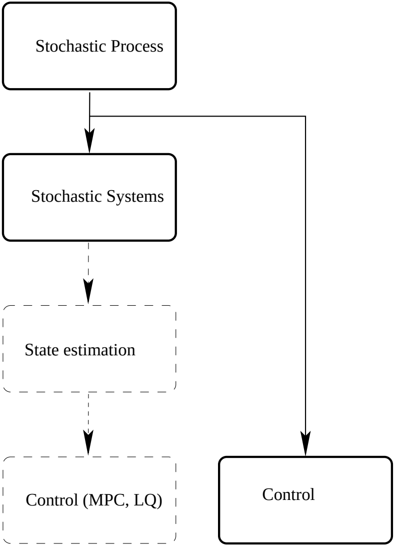
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

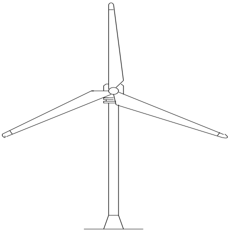
Processes II (L11-12)

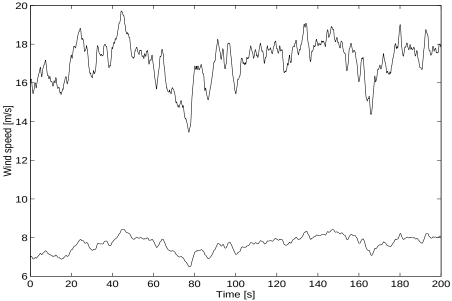
Stochastic Processes II



L12

Stochastic Processes II





Covariance function - Scalar case

(Auto) covariance function

$$r_x(s,t) = \mathbf{E}\left\{\tilde{x}_s\tilde{x}_t\right\} \qquad \tilde{x}_t = x_t - m_t \in \mathbb{R}$$

Cross covariance function

$$r_{xy}(s,t) = \mathbf{E}\left\{\tilde{x}_s\tilde{y}_t\right\}$$

Weakly (wide sense) stationary processes

$$r_x(k) = \mathbf{E}\left\{\tilde{x}_{t+k}\tilde{x}_t\right\} = \mathbf{E}\left\{\tilde{x}_t\tilde{x}_{t-k}\right\} \qquad r_{xy}(k) = \mathbf{E}\left\{\tilde{x}_{t+k}\tilde{y}_t\right\} = \mathbf{E}\left\{\tilde{x}_t\tilde{y}_{t-k}\right\}$$

$$r_x(k) = r_x(-k) \qquad r_{xy}(k) = r_{yx}(-k)$$

z_t	$r_z(k)$	$r_{zx}(k)$
$x_t + y_t$	$r_x(k) + r_y(k) + r_{xy}(k) + r_{xy}(-k)$	$r_x(k) + r_{xy}(-k)$
ax_t	$a^2 r_x(k)$	$ar_x(k)$

Spec cross

DTU

Covariance function - Vector case

(Auto) covariance function

$$r_x(s,t) = \mathbf{E}\left\{\tilde{x}_s \tilde{x}_t^\top\right\} \qquad n_x \times n_x \qquad \tilde{x}_t = x_t - m_t \in \mathbb{R}^{n_x}$$

Cross covariance function

$$r_{xy}(s,t) = \mathbf{E}\left\{\tilde{x}_s \tilde{y}_t^\top\right\} \qquad n_x \times n_y$$

Weakly stationary processes

$$r_x(k) = \mathbf{E}\left\{\tilde{x}_{t+k} \tilde{x}_t^\top\right\} = \mathbf{E}\left\{\tilde{x}_t \tilde{x}_{t-k}^\top\right\} \qquad r_{xy}(k) = \mathbf{E}\left\{\tilde{x}_{t+k} \tilde{y}_t^\top\right\} = \mathbf{E}\left\{\tilde{x}_t \tilde{y}_{t-k}^\top\right\}$$

$$r_x(k) = r_x^\top(-k) \qquad r_{xy}(k) = r_{yx}^\top(-k)$$

---

z_t	$r_z(k)$	$r_{zx}(k)$
$x_t + y_t$	$r_x(k) + r_y(k) + r_{xy}(k) + r_{xy}^T(-k)$	$r_x(k) + r_{xy}^T(-k)$
Ax_t	$A\ r_x(k)\ A^T$	$Ar_x(k)$

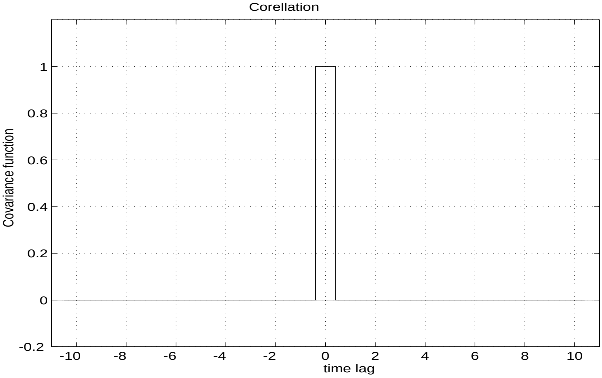
DTU

External Process models

Building block (in discrete time): White noise

$$e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$$r_e(k) = \mathbf{E}\Big\{e_t e_{t-k}\Big\} = \delta_k \sigma^2$$



External Process models

Here we will focus on:

From Model to Covariance, spectral properties.

Analysis

Later on we will look at:

From Covariance to Model

System identification

External models, ARMA process

$$A(q^{-1})y_t = C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right) \qquad r_{ey}(k) = 0 \quad \text{for} \quad k > 0$$

$$y_t + a_1 y_{t-1} + \ \dots \ + a_n y_{t-n} = e_t + c_1 e_{t-1} + \ \dots \ + c_n e_{t-n}$$

$$y_t = H(q)e_t = \frac{C(q^{-1})}{A(q^{-1})} e_t = \sum_{i=0}^{\infty} h_i e_{t-i} = h_t \star e_t$$

ARMA

$$A(q^{-1})y_t = C(q^{-1})e_t$$

MA

Also called a FIR or a zeros only model.

$$y_t = C(q^{-1})e_t$$

AR

Also called a IIR or a poles only model.

$$A(q^{-1})y_t = e_t$$

ARMA - Mean value

We assume A to have its roots inside the unit disk.

$$m_y = \frac{C(1)}{A(1)}m_e$$

$$m_y(t) = \frac{C(q^{-1})}{A(q^{-1})}m_e(t)$$

ARMA - Mean value

In the following we assume:

$$m_e = 0, \qquad m_y = 0$$

in order to reduce the complexity.

Basic (standard) assumption

The noise, e_t, is assumed to be independent of the old output history, i.e.

$$r_{ey}(k) = 0 \quad \text{for} \quad k > 0$$

Causality

For causal dynamic system

$$y_t = func(e_t, e_{t-1}, \ldots)$$

That results in:

$$r_{ye}(k) = 0 \quad \text{for} \quad k < 0$$

Moving Average (MA), FIR model, zeros only model

$$
\begin{aligned}
y_t &= C(q^{-1})e_t = \left(1 + c_1 q^{-1} + \ \dots\ + c_n q^{-n}\right) e_t \\
&= e_t + c_1 e_{t-1} + \ \dots\ + c_n e_{t-n} = \sum_{k=0}^{n} c_k e_{t-k} = h_t \star e_t
\end{aligned}
$$

$e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$

$r_{ey}(k) = 0$ for $k > 0$

$$
h_k = \begin{cases} c_k & \text{for } 0 \le k \le n \\ 0 & \text{otherwise} \end{cases} \qquad (c_0 = 1)
$$

$$
y_t \color{red}{e_{t-k}} = e_t \color{red}{e_{t-k}} + c_1 e_{t-1} \color{red}{e_{t-k}} + \ \dots\ + c_n e_{t-n} \color{red}{e_{t-k}}
$$

$$
r_{ye}(k) = \sigma^2 \begin{cases} c_k & \text{for } 0 \le k \le n \\ 0 & \text{otherwise} \end{cases} \qquad r_{ye}(k) = \sigma^2 h_k
$$

The MA-process

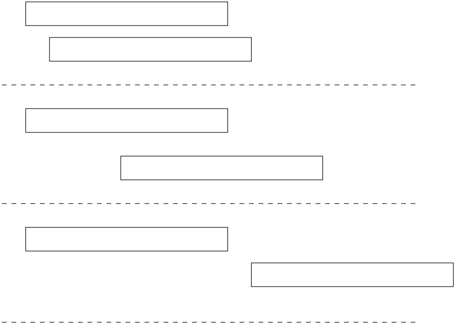
$$\begin{aligned} y_t &= C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right) \\ &= e_t + c_1 e_{t-1} + \ \dots \ + c_n e_{t-n} \end{aligned}$$

$$\begin{aligned} r_y(0) &= \mathbf{E}\Big\{\left(e_t + c_1 e_{t-1} + \ \dots \ + c_n e_{t-n}\right)^2\Big\} \\ &= \left(1 + c_1^2 + \ \dots \ + c_n^2\right)\sigma^2 \end{aligned}$$

$$\begin{aligned} r_y(1) = \mathbf{E}\Big\{y_t y_{t-1}\Big\} &= \mathbf{E}\Big\{\left(e_t + c_1 e_{t-1} + c_2 e_{t-2} \dots \ + c_n e_{t-n}\right) \times \\ &\qquad \left(e_{t-1} + c_1 e_{t-2} + \ \dots \ + c_{n-1} e_{t-n} + c_n e_{t-n-1}\right)\Big\} \\ &= \left(c_1 \ + \ c_2 c_1 + \ \dots \ + c_n c_{n-1}\right)\sigma^2 \end{aligned}$$

DTU

The MA-process



$r_y(k) = 0$ for $k > n$

$$r_y(k) = \Big[c_k c_0 + c_{k+1} c_1 + \ldots + c_n c_{n-k}\Big]\sigma^2 \qquad \text{for } 1 \leq k \leq n \qquad (0 \leq k \leq n)$$

The MA-process

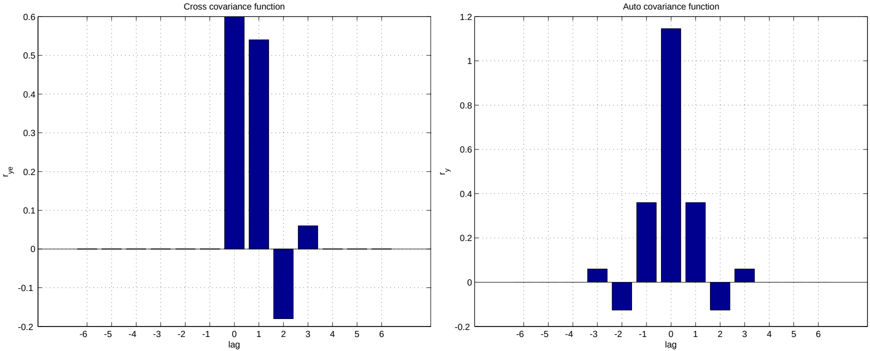
$$y_t = C(q^{-1})e_t$$

$$r_{ye}(k) = \sigma^2 \begin{cases} c_k & \text{for } 0 \leq k \leq n \\ 0 & \text{otherwise} \end{cases} = \sigma^2 h_k$$

$$r_y(k) = \sigma^2 \begin{cases} 1 + c_1^2 + c_2^2 + ... + c_n^2 & \text{for } k = 0 \\ c_{|k|} + c_{|k|+1}c_1 + ... + c_n c_{n-|k|} & \text{for } 1 \leq |k| \leq n \\ 0 & \text{otherwise} \end{cases}$$

MA-process

$$y_t = \left(1 + 0.9q^{-1} - 0.3q^{-2} + 0.1q^{-3}\right) e_t \qquad e_t \in \mathbf{N}_{iid}(0, 0.6)$$



Pause

Auto Regressive Moving Average, IIR model

Stochastic Process Model:

$$A(q^{-1})y_t = C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right) \qquad r_{ey}(k) = 0 \text{ for } k > 0$$

$$y_t + a_1 y_{t-1} + \ \dots \ + a_n y_{t-n} = e_t + c_1 e_{t-1} + \ \dots \ + c_n e_{t-n}$$

$$y_t = H(q)e_t = \frac{C(q^{-1})}{A(q^{-1})} e_t = \sum_{i=0}^{\infty} h_i e_{t-i} = h_t \star e_t = \sum_{i=0}^{\infty} h_i q^{-i} \ e_t$$

Necessary condition for stationarity: $A(z^{-1}) \neq 0$ for $|z| \geq 1$.

Notice, causality is ensured.

The ARMA process

$$y_t + a_1 y_{t-1} + \ldots + a_n y_{t-n} = e_t + c_1 e_{t-1} + \ldots + c_n e_{t-n}$$

$r_{ey}(i) = 0$ for $i > 0$

Multiply with e_{t-k} and perform the expectation.

$$y_t e_{t-k} + a_1 y_{t-1} e_{t-k} + \ldots + a_n y_{t-n} e_{t-k} = e_t e_{t-k} + c_1 e_{t-1} e_{t-k} + \ldots + c_n e_{t-n} e_{t-k}$$

$$r_{ye}(k) + a_1 r_{ye}(k-1) + \ldots + a_n r_{ye}(k-n) = r_e(k) + c_1 r_e(k-1) + \ldots + c_n r_e(k-n)$$

Results:

$$A(q^{-1}) r_{ye}(k) = C(q^{-1}) \delta_k \sigma^2 \qquad r_{ye}(k) = 0 \text{ for } k < 0$$

or

$$r_{ye}(k) = h_k \sigma^2 = \begin{cases} 0 & \text{for } k < 0 \\ h_k \sigma^2 & \text{for } k \geq 0 \end{cases}$$

The ARMA process

▸ cross ▸ example

k	-2	-1	0	1	2	3	...
$r_{ye}(k)$	0	0	$h_0\sigma^2$	$h_1\sigma^2$	$h_2\sigma^2$	$h_3\sigma^2$	...

k	2	1	0	-1	-2	-3	...
$r_{ey}(k)$	0	0	$h_0\sigma^2$	$h_1\sigma^2$	$h_2\sigma^2$	$h_3\sigma^2$	...

k	...	-3	-2	-1	0	1	2
$r_{ey}(k)$	...	$h_3\sigma^2$	$h_2\sigma^2$	$h_1\sigma^2$	$h_0\sigma^2$	0	0

$$r_{ey}(k) = \sigma^2 \hbar_k \quad \hbar_k = h_{-k}$$

$r_{ey}(k) = 0$ for $k > 0$ (we actually assumed this)

The ARMA process

$$y_t + a_1 y_{t-1} + \dots + a_n y_{t-n} = e_t + c_1 e_{t-1} + \dots + c_n e_{t-n}$$

Multiply with y_{t-k} and perform the expectation.

$$y_t y_{t-k} + a_1 y_{t-1} y_{t-k} + \dots + a_n y_{t-n} y_{t-k} = e_t y_{t-k} + c_1 e_{t-1} y_{t-k} + \dots + c_n e_{t-n} y_{t-k}$$

$$r_y(k) + a_1 r_y(k-1) + \dots + a_n r_y(k-n) = r_{ey}(k) + c_1 r_{ey}(k-1) + \dots + c_n r_{ey}(k-n)$$

Results in: **The Yule-Walker equation:**

$$A(q^{-1}) r_y(k) = C(q^{-1}) r_{ey}(k)$$

$$r_{ey}(k) = \sigma^2 \hbar_k \quad \hbar_k = h_{-k}$$

$$r_y(k) = \sigma^2 h_k \star \hbar_k$$

For a first order process the Yule-Walker

$$A(q^{-1})r_y(k) = C(q^{-1})r_{ey}(k)$$

becomes:

$$r_y(k) + a_1 r_y(k-1) = r_{ey}(k) + c_1 r_{ey}(k-1)$$

$k = 0$

$$r_y(0) + a_1 r_y(-1) = r_{ey}(0) + c_1 r_{ey}(-1) = \sigma^2(h_0 + c_1 h_1)$$

cross

$k = 1$

$$r_y(1) + a_1 r_y(0) = r_{ey}(1) + c_1 r_{ey}(0) = \sigma^2 c_1 h_0$$

$k = 2$

$$r_y(2) + a_1 r_y(1) = r_{ey}(2) + c_1 r_{ey}(1) = 0$$

$k > n \;\; = 1$ in our case.

$$r_y(k) + a_1 r_y(k-1) = 0$$

or

$$r_y(k) = -a_1 r_y(k-1) \qquad \text{for} \qquad k > n$$

The symmetric $r_y(k), \;\; k > 1$ can be determined recursively if $r_y(1)$ is known.

ARMA process

But:

$$r_y(0) + a_1 r_y(-1) = \sigma^2(h_0 + c_1 h_1)$$

$$r_y(1) + a_1 r_y(0) = \sigma^2 c_1 h_0$$

can also be written as:

$$\begin{bmatrix} 1 & a_1 \\ a_1 & 1 \end{bmatrix} \begin{bmatrix} r_y(0) \\ r_y(1) \end{bmatrix} = \sigma^2 \begin{bmatrix} h_0 + c_1 h_1 \\ c_1 h_0 \end{bmatrix}$$

from which $r_y(0)$ and $r_y(1)$ can be found.

In general:

$$\begin{bmatrix} 1 & \dots & a_n \\ \vdots & \ddots & \vdots \\ a_n & \dots & 1 \end{bmatrix} \begin{bmatrix} r_y(0) \\ \vdots \\ r_y(n) \end{bmatrix} = \sigma^2 \begin{bmatrix} \vdots \end{bmatrix}$$

and

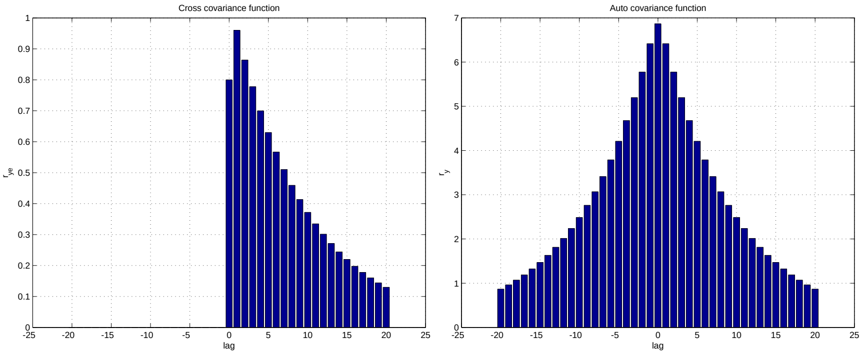
$$r_y(k) = -a_1 r_y(k-1) \dots - a_n r_y(k-n) \qquad \text{for} \qquad k > n$$

$$\begin{bmatrix} A_1 + A_2 \end{bmatrix} r = hs$$

$$A_1 = \begin{bmatrix} 1 & a_1 & a_2 & a_3 & a_4 \\ a_1 & a_2 & a_3 & a_4 & 0 \\ a_2 & a_3 & a_4 & 0 & 0 \\ a_3 & a_4 & 0 & 0 & 0 \\ a_4 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & a_1 & 1 & 0 & 0 \\ 0 & a_2 & a_1 & 1 & 0 \\ 0 & a_3 & a_2 & a_1 & 1 \end{bmatrix}$$

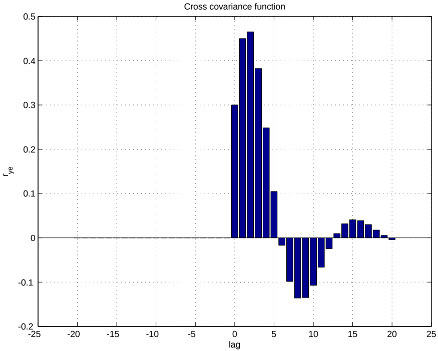
First order example

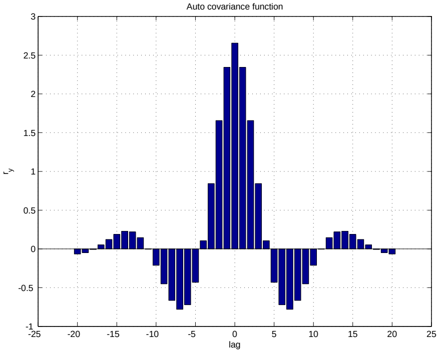
$$(1 - 0.9q^{-1})y_t = (1 + 0.3q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}(0, 0.8)$$



Second order example

$$(1 - 1.5q^{-1} + 0.7q^{-2})y_t = e_t \qquad e_t \in \mathbf{N}_{iid}(0, 0.3)$$





$$y(t_c) = H(\frac{d}{dt_c})e(t_c)$$

$$H(s) = \frac{s^m + c_1 s^{m-1} + \ \dots \ + c_m}{s^n + a_1 s^{n-1} + \ \dots \ + a_n} \quad n > m$$

$$r_{ye}(t_c) = \sigma^2 h_{t_c} \qquad \hbar_k = h_{-t_c}$$

$$r_y(t_c) = \sigma^2 h_{t_c} \star \hbar_{t_c}$$

Wind model

$$H_w(s) = \frac{k}{(1 + p_1 s)(1 + p_2 s)}$$

Pause

Spectrum

$$\Psi_{xy}(z) = \mathcal{Z}_b[r_{xy}(k)] = \sum_{k=-\infty}^{\infty} r_{xy}(k) z^{-k}$$

$$\Psi_x(z^{-1}) = \Psi_x(z) \quad \text{Symmetry}$$

Spectral density - Auto

$$\phi_x(\omega) = \Psi_x(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-j\omega k} = \mathcal{F}[r_x(k)] \qquad \omega \in [-\pi;\ \pi]$$

$$\phi_x(\omega) \in \mathbb{R} \qquad \phi_x(\omega) \geq 0 \qquad \phi_x(-\omega) = \phi_x(\omega)$$

$$r_x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\omega) e^{j\omega k} d\omega \qquad r_x(0) = \text{Var}\{x_t\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\omega) d\omega$$

ζ_t	$\Psi_\zeta(z)$
$x_t + y_t$	$\Psi_x(z) + \Psi_y(z) + \Psi_{xy}(z) + \Psi_{xy}(z^{-1})$
$a x_t$	$a^2 \Psi_x(z)$

CoVar

Spectral properties - Vector case

Spectrum

$$\Psi_{xy}(z) = \mathcal{Z}_b[r_{xy}(k)] = \sum_{k=-\infty}^{\infty} r_{xy}(k) z^{-k} \qquad \Psi_x(z^{-1}) = \Psi_x^{\top}(z)$$

Spectral density

$$\phi(\omega) = \Psi(e^{j\omega}) = \sum_{k=-\infty}^{\infty} r_x(k) e^{-j\omega k} = \mathcal{F}[r_x(k)] \qquad \omega \in [-\pi;\ \pi]$$

$$\phi(\omega) \in \mathbb{C}^{n\times n} \qquad \phi(\omega) \geq 0 \qquad \phi(-\omega) = \bar{\phi}^T(\omega)$$

$$r_x(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(\omega) e^{j\omega k} d\omega \qquad r_x(0) = \text{Var}\{x_t\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(\omega) d\omega$$

ζ_t	$\Psi_\zeta(z)$
$x_t + y_t$	$\Psi_x(z) + \Psi_y(z) + \Psi_{xy}(z) + \Psi_{xy}^T(z^{-1})$
Φx_t	$\Phi \Psi_x(z) \Phi^T$

DTU

Example

Covariance function

k	...	-3	-2	-1	0	1	2	3	...
$r_x(k)$		0.5	1	1.5	2	1.5	1	0.5	

Spectrum

$$\Psi(z) = \sum_{k=-\infty}^{\infty} r_x(k) z^{-k} = 2 + 1.5\left(z + z^{-1}\right) + 1\left(z^2 + z^{-2}\right) + 0.5\left(z^3 + z^{-3}\right) + \ldots$$

Spectral density

$$\phi(\omega) = \Psi(e^{j\omega}) = 2 + 3cos(\omega) + 2cos(2\omega) + 1cos(3\omega) + \ldots \qquad \omega \in [-\pi;\ \pi]$$

$$\phi(\omega) \in \mathbb{R} \qquad \phi(\omega) \geq 0 \qquad \phi(-\omega) = \phi(\omega)$$

$$e^{j\omega} = cos(\omega) + j\ sin(\omega) \qquad e^{jk\omega} = cos(k\omega) + j\ sin(k\omega)$$

Frequency domains

Normalized angular frequency

$$\omega \in [-\pi, \pi] \qquad \omega = \bar{\omega} T_s$$

Angular frequency

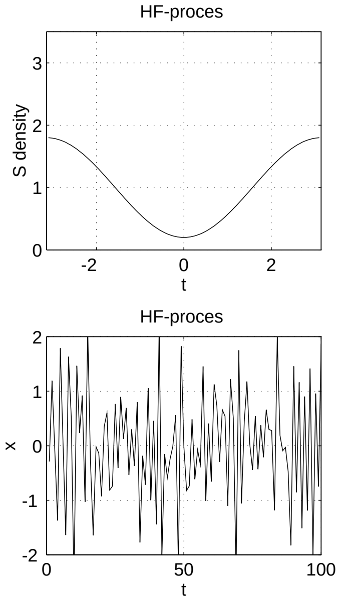
$$\bar{\omega} \in [-\frac{\pi}{T_s}, \frac{\pi}{T_s}] \qquad \bar{\omega} = 2\pi f$$

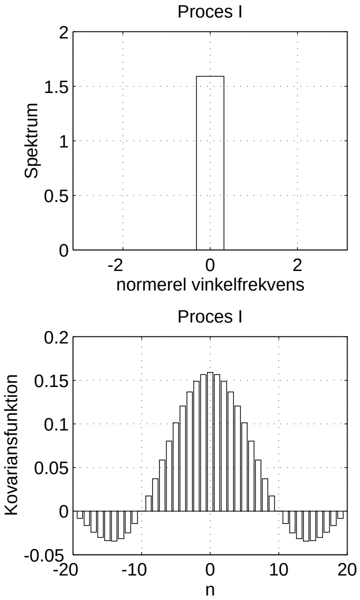
Frequency

$$f \in [-\frac{1}{2T_s}, \frac{1}{2T_s}] = [-\frac{1}{2} f_s, \frac{1}{2} f_s]$$

▸ L12

DTU


HF-proces
LF-proces
S density
x
t
HF-proces
LF-proces


Proces I
Proces II
Spektrum
normerel vinkelfrekvens
Kovariansfunktion
n
2
1.5
1
0.5
0
-2
0
2
0.2
0.15
0.1
0.05
0
-0.05
-20
-10
0
10
20

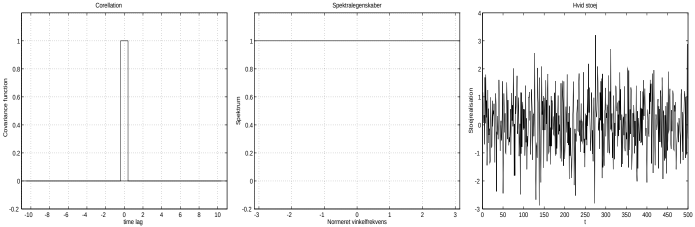
White noise

From optics

$$\phi(\omega) = c \qquad\qquad r_k = \begin{cases} c & \text{for } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

When and why

$$e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$



Bits from the Z-transform

$$\mathcal{Z}_b[h_k] = \sum_{-\infty}^{\infty} h_i z^{-i} = \sum_{0}^{\infty} h_i z^{-i} = H(z) = \frac{C(z^{-1})}{A(z^{-1})}$$

Reversing the time:

$$\mathcal{Z}_b[h_{-k}] = \sum_{-\infty}^{\infty} h_{-i} z^{-i} = \sum_{-\infty}^{\infty} h_i z^{i} = H(z^{-1})$$

Convolution:

$$\mathcal{Z}_b[h_t * u_t] = H(z)U(z)$$

$$r_y(k) = \sigma^2 h_k \star \hbar_k \qquad r_{ye}(k) = h_k \sigma^2$$

$$\Psi_y(z) = H(z)H(z^{-1})\sigma^2 \qquad \Psi_{ye}(z) = H(z)\sigma^2$$

$$\phi_y(\omega) = H(e^{j\omega})H(e^{-j\omega})\sigma^2 = |H(e^{j\omega})|^2\sigma^2 \in \mathbb{R} \qquad \phi_{ye}(\omega) = H(e^{j\omega})\sigma^2 \in \mathbb{C}$$

C-time process

$$\Psi_y(s) = H(s)H(-s)\sigma^2 \qquad \Psi_{ye}(z) = H(s)\sigma^2$$

$$\phi_y(\omega) = H(j\omega)H(-j\omega)\sigma^2 = |H(j\omega)|^2\sigma^2 \in \mathbb{R} \qquad \phi_{ye}(\omega) = H(j\omega)\sigma^2 \in \mathbb{C}$$

End L11

Recap L11

Building block (in discrete time):

$$e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right) \qquad r_e(k) = \mathbf{E}\left\{e_t e_{t-k}\right\} = \delta_k \sigma^2$$

Stochastic Process Model

$$y_t = H(q)e_t = \frac{C(q^{-1})}{A(q^{-1})}e_t = \sum_{i=0}^{\infty} h_i e_{t-i} = h_t \star e_t \qquad r_{ey}(k) = 0 \text{ for } k > 0$$

$(A, C, \sigma^2) \rightarrow r_y(k), r_{ye}(k)$ **(Yule-Walker)**

$$r_{ye}(k) = \sigma^2 h_k \qquad r_y(k) = \sigma^2 h_k \star \hbar_k$$

$$\phi_{ye}(\omega) = H(e^{j\omega})\sigma^2 \in \mathbb{C} \qquad \phi_y(\omega) = H(e^{j\omega})H(e^{-j\omega})\sigma^2 \in \mathbb{R}$$

The same (in some sense) goes for Continuous time and in the vector case.

Learning objectives

- Process models (external)
  - Spectral properties
  - Spectral factorization
  - The representations theorem
- Stochastic systems (external)

Sepc example

DTU

Example 1

$$y_t = H(q)e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$$H(q) = \frac{2 - q^{-1}}{1 - 0.8q^{-1}} \qquad \sigma^2 = 1$$

$$\Psi_y(z) = H(z)H(z^{-1})\sigma^2 = \frac{2 - z^{-1}}{1 - 0.8z^{-1}} \times \frac{2 - z}{1 - 0.8z} \times 1 = \frac{-2z^{-1} + 5 - 2z}{-0.8z^{-1} + 1.64 - 0.8z}$$

$$\phi(\omega) = \Psi(e^{j\omega}) = \frac{5 - 4\,cos(\omega)}{1.64 - 1.6\,cos(\omega)}$$

$$e^{\sigma + j\theta} = e^{\sigma}\left(cos(\theta) + jsin(\theta)\right) \qquad\qquad z \;\rightarrow\; e^{j\omega}$$

$$(z^n + z^{-n}) \;\rightarrow\; cos(n\omega) + jsin(n\omega) + cos(-n\omega) + jsin(-n\omega) = 2cos(n\omega)$$

DTU

Generally - Specifically

$$H(z) = \frac{b_0 + b_1 z^{-1} + ... + b_{n_b}}{1 + a_1 z^{-1} + ... + a_{n_a}}$$

$$\Psi_y(z) \frac{-2z^{-1} + 5 - 2z}{-0.8z^{-1} + 1.64 - 0.8z}$$

$$\phi(\omega) = \frac{5 - 4\, cos(\omega)}{1.64 - 1.6\, cos(\omega)}$$

Spectrum

$$\psi(z) = \frac{\bar{b}_0 + \sum_{i=1}^{n_b} \bar{b}_i (z^i + z^{-i})}{\bar{a}_0 + \sum_{i=1}^{n_a} \bar{a}_i (z^i + z^{-i})}$$

Spectral density

$$\phi(\omega) = \frac{\bar{b}_0 + \sum_{i=1}^{n_b} 2\bar{b}_i \cos(i\omega)}{\bar{a}_0 + \sum_{i=1}^{n_a} 2\bar{a}_i \cos(i\omega)}$$

$$(b_0 + b_1 z^{-1} + b_2 z^{-2})(b_0 + b_1 z^1 + b_2 z^2) = (b_0^2 + b_1^2 + b_2^2) + (z^1 + z^{-1})(b_1 b_0 + b_2 b_1) + (z^2 + z^{-2}) b_2 b_0$$

$$\bar{b}_i = \sum_{j=i}^{n_b} b_j b_{j-i} \qquad \bar{a}_i = \sum_{j=i}^{n_b} a_j a_{j-i}$$

spec.m, spec2sd.m

DTU

C-time

$$y_t = H(\frac{d}{dt})e_t$$

$$H(s) = \frac{1}{1+s\tau} \qquad \sigma^2 = 1$$

$$\psi(s) = H(s)H(-s)\sigma^2 = \frac{1}{1+s\tau} \times \frac{1}{1-s\tau} = \frac{1}{1-s^2\tau^2}$$

$$\phi(\omega) = \psi(j\omega) = \frac{1}{1+\omega^2\tau^2}$$

$$(s^2 + a_1 s + a_2)(s^2 - a_1 s + a_2) = s^4 + (a_1 - a_1)s^3 + (a_2 + a_2 - a_1^2)s^2 + (a_1 a_2 - a_1 a_2)s + a_2^2$$

$$= s^4 + (2a_2 - a_1^2)s^2 + a_2^2$$

Example(s)

$$y_t = H(q)e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$$\Psi(z) = H(z)H(z^{-1})\sigma^2$$

$$\phi(\omega) = H(e^{j\omega})H(e^{-j\omega})\sigma^2 = |H(e^{j\omega})|^2\sigma^2$$

$$e^{jk\omega} = cos(k\omega) + j\ sin(k\omega)$$

$$cos(k\omega) = \frac{e^{jk\omega} + e^{-jk\omega}}{2} \rightarrow \frac{z^k + z^{-k}}{2}$$

Example 1 - computerized

$$H(q) = \frac{2 - q^{-1}}{1 - 0.8q^{-1}} \qquad \sigma^2 = 1$$

```
a=[1 -0.8];
b=[2 -1];
[as,bs]=spec(a,b);
bs=bs*s2;

as =
   -0.8000    1.6400   -0.8000
bs =
    -2     5    -2
[asd,bsd]=spec2sd(as,bs)
asd =
    1.6400   -1.6000
bsd =
     5    -4
```

$$\psi(z) = \frac{-2z^{-1} + 5 - 2z}{-0.8z^{-1} + 1.64 - 0.8z}$$

$$\phi(\omega) = \frac{5 - 4\,cos(\omega)}{1.64 - 1.6\,cos(\omega)}$$

Example 2 - only results

$$H(q) = \frac{1 - 0.5\ q^{-1}}{1 - 0.8\ q^{-1}} \qquad \sigma^2 = 4$$

```
b =
    1.0000   -0.5000
s2 =
     4

Transfer function:
z - 0.5
-------
z - 0.8

Sampling time: 1
as =
   -0.8000    1.6400   -0.8000
bs =
    -2      5     -2
asd =
    1.6400   -1.6000
bsd =
     5    -4
```

$$\phi(\omega) = \frac{5 - 4\ cos(\omega)}{1.64 - 1.6\ cos(\omega)}$$

Notice same spec. dens.

Example 3

$$H(q) = \frac{1 - 2\,q^{-1}}{1 - 0.8\,q^{-1}} \qquad \sigma^2 = 1$$

```
b =
      1     -2
a =
    1.0000   -0.8000
s2 =
      1

Transfer function:
 z - 2
-------
z - 0.8

Sampling time: 1
as =
   -0.8000    1.6400   -0.8000
bs =
    -2      5     -2
asd =
    1.6400   -1.6000
bsd =
      5     -4
```

$$\phi(\omega) = \frac{5 - 4\,cos(\omega)}{1.64 - 1.6\,cos(\omega)}$$

Notice same spec. dens.

Spectral factorization

Given $H(q)$ and e_t, then

$$\phi(\omega) = H(e^{-j\omega})H(e^{j\omega})\sigma^2 = |H(e^{j\omega})|^2\sigma^2$$

What about the other way around.

Given a stationary process and $\phi(\omega) \geq 0$, then there exists and $H(q)$ with no zeroes outside stability area (and all poles inside) such that:

$$\phi(\omega) = H(e^{-j\omega})H(e^{j\omega})\sigma^2$$

if $\phi(\omega)$ is rational.

Ie. a rational function in z, $e^{j\omega}$ or $cos(\omega)$ (s^2, ω^2 in continuous time).

sfak see later.

Example 4 - MA process

Let's try the other way around

$$\phi(\omega) = 0.372 - 0.256\ cos(\omega)$$

$$\psi(z) = -0.128\ z^{-1} + 0.372 - 0.128\ z$$

```
bsd =
    0.3720   -0.2560
bs =
   -0.1280    0.3720   -0.1280

b=sfak(bs)

b =
    0.5665   -0.2259

s2=b(1)^2;
s2 =
    0.3210
b=polmon(b) % b=b/sqrt(s2)
b =
    1.0000   -0.3988
```

Example 4 - MA process - just checking

```
b =
    1.0000   -0.3988
bs=spec(b)*s2;
bs =
   -0.1280    0.3720   -0.1280         check 1


bsp=spec2sd(bs);
bsp =
    0.3720   -0.2560                   check 2
```

$$H(q) = B(q) = 1 - 0.4\, q^{-1} \qquad \sigma^2 = 0.321$$

Example 4

$$\phi(\omega) = \frac{0.372 - 0.256\ cos(\omega)}{1 - 0.94\ cos(\omega)}$$

```
bsd =
    0.3720   -0.2560
asd =
    1.0000   -0.9400

as =
   -0.4700    1.0000   -0.4700
bs =
   -0.1280    0.3720   -0.1280

a=sfak(as); b=sfak(bs);

a =
    0.8189   -0.5739
b =
    0.5665   -0.2259
```

$$H(q) = \frac{0.5665 - 0.2259\, q^{-1}}{0.8189 - 0.5739\, q^{-1}} \qquad \sigma^2 = 1$$

$$H(q) = \frac{0.6918 - 0.2759\, q^{-1}}{1 - 0.7\, q^{-1}} \qquad \sigma^2 = 1$$

denominator normalized

$$H(q) = \frac{1 - 0.4\, q^{-1}}{1 - 0.7\, q^{-1}} \qquad \sigma^2 = (0.6918)^2 = 0.4786$$

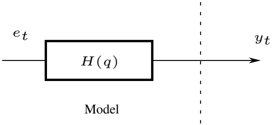
numerator normalized

Just checking

```
[as,bs]=spec(a,b);
bs=bs*s2



as =
   -0.7009    1.4912   -0.7009
bs =
   -0.1909    0.5547   -0.1909

[asd,bsd]=spec2sd(as,bs)

asd =
    1.4912   -1.4018
bsd =
    0.5547   -0.3818

bsd=bsd/asd(1)
asd=asd/asd(1)

bsd =
    0.3720   -0.2560
asd =
    1.0000   -0.9400
```

The representation theorem

Given a weak stationary stochastic process with a rational spectral density $\phi(\omega) \geq 0$. This process can be modelled as:



Notice:

- All poles inside stability area
- No zeroes outside stability area

Stochastic Systems

Stochastic systems

$$y_t = Gu_t + H_d v_t \qquad H_d = \frac{C}{A}$$

The deterministic part (assuming u_t is deterministic i.e. known).

$$\mathbf{E}\Big\{y_t\Big\} \;=\; m_t = G(q)u_t$$

The stochastic part. H_d is asymptotic stable. v_t is a weak stationary process.

$$A(q^{-1})r_{yv}(k) = C(q^{-1})r_v(k)$$

$$A(q^{-1})r_y(k) = C(q^{-1})r_{vy}(k)$$

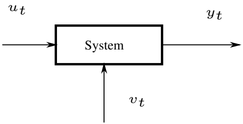
$$\Psi_{yv}(z) = H_d(z)\Psi_v(z)$$

$$\Psi_y(z) = H_d(z)H_d(z^{-1})\Psi_v(z)$$

$$\phi_{yv}(\omega) = H_d(e^{j\omega})\phi_v(\omega) \qquad \phi_y(\omega) = |H_d(e^{j\omega})|^2\phi_v(\omega)$$

In vector case as in C-time as well.

Stochastic systems



$$y_t = Gu_t + H_d v_t$$

H_d is asymptotic stable.The process v_t is assumed to be weakly stationary process with a rational spectrum ie.

$$v_t = H_n e_t$$

where H_n have no zeros (and no poles) outside the stability area.

$$y_t = Gu_t + He_t \qquad H = H_d H_n$$

where e_t is white noise and H is asymptotic stable.

What if H_d have unstable zeroes?

$$\begin{aligned}\Psi(z) &= H_d(z)H_n(z)H_d(z^{-1})H_n(z^{-1}) \\ &= H(z)H(z^{-1})\end{aligned}$$

where H have no zeros outside stability area.

AC-gain

$$x_{i+1} = Ax_i + Be_i$$

$$y_i = Cx_i + De_i$$

$$y_i = \left[\; C\big(qI - A\big)^{-1}B + D \; \right]e_i = H(q)e_i = \frac{C_p(q^{-1})}{A_p(q^{-1})}e_i$$

DC-gain, gain af $\omega = 0$.

$$K_{dc} = \frac{y_\infty}{e_0} = H(1)$$

AC-gain, Variance gain, Average (over frequencies) gain.

$$K_{ac} = \frac{var(y_i)}{var(e_i)}$$

if stochastic processes

If $e_i \in \mathbf{N}_{iid}(0, \sigma^2)$

$$P_x = AP_xA^T + B\sigma^2B^T$$

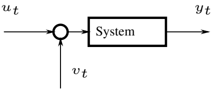
trfvar, dlyap

$$P_y = CP_xC^T + D\sigma^2D^T$$

$$P_y = \int_- \pi^\pi H(e^{j\omega})H(e^{-j\omega})d\omega$$

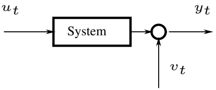
ARMAX structure

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$



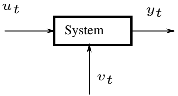
Box-Jenkins structure

$$y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$



L structure

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$



Continuous time spectrum

$\Psi_x(s) = \mathcal{L}_b[r_x(t)]$ $\Psi_{xy}(s) = \mathcal{L}_b[r_{xy}(t)]$ $s \in \mathbb{C}$ $t \in \mathbb{R}$

$\phi_x(\omega) = \Psi_x(j\omega) \in \mathbb{C}^{n\times n}$ $\phi_{xy}(\omega) = \Psi_{xy}(j\omega)$ $\omega \in \mathbb{R}$

$\Psi_x(s) = \Psi_x^\top(-s)$ $\Psi_{xy}(s) = \Psi_{yx}^\top(-s)$

$\phi_x(\omega) = \phi_x^\top(-\omega) \geq 0$ $\phi_{xy}(\omega) = \phi_{yx}^\top(-\omega)$

$$Var = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_x(\omega)d\omega$$

$\phi_x(\omega)$ most decrease for $|\omega| \to \infty$.

$$y(t_c) = H_c(\frac{d}{dt_c})e(t_c)$$

$$\phi_y(\omega) = H_c(j\omega)H_c(-j\omega)\sigma^2 = |H_c(j\omega)|^2\sigma^2$$

$$\phi_{ye}(\omega) = H_c(j\omega)\sigma^2$$

Given a stationary process and $\phi(\omega) \geq 0$, then there exists and $H(s)$ with no zeroes outside stability area (and all poles inside) such that:

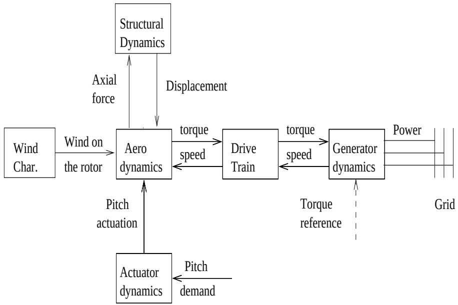
$$\phi(\omega) = H(j\omega)H(-j\omega)\sigma^2$$

if $\phi(\omega)$ is rational.

Approximative spectral density for wind:

$$\phi_v(\omega) = \frac{k^2}{(1+p_1^2\omega^2)(1+p_2^2\omega^2)}$$

Wind characteristic



Point wind speed

$$v_{wind} = v_m + v_p$$

$$\frac{f \cdot S_p(f)}{v_{*0}^2} = \frac{0.5 f_i}{1 + 2.2 f_i^{\frac{5}{3}}} \cdot \left(\frac{h_i}{L}\right)^{\frac{2}{3}} + \frac{105 f_{ru}}{(1 + 33 f_{ru})^{\frac{5}{3}}} \cdot \frac{(1 - \frac{h}{h_i})^2}{(1 + 15\frac{h}{h_i})^{\frac{2}{3}}}$$

$$f_i = \frac{f \cdot h_i}{v_m} \qquad f_{ru} = \frac{\frac{f \cdot h}{v_m}}{1 + 15\frac{h}{h_i}}$$

List of Symbol	
v_m	the mean wind speed [m/s]
v_{*0}	friction velocity [m/s]
h	measuring height [m]
h_i	height of convective boundary layer [m] $h_i = 1000$ m
L	Monon-Obukhov length [m]
f	frequency $f = \frac{\omega}{2\pi}$ [Hz]

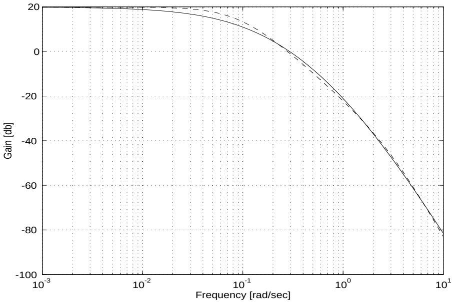
DTU

Effective wind speed

$$S_{ef}(f) = S_p(f)F(f) \qquad F(f) = \frac{1}{(1 + \frac{8\sqrt{\pi}}{3}\frac{R}{v_m}f)(1 + 4\sqrt{\pi}\frac{R}{v_m}f)}$$

$$\Phi_v(\omega) = \frac{k^2}{(1+p_1^2\omega^2)(1+p_2^2\omega^2)} \Phi_e(\omega) = H(j\omega)\Phi_e(\omega)H(-j\omega)$$

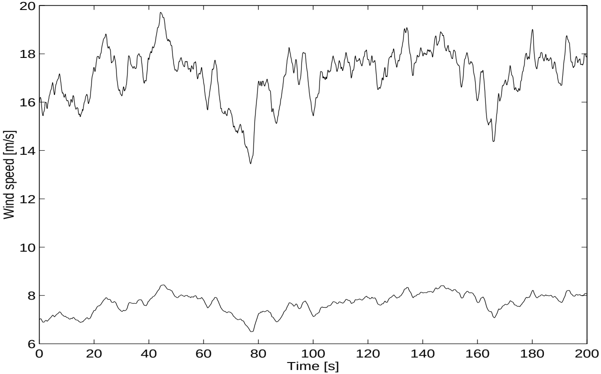
$$H(s) = \frac{k}{(1+p_1 s)(1+p_2 s)}$$



Wind model

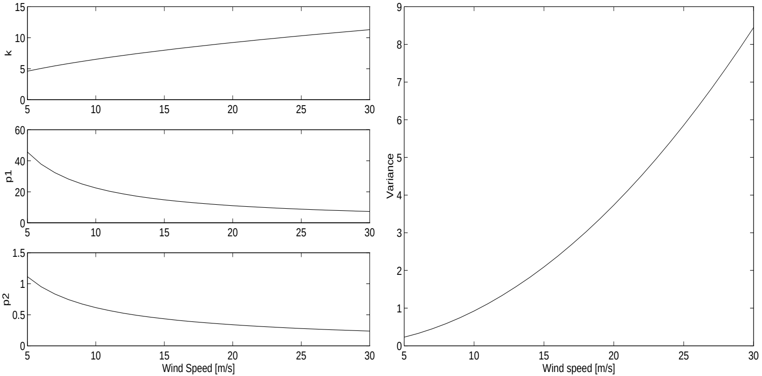
$$v_{wind} = v_m + \Delta v_{wind}$$

$$\Delta \ddot{v}_{wind} + \frac{p_1 + p_2}{p_1 p_2} \Delta \dot{v}_{wind} + \frac{1}{p_1 p_2} \Delta v_{wind} = \frac{k}{p_1 p_2} e$$



DTU

Parameter dependencies



Lesson Learned (L12)

- Spectrum of a process external model.
- Spectral density.
- Spectral factorization and representation theorem.
- Spectral analysis of a stochastic system .

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Xreg (L13-16)

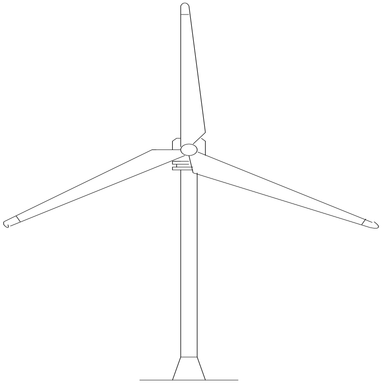
▸ L14 ▸ L15 ▸ L16

DTU

Learning objectives

- A bit motivation
- Prediction in the ARMA structure
- Prediction in the ARMAX structure
- The Diophantine equation

DTU



System and wind disturbance:

$$\begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \\ v \\ \dot{v} \end{bmatrix}_{t+1} = A_s \begin{bmatrix} \theta_\epsilon \\ \omega_r \\ \omega_g \\ \beta \\ v \\ \dot{v} \end{bmatrix}_t + B_s u_t + v_t \quad R_1$$

$$y_t = C_s x_t + e_t$$

Standard form

$$x_{t+1} = A_s x_t + B_s u_t + K\varepsilon_t$$
$$y_t = C_s x_t + \varepsilon_t$$

Innovation form

$$y_t = C_s \left[\ qI - A_s\ \right]^{-1} B_s u_t + \left(C_s \left[\ qI - A_s\ \right]^{-1} K + 1 \right) \varepsilon_t \qquad = G(q)u_t + H(q)\varepsilon_t$$

$$A(q^{-1})y_t = q^{-1}B(q^{-1})u_t + C(q^{-1})e_t$$

ARMAX form

WT linearized around $v_m = 17\ m/s$ and $f_s = 10\ Hz$

$$A(q^{-1})y_t = q^{-1}B(q^{-1})u_t + C(q^{-1})e_t$$

After model reduction:

$$\begin{aligned}
A(q^{-1}) &= 1 - 3.8q^{-1} + 5.9q^{-2} - 4.7q^{-3} + 1.9q^{-4} - 0.3q^{-5} + 3.9 \cdot 10^{-4}q^{-6} \\
B(q^{-1}) &= -875 - 2570q^{-1} + 4912q^{-2} - 558q^{-3} - 916q^{-4} - 10q^{-5} \\
C(q^{-1}) &= 1 - 0.61q^{-1}
\end{aligned}$$

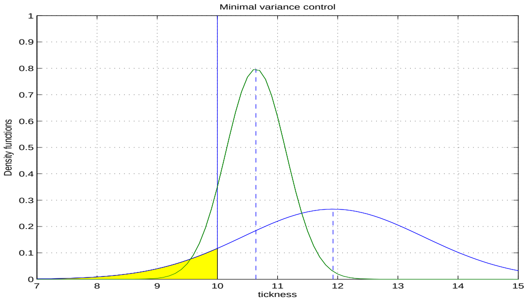
System and environment

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t$$

or more general:

$$A(q^{-1})y_t = q^{-k}\frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t + d$$

Cost function



$$J = \mathbf{E}\left\{(y_t - w)^2\right\}$$

Prediction

Prediction, Forecast, Prognosis

Has an importance in itself ($\rightarrow$ time series analysis):

- Prediction of $ value, Stock marked
- Power demand (Production planning), Energy potential in wind turbine farms
- Water level at Højer sluse

Here mainly because

- Control Design (predict the effect of the disturbances and our control action).
- System Identification

Prediction k step ahead

The system is a stationary ARMA process:

$$A(q^{-1})y_t = C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right) \perp Y_{t-1}$$

Constraints:

$$\hat{y}_{t+k|t} = func(Y_t)$$

Criteria:

$$J = \mathbf{E}\Big\{(y_{t+k} - \hat{y}_{t+k|t})^2\Big\}$$

Result:

$$\hat{y}_{t+k|t} = \mathbf{E}\Big\{y_{t+k}|Y_t\Big\}$$

Prediction in the ARMA-structure - marginal

$$y_t = \frac{C(q^{-1})}{A(q^{-1})} e_t$$ System model

$$\hat{y}_{t+k|t} = \frac{S(q^{-1})}{C(q^{-1})} y_t$$ Preditor (k step ahead)

$$\tilde{y}_{t+k|t} = G(q^{-1}) e_{t+k}$$ Prediction error

Diophantine equation:

$$C(q^{-1}) = A(q^{-1})G(q^{-1}) + q^{-k}S(q^{-1})$$

$$G(0) = 1, \quad ord(G) = k - 1$$

$$ord(S) = Max\{n_a - 1, n_c - k\}$$

L14

$$\frac{c_0 + c_1 q^{-1} + \ \dots\ + c_n q^{-n}}{1 + a_1 q^{-1} + \ \dots\ + a_n q^{-n}}$$

$$= c_0 + q^{-1} \frac{(c_1 - c_0 a_1) + (c_2 - c_0 a_2) q^{-1} + \ \dots\ (c_n - c_0 a_n) q^{1-n}}{1 + a_1 q^{-1} + \ \dots\ + a_n q^{-n}}$$

or stated shortly as:

$$\frac{C(q^{-1})}{A(q^{-1})} = g_0 + q^{-1} \frac{S_1(q^{-1})}{A(q^{-1})}$$

where

$$S_1(q^{-1}) = s_0 + s_1 q^{-1} + \ \dots\ + s_{n-1} q^{1-n}$$

The order of S_1 is $n-1$ (or less) and:

$$s_i = c_{i+1} - c_0 a_{i+1} \qquad i = 0,\ \dots\ n-1 \qquad\qquad g_0 = c_0$$

$$\frac{C(q^{-1})}{A(q^{-1})} = g_0 + q^{-1}\left\{g_1 + q^{-1}\left\{\ldots\left\{g_{k-1} + q^{-1}\frac{S_k(q^{-1})}{A(q^{-1})}\right\}\ldots\right\}\right\}$$

$$= g_0 + g_1 q^{-1} + \ \ldots \ + g_{k-1}q^{1-k} + q^{-k}\frac{S_k(q^{-1})}{A(q^{-1})}$$

or

$$\frac{C(q^{-1})}{A(q^{-1})} = G_k(q^{-1}) + q^{-k}\frac{S_k(q^{-1})}{A(q^{-1})} \tag{10}$$

where

$$G_k(q^{-1}) = g_0 + g_1 q^{-1} + \ \ldots \ + g_{k-1}q^{1-k}$$

and the order of S_k is $n-1$ (or less).

Here the coefficients, g_i, are the coefficients in the impulse response of

$$\frac{C(q^{-1})}{A(q^{-1})} = \sum_{i=0}^{\infty} g_i q^{-i}$$

The latter iff the system is asymptotic stable (as assumed).

Prediction in the ARMA structure - marginal

$$y_{t+k} = \frac{C(q^{-1})}{A(q^{-1})} e_{t+k} = G(q^{-1}) e_{t+k} + \frac{S(q^{-1})}{A(q^{-1})} e_t \qquad e_t = \frac{A(q^{-1})}{C(q^{-1})} y_t$$

$$y_{t+k} = G(q^{-1}) e_{t+k} + \frac{S(q^{-1})}{C(q^{-1})} y_t \qquad G(q^{-1}) e_{t+k} = e_{t+k} + ... + g_{k-1} e_{t+1}$$

$$\hat{y}_{t+k|t} = E\{y_{t+k} | Y_t\} = \frac{S(q^{-1})}{C(q^{-1})} y_t \qquad \tilde{y}_{t+k|t} = G(q^{-1}) e_{t+k} = e_{t+k} + ... + g_{k-1} e_{t+1}$$

$$Var\{\tilde{y}_{t+k|t}\} = \sigma^2 [1 + g_1^2 + ... + g_{k-1}^2] = \tilde{\sigma}^2$$

$$y_{t+k} | Y_t \in \mathbf{N}\left(\frac{S}{C} y_t, \tilde{\sigma}^2\right)$$

$r_{\tilde{y}}(m) = 0$ *for* $m > k$

Prediction - simultaneously

From marginal prediction

$$y_{t+k} = \frac{S_k(q^{-1})}{C(q^{-1})} y_t + G_k(q^{-1}) e_{t+k}$$

to simultaneously prediction

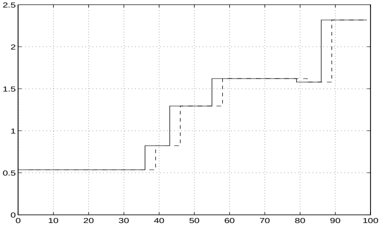
$$\begin{bmatrix} y_{t+1} \\ \vdots \\ y_{t+k} \\ \vdots \end{bmatrix} = \frac{1}{C(q^{-1})} \begin{bmatrix} S_1 \\ \vdots \\ S_k \\ \vdots \end{bmatrix} \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t+1-n} \end{bmatrix} + \begin{bmatrix} G_1 \\ \vdots \\ G_k \\ \vdots \end{bmatrix} \begin{bmatrix} e_{t+1} \\ e_{t+2} \\ \vdots \end{bmatrix}$$

$$= \begin{bmatrix} \hat{y}_{t+1} \\ \vdots \\ \hat{y}_{t+k} \\ \vdots \end{bmatrix} + \mathbf{G}E$$

$$Y_{t+1:t+k} | Y_t \in \mathbf{N}\left(\hat{Y}, \mathbf{G}\sigma^2\mathbf{G}^T\right)$$

Example - Random walk

$$w_t = \frac{1}{1-q^{-1}} e_t$$



$$1 = (1-q^{-1})(1 + g_1 q^{-1} + g_2 q^{-2} + \ldots + g_{k-1} q^{1-k}) + q^{-k} s_0$$

$$G(q^{-1}) = 1 + q^{-1} + q^{-2} + \ldots + q^{1-k} \qquad S(q^{-1}) = 1$$

$$\hat{w}_{t+k|t} = w_t$$

$$Var\left\{\tilde{w}_{t+k|t}\right\} = k\sigma^2$$

Pause

Diophantus of Alexandria

- Born between A.D. 200 and 214
- Dead between 284 and 298 (aged 84)
- sometimes called "the father of algebra"

Title page of Arithmetica


DIOPHANTI
ALEXANDRINI
ARITHMETICORVM
LIBRI SEX.
ET DE NVMERIS MVLTANGVLIS
LIBER VNVS.
AVCTORE CLAVDIO GASPARE BACHETO
LVTETIAE PARISIORVM,
Sumptibus SEBASTIANI CRAMOISY, via
Iacobæa, sub Ciconiis.
M. DC. XXI.
CVM PRIVILEGIO REGIS.

The Diophantine equation

Given A,$\bar{B}$ and C (which are general polynomials) find R and S from

$$C(q^{-1}) = A(q^{-1})R(q^{-1}) + \bar{B}(q^{-1})S(q^{-1})$$

$$\begin{aligned} C(q^{-1}) &= c_0 + c_1 q^{-1} + ... + c_{n_c} q^{-n_c} \\ \bar{B}(q^{-1}) &= b_1 q^{-1} + ... + b_{n_b} q^{-n_b} \\ A(q^{-1}) &= 1 + a_1 q^{-1} + ... + a_n q^{-n} \end{aligned}$$

Do not have an unique solution in general.

$$\begin{aligned} R(q^{-1}) &= R_0(q^{-1}) + \bar{B}(q^{-1})F(q^{-1}) \\ S(q^{-1}) &= S_0(q^{-1}) - A(q^{-1})F(q^{-1}) \end{aligned}$$

Unique solution if:

$$n_r = ord(R) = n_b - 1 \qquad\qquad n_s = Max\{n_a - 1, n_c - n_b\}$$

The Sylvester method

Let:

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} \qquad R(q^{-1}) = r_0 + r_1 q^{-1} + r_2 q^{-2}$$

then:

$$A(q^{-1})R(q^{-1}) \equiv \begin{bmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_2 & a_1 & 1 \\ 0 & a_2 & a_1 \\ 0 & 0 & a_2 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix}$$

$$C(q^{-1}) = A(q^{-1})R(q^{-1}) + \bar{B}(q^{-1})S(q^{-1})$$

$$\begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n_c} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \left(\begin{array}{ccc|ccc} 1 & \dots & 0 & 0 & \dots & 0 \\ & \ddots & \vdots & & & \vdots \\ a_1 & & \vdots & b_1 & & \vdots \\ a_2 & & 0 & b_2 & & 0 \\ \vdots & & 1 & \vdots & & b_1 \\ a_n & & a_1 & b_{n_b} & & b_2 \\ 0 & & \vdots & 0 & & \vdots \\ \vdots & & a_{n-1} & \vdots & & b_{n_b-1} \\ 0 & & a_n & 0 & & b_{n_b} \end{array}\right) \begin{pmatrix} r_0 \\ r_1 \\ \vdots \\ r_{n_r} \\ s_0 \\ s_1 \\ \vdots \\ s_{n_s} \end{pmatrix}$$

Number of equations:

$$Max\left[1+n_a+n_r, 1+n_b+n_s, 1+n_c\right]$$

Number of unknowns:

$$n_r+1+n_s+1$$

Match:

$$Max\left[n_a-n_s-1, n_b-n_r-1, n_c-n_r-n_s-1\right]=0$$

If now:

$$n_r = n_b - 1$$

Then ($[2]=0$ and $[1]$ nor $[3]$ must be lager than 0 ie.)

$$Max\left[n_a-n_s-1, n_c-n_b-n_s\right] \leq 0 \qquad Max\left[n_a-1, n_c-n_b\right] \leq n_s$$

or (since we are going for a minimum order solution)

$$n_s = Max\left[n_a-1, n_c-n_b\right]$$

The Diophantine equation

Truncated Impulse method - applicable if $\bar{B} = q^{-k}$.

$$C(q^{-1}) = A(q^{-1})R(q^{-1}) + q^{-k}S(q^{-1})$$

$$R(q^{-1}) = G(q^{-1}) = \left[\frac{C(q^{-1})}{A(q^{-1})}\right]_k$$

$$S(q^{-1}) = q^k\big(C(q^{-1}) - A(q^{-1})G(q^{-1})\big)$$

Example - $k = 1$

$$y_t - 1.7y_{t-1} + 0.7y_{t-2} = e_t + 1.5e_{t-1} + 0.9e_{t-2} \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$k = 1$

$$(1 + 1.5q^{-1} + 0.9q^{-2}) = (1 - 1.7q^{-1} + 0.7q^{-2})1 + q^{-1}(s_0 + s_1 q^{-1})$$

$$\begin{aligned} 1: &\qquad 1.5 = -1.7 + s_0 \qquad && s_0 = 3.2 \\ 2: &\qquad 0.9 = 0.7 + s_1 \qquad && s_1 = 0.2 \end{aligned}$$

$$\hat{y}_{t+1|t} = \frac{3.2 + 0.2q^{-1}}{1 + 1.5q^{-1} + 0.9q^{-2}} y_t$$

$$\tilde{y}_{t+1} = e_{t+1} \qquad Var = \sigma^2$$

$k = 2$

$$(1 + 1.5q^{-1} + 0.9q^{-2}) \quad = \quad (1 - 1.7q^{-1} + 0.7q^{-2})(1 + g_1 q^{-1}) + q^{-2}(s_0 + s_1 q^{-1})$$

$$\begin{aligned}
1: &\quad 1.5 = -1.7 + g_1 & g_1 &= 3.2 \\
2: &\quad 0.9 = 0.7 + -1.7g_1 + s_0 & s_0 &= 5.64 \\
3: &\quad 0 = 0.7g_1 + s_1 & s_1 &= -2.24
\end{aligned}$$

$$\hat{y}_{t+2|t} = \frac{5.64 - 2.24q^{-1}}{1 + 1.5q^{-1} + 0.9q^{-2}} y_t$$

$$\tilde{y}_{t+2} = e_{t+2} + 3.2e_{t+1} \qquad Var = \left[1 + (3.2)^2\right]\sigma^2 = 11.24\sigma^2$$

Prediction in a ARMAX model

▸ ARMA

I.e. k-step prediction (the horizon equals the time delay through the system)

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t \qquad Ay_{t+k} = Bu_t + Ce_{t+k}$$

Using the Diophantine equation:

$$C(q^{-1}) = A(q^{-1})G(q^{-1}) + q^{-k}S(q^{-1})$$

$$G(0) = 1 \qquad ord(G) = k-1$$

$$ord(S) = Max\{n_a - 1, n_c - k\}$$

$$\begin{aligned} y_{t+k} &= \frac{1}{C}\left(AG + q^{-k}S\right)y_{t+k} \\ &= \frac{1}{C}\left(GAy_{t+k} + Sy_t\right) \end{aligned}$$

$$y_{t+k} = \frac{1}{C}\left(GAy_{t+k} + Sy_t\right)$$

Just a copy of last result

$$\begin{aligned} y_{t+k} &= \frac{1}{C}\left(G\left[Bu_t + Ce_{t+k}\right] + Sy_t\right) \\ &= \frac{1}{C}\left(BGu_t + Sy_t\right) + Ge_{t+k} \end{aligned}$$

$$\hat{y}_{t+k|t} = \frac{1}{C(q^{-1})}\left(B(q^{-1})G(q^{-1})u_t + S(q^{-1})y_t\right)$$

$$\tilde{y}_{t+k|t} = G(q^{-1})e_{t+k} = e_{t+k} + \ldots + g_{k-1}e_{t+1}$$

MinVar

Alternative derivation

$$
\begin{aligned}
y_{t+k} &= \frac{B}{A}u_t + \frac{C}{A}e_{t+k} = \frac{B}{A}u_t + \frac{S}{A}e_t + Ge_{t+k} \\
&= \frac{B}{A}u_t + \frac{S}{A}\frac{A}{C}\left[y_t - q^{-k}\frac{B}{A}u_t\right] + Ge_{t+k} \\
&= \frac{B}{A}\left[1 - q^{-k}\frac{S}{C}\right]u + \frac{S}{C}y + Ge_{t+k} \\
&= \frac{B}{A}\frac{AG}{C}u_t + \frac{S}{C}y_t + Ge_{t+k} \\
&= \frac{1}{C}\left[BGu + Sy\right] + Ge_{t+k}
\end{aligned}
$$

▸ MV

End L13

Minimum variance control

Problem definition

System and environment

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right) \qquad e_t \perp \text{history}$$

Criterion:

$$\bar{J}_t = E\{y_{t+k}^2\}$$

Constraints:

$$u_t = funk\{\tilde{Y}_t\}$$

Prediction

DTU

Minimum variance control

$$\bar{J}_t = E\{y_{t+k}^2\} \qquad \min_{u_t(Y_t)} \mathbf{E}\left\{y_{t+k}^2\right\} = \mathbf{E}_{Y_t}\left\{\min_{u_t(Y_t)} \mathbf{E}\left\{y_{t+k}^2|Y_t\right\}\right\}$$

$$y_{t+k} = \frac{1}{C}\left[BGu_t + Sy_t\right] + Ge_{t+k}$$

$$\mathbf{E}\left\{y_{t+k}^2|Y_t\right\} = \left[\frac{1}{C}(BGu_t + Sy_t)\right]^2 + E\left\{[Ge_{t+k}]^2\right\}$$

Controller:

$$B(q^{-1})G(q^{-1})u_t = -S(q^{-1})y_t \qquad u_t = -\frac{S(q^{-1})}{B(q^{-1})G(q^{-1})}y_t$$

Design

$$C(q^{-1}) = A(q^{-1})G(q^{-1}) + q^{-k}S(q^{-1})$$

$$G(0) = 1 \quad ord(G) = k - 1 \qquad ord(S) = Max(n_a - 1, n_c - k)$$

Closed loop

$$y_t = G(q^{-1})e_t \qquad u_t = -\frac{S(q^{-1})}{B(q^{-1})}e_t$$

with variances

$$Var\{y_t\} = \sum_{i=0}^{k-1} g_i^2\sigma^2 \qquad y_t \perp Y_{t-k}$$

$$Var\{u_t\} = \textit{trfvar}(B, -S)\sigma^2$$

Example

Check the effect of k.

$$A = 1 - 1.5q^{-1} + 0.95q^{-2}$$

$$B = 1 + 0.5q^{-1} \qquad k = 1$$

$$C = 1 - 0.95q^{-1} \qquad e_t \in \mathbb{F}(0, \sigma^2) \qquad \sigma^2 = (0.1)^2$$

Design:

$$C \qquad = \qquad AG \qquad + \qquad q^{-k}S$$

$$(1 - 0.95q^{-1}) = (1 - 1.5q^{-1} + 0.95q^{-2})1 + q^{-1}(s_0 + s_1 q^{-1})$$

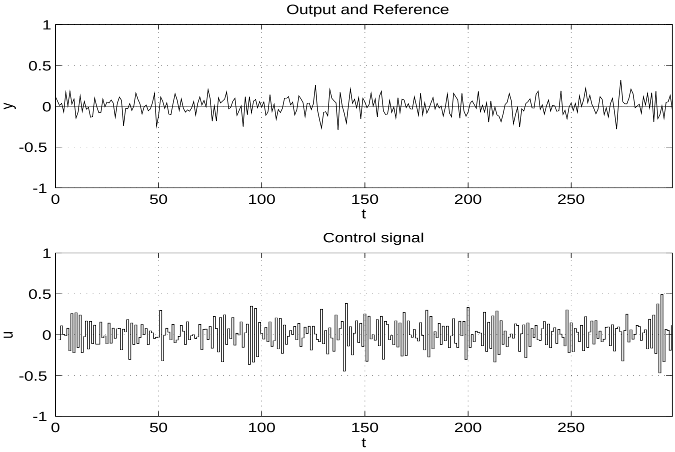
Controller

$$u_t = -\frac{S}{BG}y_t = -\frac{0.55 - 0.95q^{-1}}{1 + 0.5q^{-1}}y_t$$

$$u_t = -0.5u_{t-1} - 0.55y_t + 0.95y_{t-1}$$

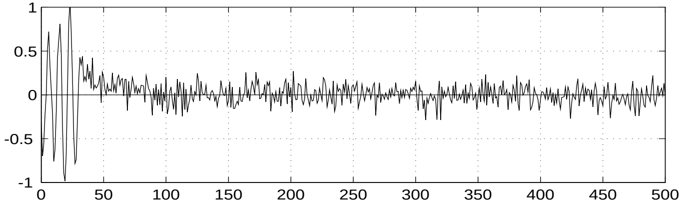
Closed loop:

$$y_t = e_t \qquad Var\{y_t\} = \sigma^2$$

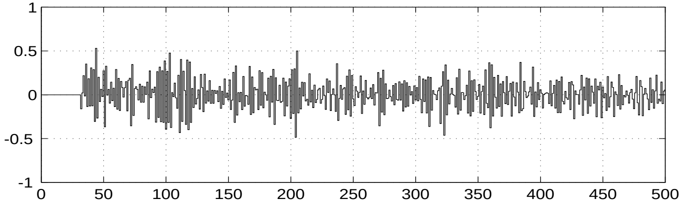


Output and Reference
y
1
0.5
0
-0.5
-1
0
50
100
150
200
250
t
Control signal
u
1
0.5
0
-0.5
-1
0
50
100
150
200
250
t

Cut in



1
0.5
0
-0.5
-1
0
50
100
150
200
250
300
350
400
450
500



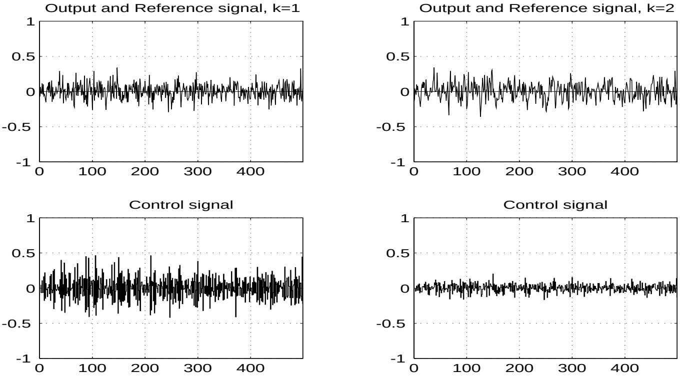
1
0.5
0
-0.5
-1
0
50
100
150
200
250
300
350
400
450
500

$k = 2$

$$(1-0.95q^{-1}) = (1-1.5q^{-1}+0.95q^{-2})(1+g_1 q^{-1}) + q^{-2}(s_0 + s_1 q^{-1})$$

$$u_t = -\frac{S}{BG} y_t \qquad \boxed{G = 1 + 0.55q^{-1}}$$

$$R = 1 + 1.05q^{-1} + 0.275q^{-2} \qquad S = -0.125 - 0.53q^{-1}$$



Problems with zeros

$$A = 1 - 1.5q^{-1} + 0.95q^{-2}$$

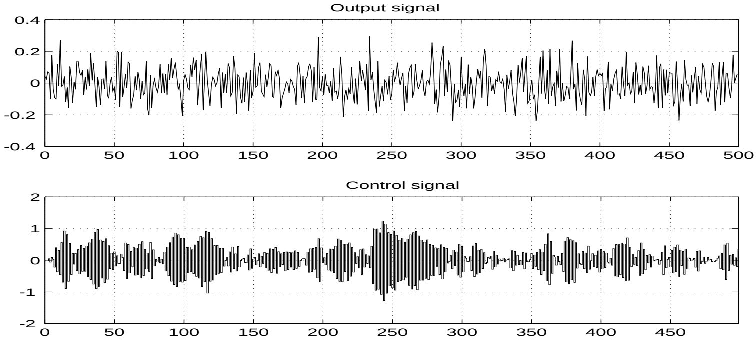
$$B = 1 + 0.95q^{-1} \qquad k = 1$$

$$C = 1 - 0.95q^{-1} \qquad e_t \in \mathbb{F}(0, \sigma^2) \qquad \sigma^2 = (0.1)^2$$

$$R = BG = 1 + 0.95q^{-1} \qquad S = 0.55 - 0.95q^{-1}$$

ie.

$$u_t = -0.95u_{t-1} - 0.55y_t + 0.95y_{t-1}$$



Minimum variance control

Is a basic stochastic control strategy, but have problems with:

- Set point
- Constant disturbance
- Large control effort (Detuning)
- Non damped system zeroes (in B and C).

MV_0-control

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

Cost:

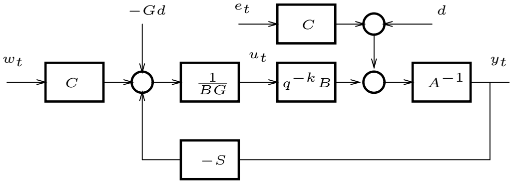
$$J = \mathbf{E}\left\{(y_{t+k} - w_t)^2\right\}$$

Controller:

$$BGu_t = Cw_t - Sy_t - Gd \qquad u_t = \frac{C}{BG}w_t - \frac{S}{BG}y_t - \frac{1}{B}d$$

Design:

$$C = AG + q^{-k}S \qquad G(0) = 1 \qquad ord(G) = k - 1 \qquad ord(S) = max(n_a - 1, n_c - k)$$



Realization

Why:

$$y_{t+k} = \frac{1}{C}\left[BGu_t + Sy_t + Gd\right] + Ge_{t+k}$$

$$y_{t+k} - w_t = \frac{1}{C}\left[BGu_t + Sy_t - Cw_t + Gd\right] + Ge_{t+k}$$

$$E\left\{ (y_{t+k} - w_t)^2 \middle| Y_t \right\} = \left\{ \frac{1}{C}\left[BGu_t + Sy_t - Cw_t + Gd\right] \right\}^2 + Var\{Ge_{t+k}\}$$

Closed loop

$$y_t = q^{-k} w_t + Ge_t$$

$$u_t = \frac{A}{B} w_t - \frac{S}{B} e_t - \frac{1}{B} d$$

▶ L15

Other structures

General L-structure

$$Ay_t = q^{-k}\frac{B}{F}u_t + \frac{C}{D}e_t + d$$

$$J = \mathbf{E}\Big\{(y_{t+k} - w_t)^2\Big\}$$

Box-Jenkins

$$y_t = q^{-k}\frac{B}{F}u_t + \frac{C}{D}e_t + d$$

$$J = \mathbf{E}\Big\{(y_{t+k} - w_t)^2\Big\}$$

Design

$$C = AD\,G + q^{-k}S$$

$$u_t = \frac{C}{BG}\frac{F}{D}w_t - \frac{S}{BG}\frac{F}{D}y_t - \frac{1}{B}d$$

Design

$$C = D\,G + q^{-k}S$$

$$u_t = \frac{C}{BG}\frac{F}{D}w_t - \frac{S}{BG}\frac{F}{D}y_t - \frac{1}{B}d$$

Closed loop

$$y_t = q^{-k}w_t + Ge_t$$

$$u_t = \frac{FA}{B}w_t - \frac{S}{B}\frac{F}{D}e_t - \frac{1}{B}d$$

Closed loop

$$y_t = q^{-k}w_t + Ge_t$$

$$u_t = \frac{F}{B}w_t - \frac{S}{B}\frac{F}{D}e_t - \frac{1}{B}d$$

Example

$$A = 1 - 1.5q^{-1} + 0.95q^{-2}$$

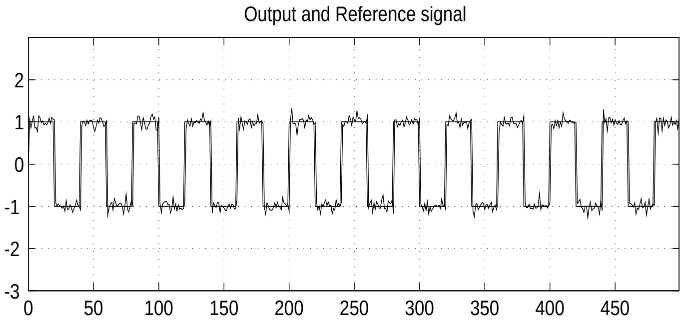
$$B = 1 + 0.5q^{-1} \qquad k = 1$$

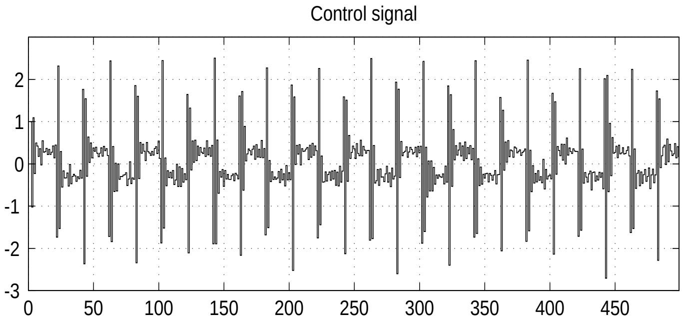
$$C = 1 - 0.95q^{-1} \qquad \sigma^2 = (0.1)^2$$

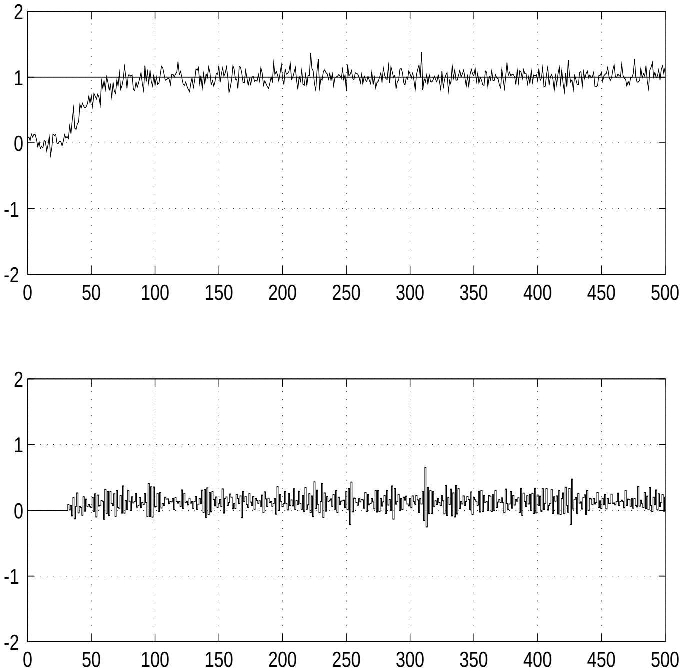
$$Q = C = 1 - 0.95q^{-1} \qquad R = BG = 1 + 0.5q^{-1}$$

$$S = 0.55 - 0.95q^{-1}$$

$$u_t = -0.5u_{t-1} + w_t - 0.95w_{t-1} - 0.55y_t + 0.95y_{t-1}$$


Output and Reference signal
2
1
0
-1
-2
-3
0
50
100
150
200
250
300
350
400
450


Control signal
2
1
0
-1
-2
-3
0
50
100
150
200
250
300
350
400
450


2
1
0
-1
-2
0
50
100
150
200
250
300
350
400
450
500
2
1
0
-1
-2
0
50
100
150
200
250
300
350
400
450
500

DTU

MV_0

Still problems with

- Control effort
- Non damped zeroes

End L14

Learning objectives - L15

MV_0 ▸ MV_0 has problems with control effort and non damped zeros

- PZ - controller
- Poleplacement controller
- MV_1 controller (I and II)
- Generalized Minimum Variance control
- MV_2 controller

PZ-control

PZ (Pole zero control, Change location (s) of pole(s) and zero(s).)

Detuning

y_t *close to* $y_m(t) = q^{-k}\frac{B_m}{A_m}w_t$ *rather than* $q^{-k}w_t$

Example:

System:

$$(1 - 0.98q^{-1})y_t = q^{-2}(1 + 0.3q^{-1})u_t + (1 + 0.74q^{-1})e_t$$

Goal:

$$y_m(t) = q^{-2}\frac{0.6}{1 - 0.4q^{-1}}w_t$$

PZ-control

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

Cost:

$$J = \mathbf{E}\Big\{(A_m y_{t+k} - B_m w_t)^2\Big\}$$

Controller:

$$BGu_t = B_m C w_t - S y_t - Gd$$

Design:

$$A_m C = AG + q^{-k}S$$

$$G(0) = 1 \qquad ord(G) = k - 1 \qquad ord(S) = max(n_a - 1, n_c + n_{a_m} - k)$$

DTU

PZ - Proof

Why:

$$A_m y_{t+k} = \frac{A_m C}{C} y_{t+k} = \frac{1}{C}\left[A G y_{t+k} + S y_t \right]$$

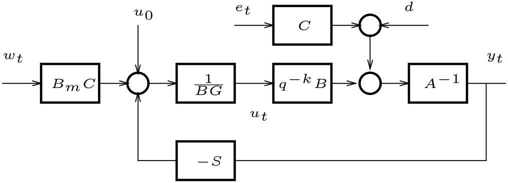
$$= \frac{1}{C}\left[G\Big(B u_t + C e_{t+k} + d \Big) + S y_t \right]$$

$$A_m y_{t+k} = \frac{1}{C}\Big[B G u_t + S y_t + G d \Big] + G e_{t+k}$$

$$A_m y_{t+k} - B_m w_t = \frac{1}{C}\Big[B G u_t + S y_t - C B_m w_t + G d \Big] + G e_{t+k}$$

Controller:

$$BGu_t = B_m C w_t - S y_t - G d$$



Closed loop

$$y_t = q^{-k} \frac{B_m}{A_m} w_t + \frac{G}{A_m} e_t \qquad u_t = \frac{A}{B} \frac{B_m}{A_m} w_t - \frac{S}{B} \frac{1}{A_m} e_t - \frac{1}{B} d$$

Detuned, but still problems with not well damped system zeroes.

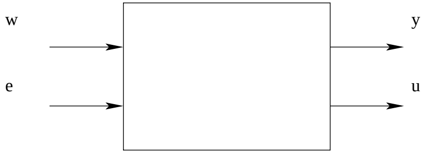
Closed Loop

System:

$$Ay = \bar{B}u + Ce \qquad \bar{B} = q^{-k}B$$

Controller:

$$Ru = Qw - Sy$$



Closed Loop

System:

$$Ay = \bar{B}u + Ce + d \qquad \bar{B} = q^{-k}B$$

Controller:

$$Ru = Qw - Sy - \gamma$$

Multiply system with (R)

$$ARy = \bar{B}Ru + CRe + Rd$$

Multiply controller with ($\bar{B}$)

$$\bar{B}Ru = \bar{B}Qw - \bar{B}Sy - \bar{B}\gamma$$

and obtain

$$(AR + \bar{B}S)y = \bar{B}Qw + RCe + (Rd - \bar{B}\gamma)$$

System:

$$Ay = \bar{B}u + Ce + d$$

Controller:

$$Ru = Qw - Sy - \gamma$$

Multiply system with (S)

$$ASy = \bar{B}Su + CSe + Sd$$

Multiply controller with (A)

$$ARu = AQw - ASy - A\gamma$$

and obtain

$$(AR + \bar{B}S)u = AQw - CSe - (A\gamma + Sd)$$

GSP-control

General stochastic pole placement controller

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

Argumentation:

Goal:

$$y_m(t) = q^{-k}\frac{B_m}{A_m}w_t$$

close related to:

$$\bar{J}_t = E\left\{(A_m y_{t+k} - B_m w_t)^2\right\}$$

Previously (PZ control) we looked at:

$$A_m y_{t+k} - B_m w_t = \frac{1}{C}\Big[BGu_t + Sy_t - CB_m w_t + Gd\Big] + Ge_{t+k}$$

We had then problems with undamped system zeros.

▸ GSP alternative derivation

DTU

GSP control

We might change the PZ set up:

$$A_m y_{t+k} - B_m w_t = \frac{1}{C}\left[BGu_t + Sy_t - CB_m w_t + Gd\right] + Ge_{t+k}$$

Now, the problem with the system zeroes:

$$B = B_+ B_-$$

If we don't cancel the zeros in B_-, then we have to keep them in B_m ie.

$$B_m = B_- B_{m_1}$$

Let's try the following Diophantine equation $(C \to A_o)$

$CA_m = AG + q^{-k}S$	$ord(G) = k-1$	PZ
$A_o A_m = AG + q^{-k}B_- S$	$ord(G) = k + n_{b_-} - 1$	GSP

then

$$A_m y_{t+k} - B_m w_t = \frac{B_-}{A_o}\left[B_+ Gu_t + Sy_t - A_o B_{m_1} w_t + \frac{G}{B_-} d\right] + \frac{C}{A_o} Ge_{t+k}$$

$$A_o A_m = AG + q^{-k} B_- S \qquad \text{(a copy)}$$

Why:

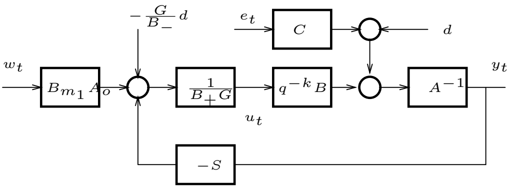
$$\begin{aligned} A_m y_{t+k} &= \frac{A_m A_o}{A_o} y_{t+k} = \frac{1}{A_o}\left[AG y_{t+k} + B_- S y_t\right] \\ &= \frac{1}{A_o}\left[G\left(Bu_t + Ce_{t+k} + d\right) + B_- S y_t\right] \end{aligned}$$

$$A_m y_{t+k} = \frac{1}{A_o}\left[B_+ B_- G u_t + B_- S y_t + Gd\right] + Ge_{t+k}$$

$$A_m y_{t+k} - B_m w_t = \frac{B_-}{A_o}\left[B_+ G u_t + S y_t - A_o B_{m_1} w_t + \frac{G}{B_-} d\right] + Ge_{t+k}$$

Controller

$$B_+ G u_t = B_{m_1} A_o w_t - S y_t - \frac{G}{B_-} d$$



Design

1. Factorize $B = B_+ B_-$
2. Choose A_m, B_{m_1}
$$DC\left[\frac{B_{m_1} B_-}{A_m}\right] = 1$$
3. Choose A_o and solve
$$A_o A_m = AG + q^{-k} B_- S$$
for S and G
4. Use the controller:
$$B_+ G u_t = B_{m_1} A_o w_t - S y_t - \frac{G}{B_-} d$$

DTU

GSP - Alternative derivation

If the system ▸ Closed loop

$$Ay = \bar{B}u + Ce + d \qquad \text{is closed with} \qquad Ru = Qw - Sy - \gamma$$

then the closed loop is given by

$$y = \frac{\bar{B}Q}{AR + \bar{B}S} w + \frac{RC}{AR + \bar{B}S} e + \frac{Rd - \bar{B}\gamma}{AR + \bar{B}S}$$

Now factorize $\bar{B} = q^{-k} B_+ B_-$ then

$$\frac{q^{-k} B_+ B_- Q}{AR + q^{-k} B_+ B_- S} = q^{-k} \frac{B_m}{A_m}$$

Furthermore let $R = B_+ G$ (R must have B_+ as a factor)

$$\frac{q^{-k} B_+ B_- Q}{AB_+ G + q^{-k} B_+ B_- S} = \frac{q^{-k} B_- Q}{AG + q^{-k} B_- S} = q^{-k} \frac{B_m}{A_m}$$

Since we don't want to cancel B_- it must be contained in B_m, i.e. $B_m = B_- B_{m_1}$. This results in

$$\frac{q^{-k} B_- Q}{AG + q^{-k} B_- S} = \frac{q^{-k} B_- B_{m_1}}{A_m}$$

or (sufficient condition)

$$\frac{Q}{AG + q^{-k} B_- S} = \frac{B_{m_1}}{A_m}$$

GSP - Alternative derivation

$$\frac{Q}{AG + q^{-k}B_-S} = \frac{B_{m_1}}{A_m}$$

just a copy

The only way B_{m_1} can be introduced is via Q. Factorize Q into $Q = B_{m_1}A_o$, then

$$\frac{B_{m_1}A_o}{AG + q^{-k}B_-S} = \frac{B_{m_1}}{A_m}$$

or

$$AG + q^{-k}B_-S = A_mA_o$$

Back tracking:

$$R = B_+G \qquad Q = B_{m_1}A_o \qquad \gamma = \frac{G}{B_-}d$$

$$Ru_t = Qw_t - S_t - \gamma$$

$$y_t = q^{-k} \frac{B_{m_1} B_-}{A_m} w_t + \frac{G}{A_m} \frac{C}{A_o} e_t$$

$$u_t = \frac{A}{A_m} \frac{B_{m_1}}{B_+} w_t - \frac{S}{A_m B_+} \frac{C}{A_o} e_t - \frac{1}{B} d$$

GSP control of wind turbine

($v_m = 17\ m/s$ and $f_s = 10\ Hz$)

$$A(q^{-1})y_t = q^{-1}B(q^{-1})u_t + C(q^{-1})e_t$$

System and wind:

$$\begin{aligned}
A(q^{-1}) &= 1 - 3.8058q^{-1} + 5.9125q^{-2} - 4.6877q^{-3} + 1.8890q^{-4} \\
&\quad -0.30817q^{-5} + 3.8557 \cdot 10^{-4}q^{-6} \\
B(q^{-1}) &= -875.47 - 2570.7q^{-1} + 4912.7q^{-2} - 558.79q^{-3} \\
&\quad -916.30q^{-4} - 9.8661q^{-5} \\
C(q^{-1}) &= 1 - 0.60653q^{-1}
\end{aligned}$$

Zeroes:

p_{b1} = -4.2716

p_{b2} = -0.34865

p_{b3} = -0.010846

p_{b4} = 0.70469

p_{b5} = 0.99005

Factorize B

$$\begin{aligned} B^-(q^{-1}) &= (1 - p_{b1}q^{-1}) \\ B^+(q^{-1}) &= b_0(1 - p_{b2}q^{-1})(1 - p_{b3}q^{-1})(1 - p_{b4}q^{-1})(1 - p_{b5}q^{-1}) \end{aligned}$$

$$H_m(q) = \frac{0.007197 + 0.03074q^{-1}}{1 - 2.399q^{-1} + 2.100q^{-2} - 0.6585q^{-3} - 0.0613q^{-4} + 0.0566q^{-5}}$$

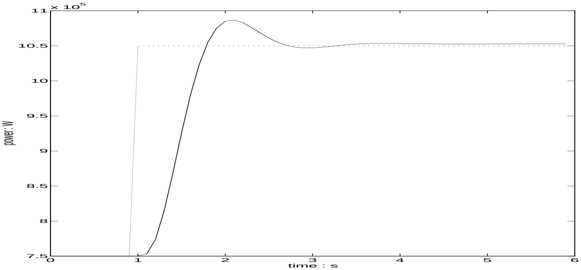
$$B_m^1(q^{-1}) = -1.1743\ 10^{-5}$$

Solution:

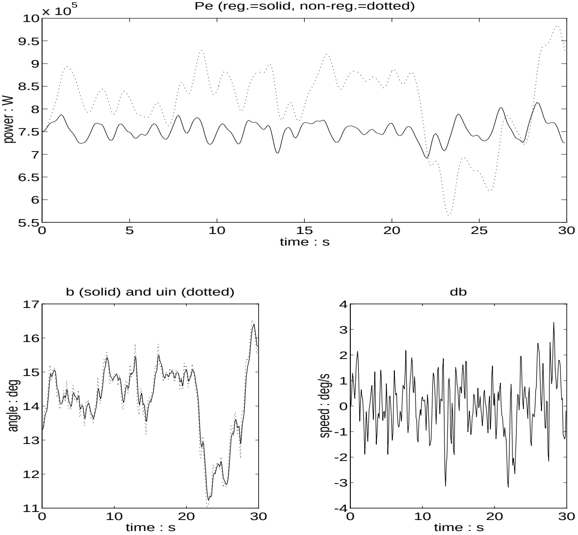
$$\begin{aligned} S(q^{-1}) &= -2.1577 \cdot 10^{-4} + 6.1771 \cdot 10^{-4}q^{-1} - 6.8528 \cdot 10^{-4}q^{-2} + \\ &\quad 3.4421 \cdot 10^{-4}q^{-3} - 6.5465 \cdot 10^{-5}q^{-4} + 9.8279 \cdot 10^{-8}q^{-5} \end{aligned}$$

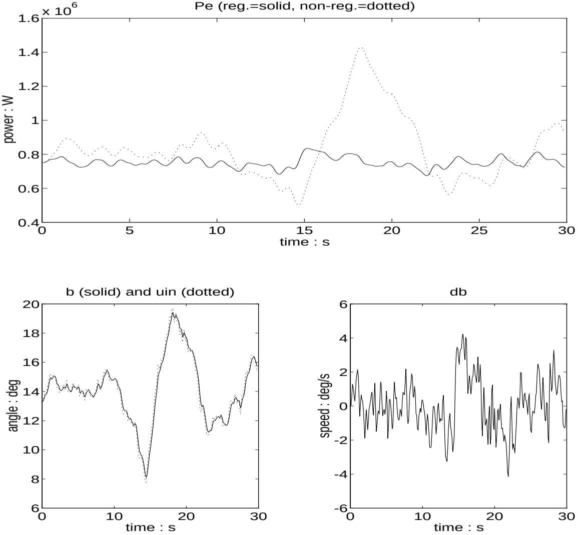
$$Q(q^{-1}) = B_m^1(q^{-1})C(q^{-1}) = -1.1743 \cdot 10^{-5} + 7.1226 \cdot 10^{-6}q^{-1}$$

$$\begin{aligned} R(q^{-1}) &= B^+(q^{-1})G(q^{-1}) = 1 - 0.66670q^{-1} - 0.80045q^{-2} + 0.30604q^{-3} + 0.16603q^{-4} + \\ &\quad 1.7641 \cdot 10^{-3}q^{-5} + 4.7409 \cdot 10^{-9}q^{-6} + 5.9708 \cdot 10^{-12}q^{-7} + \\ &\quad 7.5285 \cdot 10^{-15}q^{-8} + 9.1726 \cdot 10^{-18}q^{-9} \end{aligned}$$



x 10^5
11
10.5
10
9.5
9
8.5
8
7.5
0
1
2
3
4
5
6
time : s
power : W


Pe (reg.=solid, non-reg.=dotted)
× 10^5
power : W
time : s
b (solid) and uin (dotted)
angle : deg
time : s
db
speed : deg/s
time : s


Pe (reg.=solid, non-reg.=dotted)
× 10^6
power : W
time : s
b (solid) and uin (dotted)
angle : deg
time : s
db
speed : deg/s
time : s

Design

1. Choose $B_- = B$ (i.e. $B_+ = 1$).
2. Choose A_m, B_{m_1} and A_o
$$DC\left[\frac{B_{m_1}B}{A_m}\right] = 1$$
3. Solve
$$A_oA_m = AG + q^{-k}BS$$
for S and G. Here $R = G$.
4. Use the controller:
$$Ru_t = B_{m_1}A_ow_t - Sy_t - \frac{G}{B}d$$

(All zeros cancelled gives PZ control)

Closed loop

$$y_t = q^{-k} \frac{B_{m_1} B}{A_m} w_t + \frac{G}{A_m} \frac{C}{A_o} e_t$$

$$u_t = \frac{A B_{m_1}}{A_m} w_t - \frac{S}{A_m} \frac{C}{A_o} e_t - \frac{1}{B} d$$

Preliminaries (Just a copy)

Incomplete state information:

$$\min_{u(y)} \mathbf{E}\Big\{ I(x,u) \Big\} = \mathbf{E}\Big\{ \min_{u(y)} \mathbf{E}\Big\{ \ I(x,u) | \ y \Big\} \Big\}$$

$$x|y \in \mathbb{F}\left(\hat{x}, P\right)$$

$$\mathbf{E}\Big\{ x^\top S x | y \Big\} = \hat{x}^\top S \hat{x} + tr\left(SP\right)$$

MV_1 Controller I

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

Cost:

$$J = \mathbf{E}\Big\{(y_{t+k} - w_t)^2 + \rho u_t^2\Big\}$$

Controller:

$$[BG + \alpha C]u_t = Cw_t - Sy_t - Gd$$

Design:

$$C = AG + q^{-k}S \qquad \alpha = \frac{\rho}{b_0}$$

$$G(0) = 1 \qquad ord(G) = k - 1 \qquad ord(S) = max(n_a - 1, n_c - k)$$

Closed loop:

$$y_t = q^{-k}\frac{B}{B + \alpha A}w_t + \frac{BG + \alpha C}{B + \alpha A}e_t$$

$$u_t = \frac{A}{B + \alpha A}w_t - \frac{S}{B + \alpha A}e_t - \frac{1}{B}d$$

DTU

Optimality

$$J = \mathbf{E}\Big\{(y_{t+k} - w_t)^2 + \rho u_t^2\Big\}$$

$$I = \mathbf{E}\Big\{ (y_{t+k} - w_t)^2 + \rho u_t^2 \Big| Y_t\Big\}$$
$$= (\hat{y}_{t+k} - w_t)^2 + \rho u_t^2$$

$$y_{t+k} = \frac{1}{C}(BGu_t + Sy_t + Gd) + Ge_{t+k}$$

$$\hat{y}_{t+k} = \frac{1}{C}(BGu_t + Sy_t + Gd)$$

Stationarity

$$2(\hat{y}_{t+k} - w_t)b_0 + 2\rho u_t = 0$$

$$\frac{1}{C}(BGu_t + Sy_t + Gd - Cw_t + \alpha C u_t) = 0$$

$$\alpha = \frac{\rho}{b_0}$$

$$(BG + \alpha C)u_t + Sy_t + Gd - Cw_t = 0$$

MV_1 Controller II

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

Cost:

$$J = \mathbf{E}\left\{(y_{t+k} - w_t)^2 + \rho(u_t - u_{t-1})^2\right\}$$

Controller:

$$[BG + \alpha(1 - q^{-1})C]u_t = Cw_t - Sy_t - Gd$$

Design:

$$C = AG + q^{-k}S \qquad \alpha = \frac{\rho}{b_0}$$

$$G(0) = 1 \qquad ord(G) = k - 1 \qquad ord(S) = max(n_a - 1, n_c - k)$$

Closed loop:

$$y_t = q^{-k}\frac{B}{B + \alpha(1 - q^{-1})A}w_t + \frac{BG + \alpha C}{B + \alpha(1 - q^{-1})A}e_t$$

$$u_t = \frac{A}{B + \alpha(1 - q^{-1})A}w_t - \frac{S}{B + \alpha(1 - q^{-1})A}e_t - \frac{1}{B}d$$

GMV Control (Clarke-Gawthrop)

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

Cost:

$$\bar{J} = E\left\{[\tilde{\boldsymbol{y}}_{t+k} - \tilde{\boldsymbol{w}}_t]^2 + \rho\tilde{\boldsymbol{u}}_t^2\right\}$$

$$\tilde{\boldsymbol{y}}_t = H_y(q)y_t \qquad \tilde{\boldsymbol{w}}_t = H_w(q)w_t \qquad \tilde{\boldsymbol{u}}_t = H_u(q)u_t$$

$$H_y(q) = \frac{B_y(q^{-1})}{A_y(q^{-1})} \qquad H_u(q) = \frac{B_u(q^{-1})}{A_u(q^{-1})} \qquad H_w(q) = \frac{B_w(q^{-1})}{A_w(q^{-1})}$$

Interpretation: Minimal Variance control of

$$\xi_t = \frac{B_y(q^{-1})}{A_y(q^{-1})}y_t + q^{-k}\left(\alpha\frac{B_u(q^{-1})}{A_u(q^{-1})}u_t - \frac{B_w(q^{-1})}{A_w(q^{-1})}w_t\right) \qquad \alpha = \frac{\rho}{b_0}$$

Controller:

$$[A_uBG + \alpha CB_u]\,u_t = C\frac{A_uB_w}{A_w}w_t - S\frac{A_u}{A_y}y_t - A_uGd$$

GMV

$$[A_u BG + \alpha C B_u]\, u_t = C \frac{A_u B_w}{A_w} w_t - S \frac{A_u}{A_y} y_t - A_u G d$$

Design:

$$B_y C = A_y A G + q^{-k} S$$

$$\alpha = \frac{\rho}{b_0}$$

Closed loop: See book.

y_t *close to* $y_m(t) = q^{-k}\frac{B_m}{A_m}w_t$ *but what about the noise*

PZ Control

$$J = \mathbf{E}\Big\{(A_m y_{t+k} - B_m w_t)^2\Big\} \qquad y_t = q^{-k}\frac{B_m}{A_m}w_t + \frac{G}{A_m}e_t$$

Variation of MV_0 Control:

$$J = \mathbf{E}\Big\{(y_{t+k} - \frac{B_m}{A_m}w_t)^2\Big\} \qquad y_t = q^{-k}\frac{B_m}{A_m}w_t + Ge_t$$

Yet another version

$$J = \mathbf{E}\Big\{(\frac{A_m}{B_m}y_{t+k} - w_t)^2\Big\} \qquad y_t = q^{-k}\frac{B_m}{A_m}w_t + \frac{B_m}{A_m}Ge_t$$

These have different noise characteristics.

MV_2 Control

System:

$$Ay_t = q^{-k}Bu_t + Ce_t + d$$

Cost:

$$J = \mathbf{E}\Big\{\Big(\frac{A_e}{B_e}y_{t+k} - \frac{A_eB_m}{B_eA_m}w_t\Big)^2\Big\}$$

$$H_y = \frac{A_e}{B_e} \qquad H_w = \frac{A_eB_m}{B_eA_m} \qquad \rho = 0$$

Controller:

$$BGu_t = C\frac{B_w}{A_w}w_t - S\frac{1}{A_y}y_t - Gd$$

Design:

$$A_eC = B_eAG + q^{-k}S$$

Closed loop:

$$y_t = q^{-k}\frac{B_m}{A_m}w_t + \frac{B_e}{A_e}Ge_t$$

$$u_t = \frac{AB_m}{BA_m}w_t - \frac{SB_e}{BA_e}e_t - \frac{1}{B}d$$

▸ L16

- PZ - controller
- Poleplacement controller (GSP)
- MV_1 controller (I and II)
- Generalized Minimum Variance control (GMV)
- MV_2 controller

End L15

- MV_0 (Minimum variance control)
  - large control activity
  - only well damped zeros
- PZ (pole-zero-placement)
  - reduced requirement to the error (reference model)
  - only well damped zeros
- GSP (Generalized poleplacement controller)
  - factorize B.
  - handle not well damped zeros
- GMV (generalized minimum variance controller)
  - weight on control activity
  - frequency weights or reference model(s)
  - handle not well damped zeros

All based on a one step/instant horizon

GMV

System

$$A y_t = \bar{B} u_t + C e_t \qquad \bar{B} = q^{-k} B \quad = b_0 q^{-k} + b_1 q^{-k-1} + \ \ldots$$

Cost:

$$\bar{J}_t = \lim_{N \to \infty} \frac{1}{N} \sum_{i=t+1}^{t+N} \mathbf{E}\left\{ y_i^2 + \rho^2 u_i^2 \right\}$$

Controller:

$$R u_t = -S y_t \qquad u_t = -\frac{S(q^{-1})}{R(q^{-1})}\, y_t$$

Design:

Solve for R and S:

$$A_m C = AR + \bar{B} S$$

where A_m is a stable solution to

$$A_m(z^{-1}) A_m(z) = \bar{B}(z^{-1}) \bar{B}(z) + \rho^2 A(z^{-1}) A(z)$$

Set point and disturbance.

Controller:

$$R u_t = \eta C w_t - S y_t + u_0$$

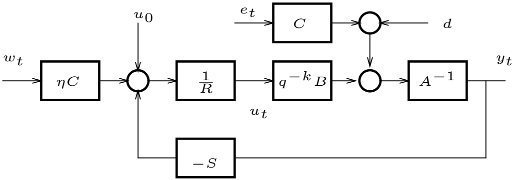
Design

$$A_m C = AR + \bar{B} S$$

$$A_m(z^{-1}) A_m(z) = \bar{B}(z^{-1}) \bar{B}(z) + \rho^2 A(z^{-1}) A(z)$$

$$u_0 = -\frac{R(1)}{\bar{B}(1)} d \qquad \eta = \frac{A_m(1)}{\bar{B}(1)}$$

Notice, its ability to handle not well define time delays.



Special cases

Design:

$$A_m C = AR + \bar{B}S$$

$$A_m(z^{-1})A_m(z) = \bar{B}(z^{-1})\bar{B}(z) + \rho^2 A(z^{-1})A(z)$$

$$Ru_t = \eta C w_t - S y_t + u_0$$

$\rho = 0$ and B is stable.

In that case $A_m = B$ is the solution to

$$A_m(z^{-1})A_m(z) = \bar{B}(z^{-1})\bar{B}(z) + \rho^2 A(z^{-1})A(z)$$

The Diophantine equation becomes:

$$BC = AR + \bar{B}S$$

With $\bar{B} = q^{-k}B$ we must have $R = BG$ ie. then the Diophantine equation becomes

$$C = AG + q^{-k}S$$

and the controller becomes:

$$BGu_t = Cw_t - Sy_t - Gd$$

i.e. a MV_0 controller.

$\rho = 0$ and B is having non stable zeros.

Here $B = B_- B_+$ and $A_m = B_+ \underline{B}$ is the solution to

$$A_m(z^{-1})A_m(z) = \bar{B}(z^{-1})\bar{B}(z) + \rho^2 A(z^{-1})A(z)$$

where $\underline{B}$ is B_- with its zeros mirrored in the stability limit i.e. $\underline{B}(z^{-1})\underline{B}(z) = B_-(z^{-1})B_-(z)$.

The Diophantine equation becomes:

$$\underline{B} B_+ C = AR + \bar{B} S$$

With $\bar{B} = q^{-k} B_+ B_-$ and $R = B_+ G$ we must have the Diophantine equation

$$\underline{B} C = AG + q^{-k} B_- S$$

and the controller:

$$R u_t = C w_t - S y_t - G d \qquad R = B_+ G$$

Often the controller:

$$\underline{B} B_+ G u_t = C w_t - S y_t - G d$$

is used. Here

$$C = AG + q^{-k} S$$

GPC - control (Clarke et al)

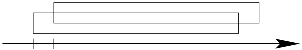
GPC: Generalized Predictive Control

$$A(q^{-1})y_t = B(q^{-1})u_{t-1} + C(q^{-1})e_t$$

Notice change of notation

$$J = \mathbf{E}\Big\{ \sum_{i=1}^{N} \big(y_{t+i} - w_{t+i}\big)^2 + \rho^2 u_{t+i-1}^2 \Big\}$$

Receding horizon:



Direct or as a QRS controller:

$$R(q^{-1})u_t = Q(q^{-1})w_t - S(q^{-1})y_t$$

The Cost Function

$$J = \mathbf{E}\Big\{ \sum_{i=1}^{N} \big(y_{t+i} - w_{t+i}\big)^2 + \rho^2 u_{t+i-1}^2 \Big\}$$

Let

$$Y = \begin{bmatrix} y_{t+1} \\ y_{t+2} \\ \vdots \\ y_{t+N} \end{bmatrix} \qquad W = \begin{bmatrix} w_{t+1} \\ w_{t+2} \\ \vdots \\ w_{t+N} \end{bmatrix} \qquad U = \begin{bmatrix} u_{t} \\ u_{t+1} \\ \vdots \\ u_{t+N-1} \end{bmatrix}$$

$$J = \mathbf{E}\Big\{ \|Y - W\|^2 + \rho^2 \|U\|^2 \Big\}$$

$$\|Z\|^2 = Z^T Z$$

Constraints

Available information (and define some quantities) :

$$U_o = \frac{1}{C(q^{-1})} \begin{bmatrix} u_{t-1} \\ \vdots \\ u_{t-n} \end{bmatrix} \qquad Y_o = \frac{1}{C(q^{-1})} \begin{bmatrix} y_t \\ \vdots \\ y_{t+1-n} \end{bmatrix}$$

$$\min_{U(Y_o)} \mathbf{E}\Big\{ \|Y - W\|^2 + \rho^2 \|U\|^2 \Big\}$$

$$\min_{U} \mathbf{E}\Big\{ \|Y - W\|^2 + \rho^2 \|U\|^2 \Big| \, Y_o \Big\}$$

$$\min_{U} \|\hat{Y} - \hat{W}\|^2 + \rho^2 \|U\|^2$$

DTU

Pause

Result 1: Linear prediction

$$Y = HU + F_u U_o + F_y Y_o + GE$$

Future output = forced response + free response + disturbance

$$Y = \begin{bmatrix} y_{t+1} \\ y_{t+2} \\ \vdots \\ y_{t+N} \end{bmatrix} \qquad U = \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+N-1} \end{bmatrix}$$

$$H = \begin{bmatrix} h_1 & 0 & \dots & 0 \\ h_2 & h_1 & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ h_N & h_{N-1} & \cdots & h_1 \end{bmatrix}$$

$$U_o = \frac{1}{C(q^{-1})} \begin{bmatrix} u_{t-1} \\ \vdots \\ u_{t-n} \end{bmatrix} \qquad Y_o = \frac{1}{C(q^{-1})} \begin{bmatrix} y_t \\ \vdots \\ y_{t+1-n} \end{bmatrix}$$

DTU

Result 1: Linear prediction

Let us consider a horizon m, where $1 \leq m \leq N$. The system description

$$A(q^{-1})y_t = B(q^{-1})u_{t-1} + C(q^{-1})e_t$$

and the Diophantine equation (truncated impulse response)

$$C = AG_m + q^{-m}S_m$$

gives the usual expression

$$y_{t+m} = \frac{1}{C}\Big[BG_m u_{t+m-1} + S_m y_t\Big] + G_m e_{t+m}$$

DTU

Result 1: Linear prediction

Consider the result from previuous slide:

$$y_{t+m} = \frac{1}{C}\Big[BG_m u_{t+m-1} + S_m y_t\Big] + G_m e_{t+m}$$

and yet another Diophantine equation

$$BG_m = CH_m + q^{-m} F_m$$

gives us that

$$y_{t+m} = H_m u_{t+m-1} + \frac{1}{C}\Big[F_m u_{t-1} + S_m y_t\Big] + G_m e_{t+m}$$

Polynomials and vectors

$$y_t = G(q^{-1})u_t \qquad = (g_0 + g_1 q^{-1} + \cdots + g_n q^{-n})u_t$$

$$y_t = \qquad\qquad g_0 u_t + g_1 u_{t-1} + \cdots + g_n u_{t-n}$$

$$y_t = \begin{pmatrix} g_0 & g_1 & \ldots & c_n \end{pmatrix} \begin{pmatrix} u_t \\ u_{t-1} \\ \vdots \\ u_{t-n} \end{pmatrix} \qquad = GU$$

DTU

Result 1: Linear prediction

$$y_{t+m} = H_m u_{t+m-1} + \frac{1}{C}\Big[F_m u_{t-1} + S_m y_t\Big] + G_m e_{t+m}$$

For example (similarily results goes for Fu, Sy and Ge)

$$H_m u_{t+m-1} = h_1 u_{t+m-1} + h_2 u_{t+m-2} + \ \ldots \ + h_m u_t$$

or (here a reversed sequence is natural)

$$H_m u_{t+m-1} = h_m u_t + h_{m-1} u_{t+1} + \ \ldots \ + h_1 u_{t+m-1}$$

$$H_m u_{t+m-1} = \begin{bmatrix} h_m & h_{m-1} & \cdots & h_1 \end{bmatrix} \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+m-1} \end{bmatrix}$$

Result 1: Linear prediction

$$y_{t+m} = H_m u_{t+m-1} + \frac{1}{C}\Big[F_m u_{t-1} + S_m y_t\Big] + G_m e_{t+m}$$

$$Y = HU + F_u U_o + F_y Y_o + GE$$

Future output = forced response + free response + disturbance

---

$$Y = \begin{bmatrix} y_{t+1} \\ y_{t+2} \\ \vdots \\ y_{t+N} \end{bmatrix} \quad U = \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+N-1} \end{bmatrix} \qquad U_o = \frac{1}{C(q^{-1})}\begin{bmatrix} u_{t-1} \\ \vdots \\ u_{t-n} \end{bmatrix} \quad Y_o = \frac{1}{C(q^{-1})}\begin{bmatrix} y_t \\ \vdots \\ y_{t+1-n} \end{bmatrix}$$

$$H = \begin{bmatrix} h_1 & 0 & \dots & 0 \\ h_2 & h_1 & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ h_N & h_{N-1} & \cdots & h_1 \end{bmatrix} \qquad F_u = \begin{bmatrix} \vdots \\ F_m \\ \vdots \end{bmatrix} \quad F_y = \begin{bmatrix} \vdots \\ S_m \\ \vdots \end{bmatrix}$$

```
function [H,Fy,Fu,G]=pred(a,b,c,N)
% Model
%
% Ay=q^{-1}Bu+Ce
%
% It is assumed that a,b,c have same order
% (ie. are padded with zeros).
%
% Y=HU+FyYo+FuUo+GE

%-----------------------------------------------------
 hu=sysimp(a,b,N); H=poltplz(hu',N); H=H(1:N,1:N);
 he=sysimp(a,c,N); G=poltplz(he',N); G=G(1:N,1:N);

 ar=a(2:end); cr=c(2:end); br=b(2:end);
 f=cr-ar; s=br*he(1)-cr*hu(1);
 Fy=f; Fu=s;
 for i=2:N
  f=[f(2:end) 0]-ar*he(i);
  Fy=[Fy; f];
  s=[s(2:end) 0]+br*he(i)-cr*hu(i);
  Fu=[Fu; s];
  end
```

pause

GPC - control

Unconstrained direct GPC control is based on the assumption that

$$\hat{Y} = HU + f \qquad f = F_y Y_o + F_u U_o$$

$$\begin{aligned} J &= \left(\hat{Y} - \hat{W}\right)^T \left(\hat{Y} - \hat{W}\right) + \rho^2 U^T U \\ &= U^T \left[H^T H + \rho^2 I\right] U + 2\left(f - \hat{W}\right)^T HU + \left(f - \hat{W}\right)^T \left(f - \hat{W}\right) \end{aligned}$$

$$\boxed{U^* = -\left[H^T H + \rho^2 I\right]^{-1} H^T \left(f - \hat{W}\right) = -K\left(f - \hat{W}\right)}$$

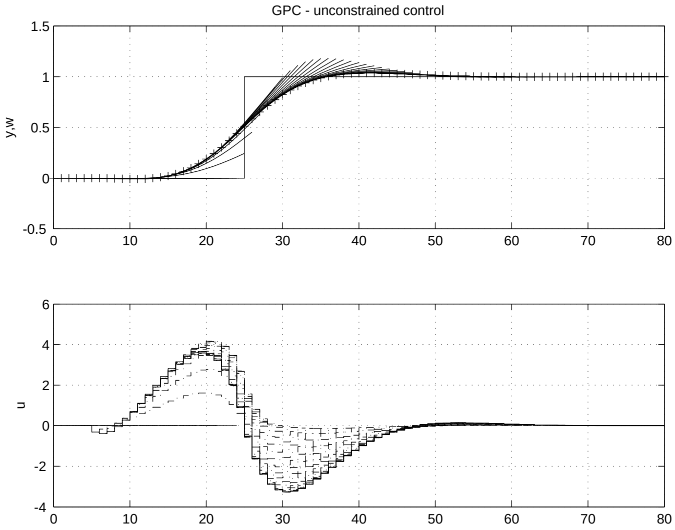
where

$$K = \left[H^T H + \rho^2 I\right]^{-1} H^T$$

Receeding horizon:

$$u_t = \left[1,\ 0,\ 0,\ \dots,\ 0\right] U = \gamma U^*$$

$$u_t = -\gamma K\left(F_y Y_o + F_u U_o - \hat{W}\right)$$



GPC - unconstrained control
y,w
1.5
1
0.5
0
-0.5
0
10
20
30
40
50
60
70
80
u
6
4
2
0
-2
-4
0
10
20
30
40
50
60
70
80

Results 2: Vectors and polynomials

Let V be a row vector:

$$V = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$$

$$VY_o = V \frac{1}{C(q^{-1})} \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t+1-n} \end{bmatrix} = V \frac{1}{C(q^{-1})} \begin{bmatrix} 1 \\ q^{-1} \\ \vdots \\ q^{1-n} \end{bmatrix} y_t$$

$$V_p(q^{-1}) = V \begin{bmatrix} 1 \\ q^{-1} \\ \vdots \\ q^{1-n} \end{bmatrix} = v_1 + v_2 q^{-1} + \dots v_n q^{1-n}$$

$$VY_o = \frac{V_p(q^{-1})}{C(q^{-1})} y_t = \frac{[V](q^{-1})}{C(q^{-1})} y_t$$

GPC as a QRS controller

Prediction of the reference $\hat{w}_{t+i} = w_t$ (random walk).

$$\hat{W} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} w_t = \underline{1} w_t$$

$$u_t = -\gamma K \Big(F_y Y_o + F_u U_o - \underline{1} w_t \Big) \qquad \text{(just a copy)}$$

$$u_t = \gamma K \underline{1} w_t - \gamma K F_y Y_o - \gamma K F_u U_o \quad = \quad \gamma K \underline{1} w_t - \frac{[\gamma K F_y](q^{-1})}{C(q^{-1})} y_t - \frac{[\gamma K F_u](q^{-1})}{C(q^{-1})} u_{t-1}$$

$$Q(q^{-1}) = C(q^{-1})[\gamma K \underline{1}](q^{-1})$$

$$R(q^{-1}) = C(q^{-1}) + q^{-1}[\gamma K F_u](q^{-1})$$

$$S(q^{-1}) = [\gamma K F_y](q^{-1})$$

$$R(q^{-1}) u_t = Q(q^{-1}) w_t - S(q^{-1}) y_t$$

DTU

```
function [Q,R,S]=dsngpc0(A,B,k,C,lam,N)
%
% GPC controller,
% R*u(t)=Q*w(t)-S*y(t)
% for the system
% A*y(t)=B*u(t-k)+C*e(t)
% such that
% J = sum_N (y-w)^2 + lam*u^2(t)
% is minimized.

%---------------------------------------------------------
B=poldel(B,k-1);
[a,b,c,n]=armaxpad(A,B,C);
[G,Fy,Fu]=pred(a,b,c,N);
K=pinv(G'*G+lam*eye(N))*G';

gamma=[1 zeros(1,N-1)];
en=ones(N,1);
Q=polmul(c,gamma*K*en);
R=polsum(c,[0 gamma*K*Fu]);
S=gamma*K*Fy;

Q=polclr(Q); R=polclr(R); S=polclr(S);

r0=R(1);
R=R/r0; S=S/r0; Q=Q/r0;
```

Lesson learned - L17

- LQG control
- GPC control

End L17

Controller realization

Realization of transferfunctions

$$y_t = H(q)u_t \qquad\qquad x_{t+1} = Ax_t + Bu_t$$

$$y_t = Cx_t + Du_t$$

$$H(q) = \frac{\bar{b}_0 + \bar{b}_1 q^{-1} + ... + \bar{b}_n q^{-n}}{1 + a_1 q^{-1} + ... + a_n q^{-n}} = \sum_{i=0}^{\infty} h_i q^{-i}$$

Canonical realization

$$A_o = \begin{bmatrix} -a_1 & 1 & ... & 0 \\ \vdots & & \ddots & \\ -a_{n-1} & 0 & ... & 1 \\ -a_n & 0 & ... & 0 \end{bmatrix} \qquad B_o = \begin{bmatrix} \bar{b}_1 - \bar{b}_0 a_1 \\ \bar{b}_2 - \bar{b}_0 a_2 \\ \vdots \\ \bar{b}_n - \bar{b}_0 a_n \end{bmatrix}$$

$$C_o = (1, 0, ..., 0) \qquad D_o = \bar{b}_0 \tag{11}$$

Controller realization

Direct realization

$$y_t = H(q)u_t \qquad H(q) = \frac{\bar{b}_0 + \bar{b}_1 q^{-1} + \ldots + \bar{b}_n q^{-n}}{1 + a_1 q^{-1} + \ldots + a_n q^{-n}}$$

$$A(q^{-1})y_t = B(q^{-1})u_t$$

$$y_t + a_1 y_{t-1} + \ \ldots\ + a_n y_{t-n} = b_0 u_t + b_1 u_{t-1} + \ \ldots\ + b_n u_{t-n}$$

$$y_t = -a_1 y_{t-1} - \ \ldots\ - a_n y_{t-n} + b_0 u_t + b_1 u_{t-1} + \ \ldots\ + b_n u_{t-n}$$

$$y_t = C_d x_t + D_d u_t$$

$$x_{t+1} = A_d x_t + B_d u_t$$

Controller realization

$$Ru_t = Qw_t - Sy_t$$

$$u_t = \frac{[Q \;\; S]}{R} \begin{bmatrix} w_t \\ -y_t \end{bmatrix}$$

$$A_o = \begin{bmatrix} -r_1 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ -r_{n-1} & 0 & \dots & 1 \\ -r_n & 0 & \dots & 0 \end{bmatrix}$$

$$B_o = \begin{bmatrix} q_1 - q_0 r_1 & s_1 - s_0 r_1 \\ q_2 - q_0 r_2 & s_2 - s_0 r_2 \\ \vdots & \\ q_n - q_0 r_n & s_n - s_0 r_n \end{bmatrix}$$

$$C_o = (1, 0, \dots, 0)$$

$$D_o = \begin{bmatrix} q_0 & s_0 \end{bmatrix}$$

Controller realization

```
%------------------------------------------------------------------------
[Ar,Br,Cr,Dr]=armax2ss(R,Q,0,S);
nr=length(Ar); Xr=zeros(nr,1);
%------------------------------------------------------------------------

measinit;                % Initilialise the measurement system
for it=1:nstp,

 w=wt(it); wf=wft(it);
 [y,t]=meas;                       % Measure output

 u=Cr*Xr+Dr*[wf;-y];               % Fixed parameter controller

 act(u);                           % Actuate control

 Xr=Ar*Xr+Br*[wf;-y];              % Fixed parameter controller

end
```

▸ MV_0 (L14)

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

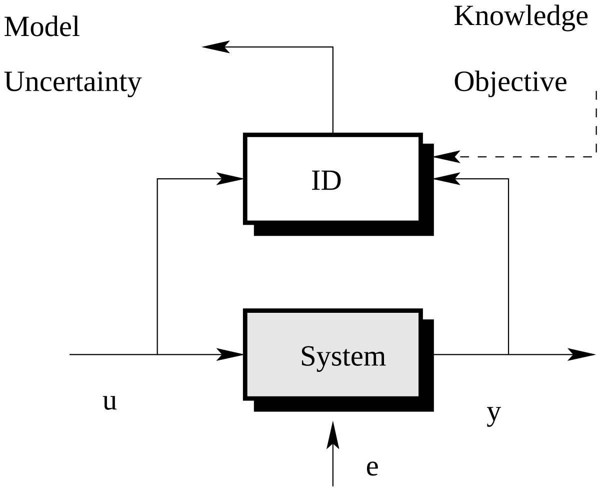
Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

System identification I - L17

The Big picture

- Stochastic system.
- Control of (known) stochastic systems
- System identification
- Adaptive control

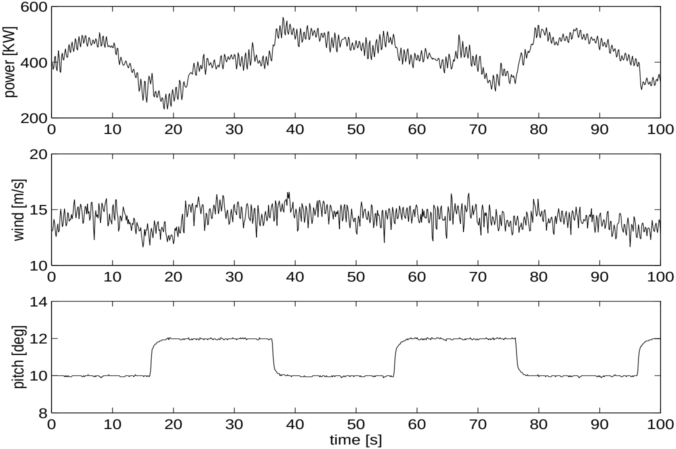
Outline - Learning objectives - L17


Model
Uncertainty
Knowledge
Objective
ID
System
u
y
e

System identification

Obtain a model from data and knowledge.

Model type reflects the application of the model.



Physical modelling - deduction

Based on:

- conservation laws
- continuity equation
- material properties

Results in:

- non linear model
- large area of validation
- rarely a model for stochastic disturbances

Statistical modelling - induction

Based on:

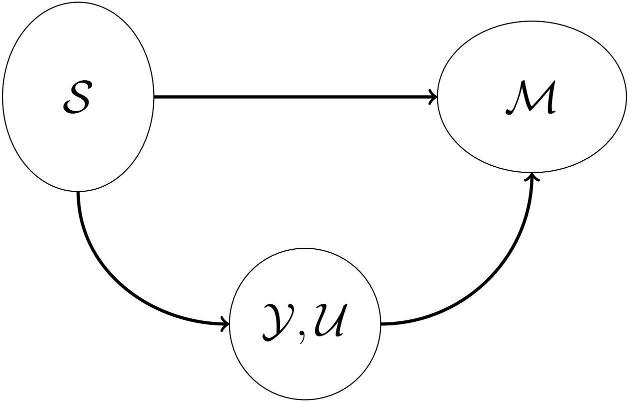
- data from actual process
- data from certain experimental conditions

Results in:

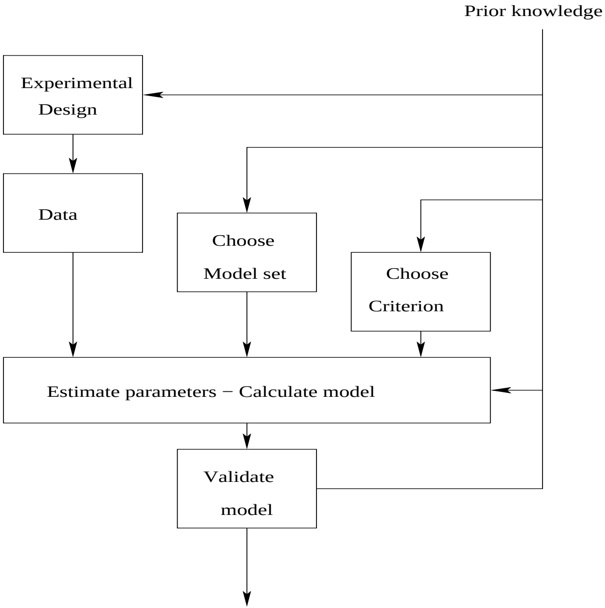
- often a linear model
- small area of validation
- often model of disturbances
- model type chosen by user

Practical system identification is based on a combination

What is a good model



The modelling process

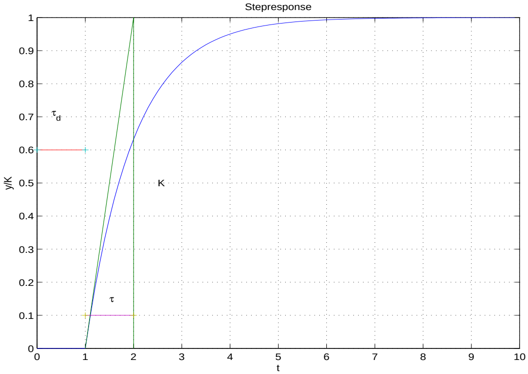


Non-parametric methods

Methods where we estimate a function or a curve.

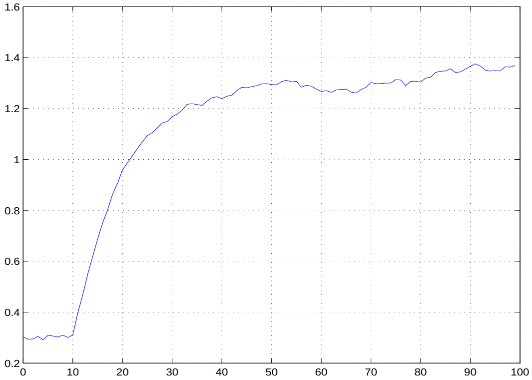
Step response methods

$$G(s) = e^{-s\tau_d} \frac{K}{1 + s\tau} \qquad \tau\dot{y}(t_c) + y(t_c) = Ku(t_c - \tau_d)$$

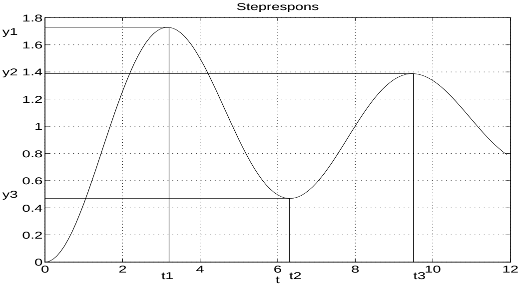


Disadvantages: repeat experiment to kill influence from noise.

Step response methods



$$G(s) = K \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 + \omega_0^2} \qquad \ddot{y}(t_c) + 2\zeta\omega_0\dot{y}(t_c) + \omega_0^2 y(t_c) = K\omega_0^2 u(t_c)$$



$$M = 1 - \frac{y_2}{y_1} \qquad T_p = t_3 - t_1$$

$$\zeta = \frac{-log(M)}{\sqrt{\pi^2 + log(M)^2}} \qquad \omega_0 = \frac{2\pi}{T_p\sqrt{1-\zeta^2}} = \frac{2}{T_p}\sqrt{\pi^2 + log(M)^2}$$

DTU

Frequency response methods

Input (exitation, probe signal)

$$u_t = \alpha \ \ cos(\omega t)$$

$$Y(z) = G(z)U(z) + Y_0(z) \qquad\qquad Y(s) = G(s)U(s) + Y_0(s)$$

Repsonse:

$$y_t = \xi(t) + Acos(\omega t + \phi)$$

$$A = |G(\omega)|\alpha \qquad \phi = arg(G(\omega))$$

Estimation:

$$I_c(N) = \frac{1}{N}\sum_{t=1}^{N} y_t cos(\omega t) \qquad\qquad I_s(N) = \frac{1}{N}\sum_{t=1}^{N} y_t sin(\omega t)$$

$$|\hat{G}(\omega)| = \frac{2}{\alpha}\sqrt{I_c^2 + I_s^2} \qquad \hat{\phi} = -arctan\frac{I_s(N)}{I_c(N)}$$

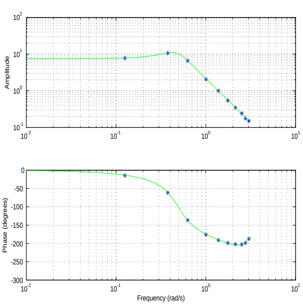
Use: FFT or ETFE

ETFE

Deterministic simulation

```
m = idpoly([1 -1.5 0.7],[0 1 0.5]);
u = iddata([],idinput([50,1,10],'sine'));
u.Period = 50;
y = sim(m,u);
me = etfe([y u]);
bode(me,'b*',m)
```

Result



Correlation methods

Dynamics:

$$y_t = G(q)u_t \qquad\qquad y_t = g_t \star u_t$$

Covariance structure:

$$r_{yu}(\tau) = g_\tau \star r_u(\tau)$$

Estimation:

$$\hat{r}_{yu}(\tau) = \frac{1}{N}\sum_{t=1}^{N} y_t u_{t-\tau} \qquad\qquad \hat{r}_u(\tau) = \frac{1}{N}\sum_{t=1}^{N} u_t u_{t-\tau}$$

Special case: If u_t is white then

$$r_{yu}(\tau) = \sigma^2 g_\tau$$

Prewhitening.

Use: cra

```
CRA    Performs correlation analysis to estimate impulse response.
   IR = CRA(Z)

   Z: The data, entered as an IDDATA object or a matrix
      with two columns Z = [y u].
   IR: The estimated impulse response (IR(1) corresponds to g(0))

   [IR,R,CL] = CRA(Z,M,NA,PLOT) gives access to
   M: The number of lags for which the functions are computed (def 20)
   NA: The order of the whitening filter.(Def 10). With NA=0, no prewhite-
       ning is performed. Then the covariance functions of the original
       data are obtained.
   PLOT: PLOT=0 gives no plots. PLOT=1 (Default) gives a plot of IR along
       with a 99 % confidence region. PLOT=2 gives a plot of all R's.

   Note that in the plot, the response to a normalized pulse input,
      u(t) = 1/T for 0<t<T, is shown, where T is the sampling interval
      of the data.

   R: The covariance/correlation information
      R(:,1) contains the lag indices
      R(:,2) contains the covariance function of y (poss. prewhitened)
      R(:,3) contains the covariance function of u (poss. prewhitened)
      R(:,4) contains the correlation function between (poss prewhitened)
        u and y (positive lags corresponds to an influence from u to y)
      CL is the 99 % significance level for the impulse response
      The plots can be redisplayed by CRA(R);
```

Spectral analysis

Dynamics

$$y_t = G(q)u_t + v_t$$

Spectral densities:

$$\Phi_{yu}(\omega) = G(e^{j\omega})\Phi_u(\omega) \qquad \Phi_y(\omega) = |G(e^{j\omega})|^2\Phi_u(\omega) + \Phi_v(\omega)$$

$$\hat{G}(e^{j\omega}) = \frac{\hat{\Phi}_{yu}(\omega)}{\hat{\Phi}_u(\omega)} \qquad \hat{\Phi}_v(\omega) = \hat{\Phi}_y(\omega) - \frac{|\hat{\Phi}_{yu}(\omega)|^2}{\hat{\Phi}_u(\omega)}$$

Estimation

$$\hat{\Phi}_y(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\hat{r}_y(k)e^{-jk\omega} \qquad \hat{\Phi}_{yu}(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\hat{r}_{yu}(k)e^{-jk\omega}$$

Windows technique. Estimate of r quite uncertain for large k.

$$\hat{\Phi}_y(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\hat{r}_y(k)w(k)e^{-jk\omega} \qquad \hat{\Phi}_{yu}(\omega) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\hat{r}_{yu}(k)w(k)e^{-jk\omega}$$

Use: spa

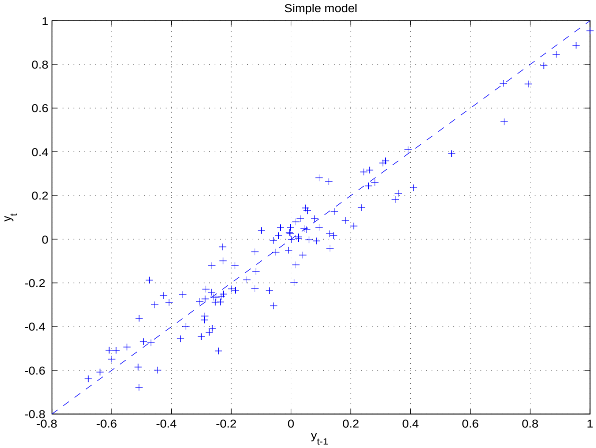
```
SPA    Performs spectral analysis.
   G = SPA(DATA)

   DATA: The output-input data as an IDDATA object. (see help IDDATA)
   G: Returned frequency response and uncertainty as an IDFRD object.
      See IDPROPS IDFRD for a description of this object.
   G also contains the spectrum of the additive noise v in the
   model y = G u + v. See IDPROPS IDFRD.

   If DATA is a time series G is returned as the estimated
   spectrum, along with its estimated uncertainty.

   The spectra are computed using a Hamming window of lag size
   min(length(DATA)/10,30), which can be changed to M by
   G = SPA(DATA,M)

   When data contains a measured input
   [Gtf,Gnoi,Gio] = SPA(DATA, ... )
   returns the information as 3 different IDFRD models:
   Gtf contains the transfer function estimate from the input to
       the output.
   Gnoi contains the spectrum of the output disturbance v.
   Gio contains the spectrum matrix for the output and input channels
       taken together.
```

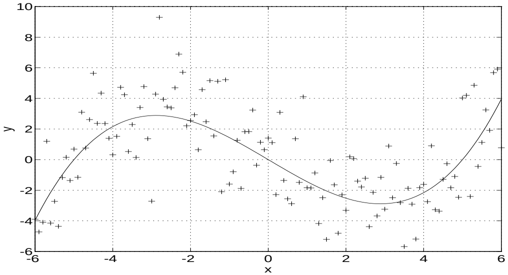
Parametric methods

A simple example on a dynamic system

System model

$$y_t = \frac{1}{1 - aq^{-1}} e_t \qquad\qquad y_t = a y_{t-1} + e_t$$





Guess (model):

$$y_t = \alpha + \beta x_t + \gamma x_t^2 + \kappa x_t^3 + e_t \qquad t = 1, \dots, N$$

$$\theta = (\alpha, \beta, \gamma, \kappa)^\top \qquad \varphi_t^\top = [\ 1\ x_t\ x_t^2\ x_t^3\ ]$$

$$y_t = \varphi_t^\top \theta + e_t \qquad t = 1, \dots, N$$

Consider a scalar system: $y_t = \varphi_t^T \theta + e_t$.

Measure of distance

$$J = \frac{1}{2} \sum_{t=1}^{N} \varepsilon_t^2 \qquad \varepsilon_t = y_t - \varphi_t^T \theta$$

$$I = \sum_{t=1}^{N} \varepsilon_t \varphi_t^T = 0$$

$$H = \sum_{t=1}^{N} \varphi_t \varphi_t^T > 0$$

The Normal equation

$$I^T = \sum_{t=1}^{N} \varphi_t (y_t - \varphi_t^T \theta) = 0$$

LS estimate

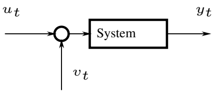
$$\hat{\theta} = \left(\sum_{t=1}^{N} \varphi_t \varphi_t^T \right)^{-1} \sum_{t=1}^{N} \varphi_t y_t$$

DTU

SYSID on external models

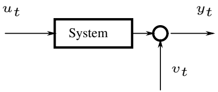
ARMAX structure

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$



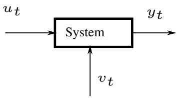
Box-Jenkins structure

$$y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$



L structure

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$



The ARX structure

ARX Models (Equation error models)

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t$$

$$y_t + a_1 y_{t-1} + \cdots + a_n y_{t-n} = b_1 u_{t-1} + \cdots + b_n u_{t-n} + e_t$$

$$y_t = -a_1 y_{t-1} \cdots - a_n y_{t-n} + b_1 u_{t-1} + \cdots + b_n u_{t-n} + e_t$$

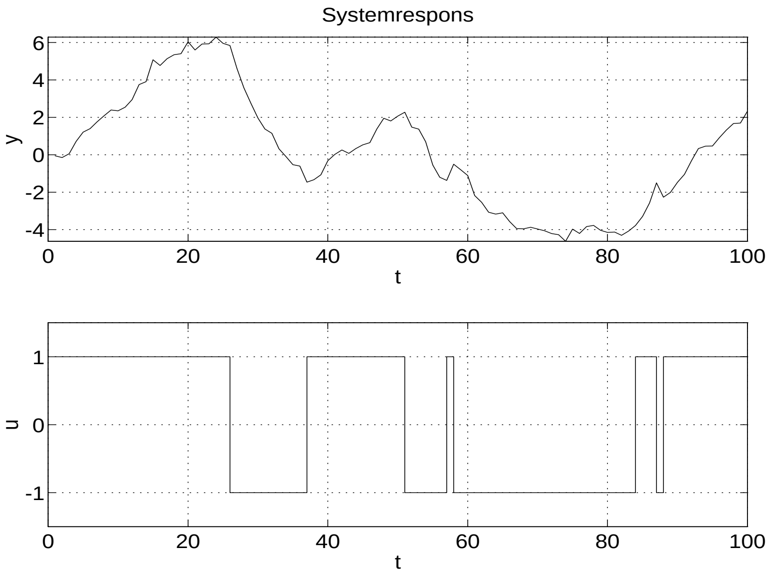
$$\begin{aligned} \varphi_t^T &= [-y_{t-1}, \cdots, -y_{t-n}, u_{t-1}, \cdots u_{t-n}] \\ \theta^T &= [\ a_1, \cdots, a_n, b_1, \cdots, b_n] \end{aligned}$$

$$y_t = \varphi_t^T \theta + e_t$$

$$\hat{\theta} = \left(\sum_{i=1}^{N} \varphi_t \varphi_t^T \right)^{-1} \sum_{i=1}^{N} \varphi_t y_t$$

A simple example

$$y_t - 0.9y_{t-1} = 0.5u_{t-1} + e_t$$



```
>>ns=[1 1 1];
>>th=arx([y u],ns);
>>present(th)
This matrix was created by the command ARX    on 10/6 2012 at 14:37
Loss fcn: 0.09331   Akaike`s FPE: 0.09712 Sampling interval 1
The polynomial coefficients and their standard deviations are
B =
         0    0.4470
         0    0.0351
A =
    1.0000   -0.9187
         0    0.0115

>>[parm,P]=th2par(th);
>>estpres(parm,P)

estimate 99% cinf

   estimat      +/-        ll        ul
   -0.9187    0.0296   -0.9483   -0.8891
    0.4470    0.0904    0.3566    0.5374
```

Analysis of LS estimate (the dynamics case)

System $y_t = \varphi_t^T \theta_0 + e_t$

Model $y_t = \varphi_t^T \hat{\theta} + \varepsilon_t$

$$\hat{\theta} = \theta_0 + \left[\frac{1}{N} \sum_{i=1}^{N} \varphi_i \varphi_i^T \right]^{-1} \left[\frac{1}{N} \sum_{i=1}^{N} \varphi_i e_i \right]$$

$$\rightarrow \theta_0 + \mathbf{E}\left\{ \varphi_t \varphi_t^T \right\}^{-1} \mathbf{E}\left\{ \varphi_t e_t \right\}$$

$$\hat{\theta} \in \mathbf{N}_a \left(\theta_\infty, \frac{\sigma^2}{N} \mathbf{E}\left\{ \varphi_t \varphi_t^T \right\}^{-1} \right)$$

Analysis of LS estimate

$$\hat{\theta} \to \theta_0 + \mathbf{E}\Big\{\varphi_t \varphi_t^T\Big\}^{-1} \mathbf{E}\Big\{\varphi_t e_t\Big\}$$

When is θ_0 the convergence point

$$\text{A}: \quad \mathbf{E}\Big\{\varphi_t e_t\Big\} = 0$$

$$\text{B}: \quad \mathbf{E}\Big\{\varphi_t \varphi_t^T\Big\} > 0$$

A: e_t has to be uncorellated with old output and input.

B: if u_t is $pe(n)$ (then B).

Persistently excitation

$$r_k = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbf{E}\left\{ u_t u_{t-k} \right\}$$

$$R_m = \begin{bmatrix} r_0 & r_1 & \cdots & r_{m-1} \\ r_1 & r_0 & & \\ \vdots & \vdots & & \vdots \\ r_{m-1} & r_{m-2} & \cdots & r_0 \end{bmatrix}$$

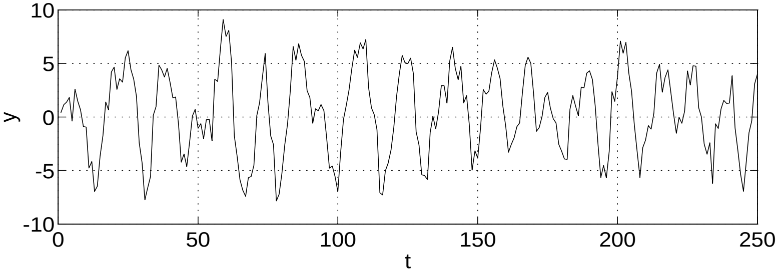
$u_t \quad pe(m)$ if $\quad R_m > 0$

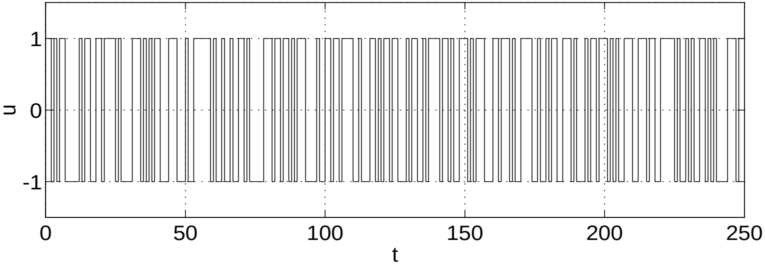
Ex 1: $u_t \in \mathbb{F}_{iid}\,(0)\,\sigma^2 \quad \rightarrow \quad pe(\infty)$

Ex 2: $u_t = const \quad \rightarrow \quad pe(1).$

Example:

$$y_t = q^{-1}\frac{0.90 + 0.085q^{-1}}{1 - 1.66q^{-1} + 0.83q^{-2}}u_t + e_t$$





$$y_t = q^{-1}\frac{0.90 + 0.085q^{-1}}{1 - 1.66q^{-1} + 0.83q^{-2}}u_t + e_t$$

ARX model:

$$(1 + a_1q^{-1} + a_2q^{-1})y_t = q^{-1}(b_0 + b_1q^{-1})u_t + \varepsilon_t$$

System description as AR(MA)X structure:

$$(1 - 1.66q^{-1} + 0.83q^{-2})y_t = q^{-1}(0.90 + 0.085q^{-1})u_t + \ (1 - 1.66q^{-1} + 0.83q^{-2})e_t$$

Residuals for correct estimate (of A and B parameters).

$$\varepsilon_t = (1 - 1.66q^{-1} + 0.83q^{-2})e_t$$

```
>>ns=[2 2 1];
>>th=arx([y u],ns);
>>present(th)
This matrix was created by the command ARX    on 10/25 2012 at 13:18
Loss fcn: 2.934   Akaike's FPE: 3.029 Sampling interval 1
The polynomial coefficients and their standard deviations are
B =
         0    0.9193    0.7678
         0    0.1100    0.1200
A =
    1.0000   -0.9767    0.1807
         0    0.0551    0.0552
>>[parm,p]=th2par(th);
>>estpres(parm,p);
estimate 99% cinf
   estimat      +/-        ll        ul
   -0.9767    0.1420   -1.1187   -0.8347
    0.1807    0.1421    0.0386    0.3228
    0.9193    0.2834    0.6359    1.2027
    0.7678    0.3091    0.4588    1.0769
```

To be compared with system:

$$(1 - 1.66q^{-1} + 0.83q^{-2})y_t = q^{-1}(0.90 + 0.085q^{-1})u_t$$

Optimization and SYSID

Optimization - a short primer

$$\hat{\theta} = \ arg\ Min\ J(\theta) \qquad \hat{\theta} = \ sol \left\{ \ I(\theta) = 0 \ \right\} \qquad I = \frac{\partial}{\partial \theta} J$$

$$J = J_0 + g_0^\top(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)^\top H_0(\theta - \theta_0) + \dots$$

$$I = g_0^T + (\theta - \theta_0)^T H_0 + \dots$$

$$\hat{\theta} = \theta_0 - H_0^{-1} g_0$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k - H_k^{-1} g_k$$

$$J = \sum_{t=1}^{N} \frac{1}{2} \varepsilon_t^2 \qquad \varepsilon_t = y_t - \varphi^\top \hat{\theta}_k$$

$$g_k = -\sum_{t=1}^{N} \varphi_t \varepsilon_t \qquad H_k = \sum_{t=1}^{N} \varphi_t \varphi_t^\top$$

$$\hat{\theta}_{k+1} = \theta_k - H_k^{-1} g_k = \hat{\theta}_k + \left(\sum_{t=1}^{N} \varphi_t \varphi_t^\top \right)^{-1} \left(\sum_{t=1}^{N} \varphi_t \varepsilon_t \right)$$

$$\hat{\theta} = \left(\sum_{t=1}^{N} \varphi_t \varphi_t^\top \right)^{-1} \left(\sum_{t=1}^{N} \varphi_t y_t \right) \qquad \text{for} \qquad \hat{\theta}_k = 0$$

In general, φ might depend on θ

$$\hat{\theta}_{k+1} = \hat{\theta}_k - H_k^{-1} g_k$$

$$J = \sum_{t=1}^{N} \frac{1}{2}\varepsilon_t^2 \qquad \varepsilon_t = y - \hat{y}_{t|t-1}(\hat{\theta}_k) \qquad \psi_t^T = \frac{\partial}{\partial\theta}\hat{y}_{t|t-1} = -\frac{\partial}{\partial\theta}\varepsilon_t$$

$$g_k = -\sum_{t=1}^{N} \psi_t \varepsilon_t \qquad H_k = \sum_{t=1}^{N} \psi_t \psi_t^T - \left(\frac{\partial}{\partial\theta}\psi_t\right)\varepsilon_t$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \left[\sum_{t=1}^{N} \psi_t \psi_t^T\right]^{-1} \sum_{t=1}^{N} \psi_t \varepsilon_t$$

ARX

Lesson learned - L17

1. System identification
2. Non-parametric methods
3. ARX and LS (PEM)
4. Identification and optimization

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

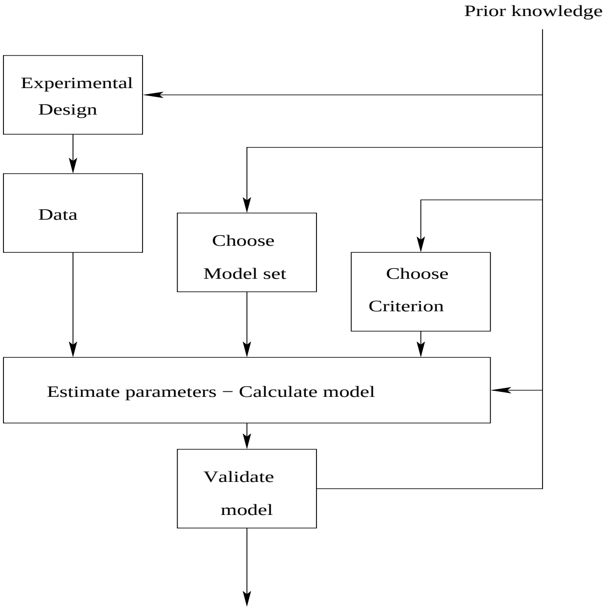
System identification (L18)

Outline - Learning objectives

With a view to recursive methods

- Resume of L17 (id of ARX models)
- Short primer on optimization/minimization
- ARX and then ?
- IV, OE and ARMAX
- L-structure (PEM, PLR methods)

The modelling process



ARX Models (from L17)

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t$$

$$y_t + a_1 y_{t-1} + \cdots + a_n y_{t-n} = b_1 u_{t-1} + \cdots + b_n u_{t-n} + e_t$$

$$\begin{aligned} \varphi_t^T &= [-y_{t-1}, \cdots, -y_{t-n}, u_{t-1}, \cdots, u_{t-n}] \\ \theta^T &= [\; a_1, \cdots, a_n, b_1, \cdots, b_n \;] \end{aligned}$$

$$y_t = \varphi_t^T \theta + e_t = \hat{y}_{t|t-1} + e_t$$

The LS estimate minimize

$$J = \frac{1}{2}\sum_{t=1}^{N} \varepsilon_t^2 \qquad \varepsilon_t = y_t - \varphi_t^T \hat{\theta} = \hat{A}(q^{-1})y_t - \hat{B}(q^{-1})u_t$$

or

$$\hat{\theta} = arg\ min \left[\frac{1}{2}\sum_{t=1}^{N} \varepsilon_t^2\right] = \left[\sum_{t=1}^{N} \varphi_t \varphi_t^T\right]^{-1} \left[\sum_{t=1}^{N} \varphi_t y_t\right]$$

Scalar Xmodels

$$y_t = G(q)u_t + v_t$$

$$v_t = H(q)e_t$$

$$e_t : f(e)$$

Parameterization:

$\theta \in G,\ H$ and f

Structures

ARX:

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t$$

ARMAX:

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$

Box-Jenkins:

$$y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

L-structure:

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

The Normal situation:

$$e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

Prediction Error Methods (PEM)

$$J = \sum_{i=1}^{N} \frac{1}{2}\varepsilon_i^2 \qquad \varepsilon_i = y_i - \hat{y}_{i|i-1}(\hat{\theta})$$

Gauss-Newton

$$\hat{\theta} = \ arg\ Min\ J(\theta)$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k - H_k^{-1} g_k$$

$$g_k = \frac{\partial}{\partial\theta} J \qquad H_k = \frac{\partial^2}{\partial^2\theta} J$$

PEM

$$J = \sum_{i=0}^{N} \frac{1}{2}\varepsilon_i^2 \qquad \varepsilon_i = y_i - \hat{y}_{i|i-1}(\hat{\theta}_k)$$

$$g_k = -\sum_{i=0}^{N} \psi_i \varepsilon_i \qquad \psi_t^T = -\frac{\partial}{\partial\theta}\varepsilon_i$$

$$H_k = \sum_{i=0}^{N} \psi_i \psi_i^T - \left(\frac{\partial}{\partial\theta}\psi_i\right)\varepsilon_i$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \left[\sum_{i=0}^{N} \psi_i \psi_i^T\right]^{-1} \sum_{i=0}^{N} \psi_i \varepsilon_i$$

For a specific model structure we need to know how

ε_i and ψ_i

are obtained.

$$Ay_t = Bu_t + e_t \qquad\qquad y_t = \varphi_t^T \theta + e_t$$

If we apply an prediction error method, ie. minimizing

$$J = \sum_{i=1}^{N} \frac{1}{2}\varepsilon_i^2 \qquad \varepsilon_i = y_i - \hat{y}_{i|i-1} = \hat{A}y_i - \hat{B}u_i = y_i - \varphi_i^T \hat{\theta}$$

the estimate of the parameters **can** be determined with the following iterations:

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \left[\sum_{i=1}^{N} \psi_i \psi_i^T\right]^{-1} \times \sum_{i=1}^{N} \psi_i \varepsilon_i$$

$$\psi_i^T = \frac{\partial}{\partial \hat{\theta}} \hat{y}_{i|i-1} = -\frac{\partial}{\partial \hat{\theta}} \varepsilon_i = \varphi_i^T$$

DTU

ARMAX - PEM

$$Ay_t = Bu_t + Ce_t$$

If we apply an prediction error method, ie. minimizing

$$J = \sum_{i=1}^{N} \frac{1}{2}\varepsilon_i^2 \qquad \varepsilon_i = y_i - \hat{y}_{i|i-1} = \frac{\hat{A}y_i - \hat{B}u_i}{\hat{C}}$$

the estimate of the parameters **can** be determined with the following iterations:

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \left[\sum_{i=1}^{N} \psi_i \psi_i^T\right]^{-1} \times \sum_{i=1}^{N} \psi_i \varepsilon_i$$

$$\psi_i^T = \frac{\partial}{\partial \hat{\theta}} \hat{y}_{i|i-1} = -\frac{\partial}{\partial \hat{\theta}} \varepsilon_i$$

$$\psi_t^T = -\frac{\partial}{\partial \theta}\varepsilon_t$$

$$C\varepsilon_t = Ay_t - Bu_t$$

or (as an example):

$$\varepsilon_t + c_1\varepsilon_{t-1} = y_t + a_1 y_{t-1} - b_0 u_t - b_1 u_{t-1}$$

$$\begin{aligned} C\frac{\partial}{\partial a_i}\varepsilon_t &= y_{t-i} \\ C\frac{\partial}{\partial b_i}\varepsilon_t &= -u_{t-i} \\ \varepsilon_{t-i} + C\frac{\partial}{\partial c_i}\varepsilon_t &= 0 \end{aligned}$$

$$\psi_t = \frac{1}{C(q^{-1})}\hat{\varphi}_t$$

The ARMAX structure

Consider the ARMAX structure:

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$

where

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \\ B(q^{-1}) &= b_0 + b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c} \end{aligned}$$

This description can easily be into

$$\begin{aligned} y_t &= -a_1 y_{t-1} - \cdots - a_{n_a} y_{t-n_a} \\ &\quad + b_0 u_t + b_1 u_{t-1} + \cdots + b_{n_b} u_{t-n_b} \\ &\quad + c_1 e_{t-1} + \cdots + c_{n_c} e_{t-n_c} \\ &\quad + e_t \end{aligned}$$

If the vectors:

$$\begin{aligned} \varphi_t &= \left(-y_{t-1}, \cdots, -y_{t-n_a}, \ u_t, \ u_{t-1}, \ \cdots, \ u_{t-n_b}, \ e_{t-1}, \ \cdots, \ e_{t-n_c} \right)^T \\ \theta &= (\ a_1, \ \cdots, \ a_{n_a}, \ b_0, \ b_1, \ \cdots, \ b_{n_b}, \ c_1, \ \cdots, \ c_{n_c})^T \end{aligned}$$

are introduced then the system can be given is the aggressor form:

$$y_t = \varphi_t^T \theta + e_t$$

There is just one slight problem. The regressors e_{t-i} is not known.

ARMAX

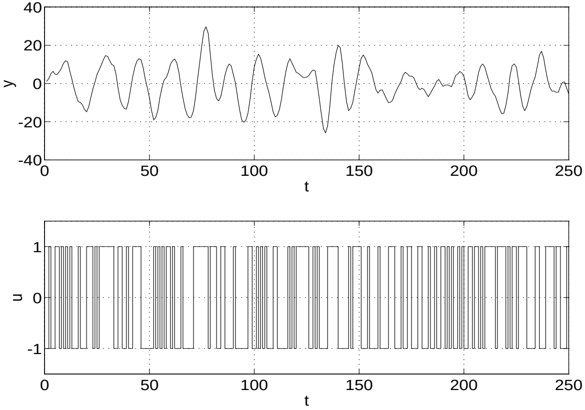
The regressors e_{t-i} are unknown, but ε_{t-i} is a good estimate (if the parameter estimate is good).

$$\hat{\varphi}_t = (-y_{t-1}, \cdots, -y_{t-n_a}, \\ u_t, u_{t-1}, \cdots, u_{t-n_b}, \\ \varepsilon_{t-1}, \cdots, \varepsilon_{t-n_c})^T$$

$$\varepsilon_i = y_i - \hat{\varphi}_i^T \hat{\theta}_k = \frac{Ay_i - Bu_i}{C}$$

Example

$$(1 - 1.66q^{-1} + 0.83q^{-2})y_t = q^{-1}(0.9 + 0.085q^{-1})u_t$$
$$+(1 + 0.8q^{-1} + 0.4q^{-2})e_t$$



$$(1 + a_1q^{-1} + a_2q^{-2})y_t = q^{-1}(b_0 + b_1q^{-1})u_t + e_t$$

DTU

```
ns=[2 2 1];
th=arx([y u],ns)
Discrete-time IDPOLY model: A(q)y(t) = B(q)u(t) + e(t)
A(q) = 1 - 1.762 q^-1 + 0.9061 q^-2

B(q) = 0.8792 q^-1 + 0.07356 q^-2

Estimated using ARX
Loss function 1.49773 and FPE 1.54644
Sampling interval: 1

[parm,p]=th2par(th);
estpres(parm,p);
estimate 99% cinf
   estimate      +/-       ll        ul
   -1.7618    0.0580   -1.8198   -1.7038
    0.9061    0.0583    0.8478    0.9644
    0.8792    0.2040    0.6752    1.0832
    0.0736    0.2100   -0.1364    0.2835
```

$$(1 - 1.66q^{-1} + 0.83q^{-2})y_t = q^{-1}(0.9 + 0.085q^{-1})u_t + (1 + 0.8q^{-1} + 0.4q^{-2})e_t$$

Example

$$(1 + a_1 q^{-1} + a_2 q^{-2}) y_t = q^{-1}(b_0 + b_1 q^{-1}) u_t + (1 + c_1 q^{-1} + c_2 q^{-2}) e_t$$

```
ns=[2 2 2 1];
th=armax([y u],ns)
Discrete-time IDPOLY model: A(q)y(t) = B(q)u(t) + C(q)e(t)
A(q) = 1 - 1.643 q^-1 + 0.7982 q^-2

B(q) = 0.9132 q^-1 + 0.199 q^-2

C(q) = 1 + 0.8649 q^-1 + 0.4143 q^-2

Estimated using ARMAX
Loss function 0.909374 and FPE 0.955418
Sampling interval: 1
```

Example

```
[parm,p]=th2par(th);
estpres(parm,p);

estimate 99% cinf

   estimat      +/-       ll        ul
   -1.6429    0.1007   -1.7435   -1.5422
    0.7982    0.0999    0.6982    0.8981
    0.9132    0.1462    0.7670    1.0595
    0.1990    0.1496    0.0494    0.3485
    0.8649    0.1680    0.6968    1.0329
    0.4143    0.1677    0.2466    0.5820
```

$$(1 - 1.66q^{-1} + 0.83q^{-2})y_t = q^{-1}(0.9 + 0.085q^{-1})u_t + (1 + 0.8q^{-1} + 0.4q^{-2})e_t$$

The L-structure

Consider the L-structure:

$$Ay_t = \frac{B}{F}u_t + \frac{C}{D}e_t + d \qquad \hat{\theta} = arg\min_{\theta}\left\{\frac{1}{2}\sum_{t=1}^{N}\varepsilon_t^2\right\}$$

$$\varepsilon_t = \frac{D}{C}\left[Ay_t - \frac{B}{F}u_t - d\right]$$

$$\psi_t = \nabla_\theta \hat{y}_t = -\nabla_\theta \varepsilon_t$$

$$\psi_t = \left(-\check{y}_{t-1}, \ldots, \check{u}_t, \ldots, -\check{y}^u_{t-1}, \ldots, \check{e}_{t-1}, \ldots, -\check{y}^e_{t-1}, \ldots, \delta\right)^T$$

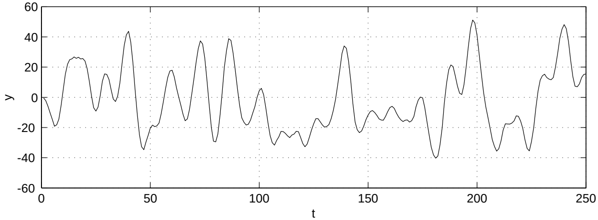
$$\check{y}_t = \frac{D}{C}y_t \qquad \check{u}_t = \frac{D}{CF}u_t \qquad \check{y}^u_t = -\frac{D}{CF}w_t$$

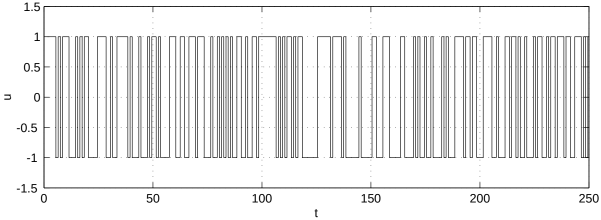
$$\check{e}_t = \frac{1}{C}\varepsilon_t \qquad \check{y}^e_t = -\frac{1}{C}\eta_t \qquad \delta = \frac{D}{C}1$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \left[\sum_{i=1}^{N}\psi_i\psi_i^T\right]^{-1} \times \sum_{i=1}^{N}\psi_i\varepsilon_i \qquad P_N = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i^2 \times \left[\sum_{i=1}^{N}\psi_i\psi_i^T\right]^{-1}$$

Example

$$(1 - 1.66q^{-1} + 0.83q^{-2})y_t = q^{-1}\frac{0.9 + 0.085q^{-1}}{1 + 0.5q^{-1}}u_t + \frac{1 + 0.8q^{-1} + 0.4q^{-2}}{1 - 0.8q^{-1}}e_t$$





Example

```
ns=[2 2 2 1 1 1];
th=pem([y u],ns)
Discrete-time IDPOLY model: A(q)y(t) = [B(q)/F(q)]u(t) + [C(q)/D(q)]e(t)
A(q) = 1 - 1.605 q^-1 + 0.838 q^-2

B(q) = 0.7857 q^-1 - 0.04615 q^-2

C(q) = 1 + 0.8219 q^-1 + 0.3656 q^-2

D(q) = 1 - 0.8786 q^-1

F(q) = 1 + 0.4723 q^-1

Estimated using PEM
Loss function 0.806865 and FPE 0.86251
Sampling interval: 1
```

Example

```
[parm,p]=th2par(th);
estpres(parm,p);

estimate 99% cinf

   estimat       +/-        ll         ul
   -1.6050     0.1013    -1.7063    -1.5037
    0.8380     0.1010     0.7370     0.9391
    0.7857     0.0962     0.6895     0.8819
   -0.0462     0.1973    -0.2434     0.1511
    0.8219     0.1858     0.6360     1.0077
    0.3656     0.1827     0.1830     0.5483
   -0.8786     0.1009    -0.9795    -0.7777
    0.4723     0.1074     0.3648     0.5797

compare([y u],th,1);
```

ML estimation

$$Y_t = (y_1, \ldots, y_t)^T$$

Likelihood function

$$\mathcal{L}(\theta; Y_t) = f(Y_t|\theta)$$

$$\hat{\theta} = arg \max_{\theta} f(Y_t|\theta)$$

$$\hat{\theta} = arg \min_{\theta} \left\{ -log\ f(Y_t|\theta) \right\}$$

The iid case

$$f(Y_t|\theta) = \prod_{i=1}^{t} f(y_i|\theta)$$

The dynamic case

$$f(Y_t|\theta) = f(y_t|Y_{t-1}, \theta) f(Y_{t-1}, \theta)$$

$$= f(y_t|Y_{t-1}, \theta) f(y_{t-1}|Y_{t-2}, \theta) f(Y_{t-2}, \theta)$$

$$= \prod_{i=1}^{t} f(y_i|Y_{i-1}, \theta)$$

$$\hat{\theta} = arg \min_{\theta} \left\{ -log \left(\prod_{i=1}^{t} f(y_i|Y_{i-1}, \theta) \right) \right\} = arg \min_{\theta} \left\{ \sum_{i=1}^{t} \left(-log \left[f(y_i|Y_{i-1}, \theta) \right] \right) \right\}$$

ML estimation

A lot of external models (ARX, ARMAX, BJ, OE and L-structure) can be put into the regression form:

$$y_i = \varphi_i^T \theta + e_i \qquad e_i \in \mathbf{N}\left(0, \sigma^2\right)$$

That means that

$$y_i | Y_{i-1}, \theta \in \mathbf{N}\left(\varphi_i^T \theta, \sigma^2\right)$$

Notice, the parameters enters via the mean value.

or that:

$$f(y_i | Y_{i-1}, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2}\frac{1}{\sigma^2}(y_i - \varphi_i^T \theta)^2\right\} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\varepsilon_i^2}{2\sigma^2}\right\}$$

ie. the LS and ML estimator coinside for the Gaussian case.

What if the destribution is double exponential ?

What if the destribution is a worst case ?

What if the parameters are in higher order moments ?

DTU

Variation of the ARX structure

Variation of the ARX structure

Problem: We have a system (model) with the structure:

$$y_t = \frac{B(q^{-1})}{A(q^{-1})} u_t + \xi_t$$

Use some variant of LS on

$$A(q^{-1}) y_t = B(q^{-1}) u_t + v_t$$

Results in IV method

We can apply an output error strategy:

$$Min \ \| y_t - \frac{B(q^{-1})}{A(q^{-1})} u_t \|^2$$

Results in OE method

Or model the disturbances and estimate the parameters.

$$A(q^{-1}) y_t = B(q^{-1}) u_t + C(q^{-1}) e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

Results in PEM or PLR method on an ARMAX model

IV method

Consider the system

$$A(q^{-1})y_t = B(q^{-1})u_t + v_t$$

and the LS estimate scheme:

$$\varepsilon_i = y_i - \varphi_i^T \hat{\theta}_n$$

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \left[\sum_{i=1}^{N} \varphi_i \varphi_i^T\right]^{-1} \left[\sum_{i=1}^{N} \varphi_i \varepsilon_i\right] \qquad \hat{\theta} \to \theta_0 + \mathbf{E}\left\{\varphi_t \varphi_t^T\right\}^{-1} \mathbf{E}\left\{\varphi_t v_t\right\}$$

If we instead use the following scheme:

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \left[\sum_{i=1}^{N} \zeta_i \varphi_i^T\right]^{-1} \left[\sum_{i=1}^{N} \zeta_i \varepsilon_i\right] \qquad \hat{\theta} \to \theta_0 + \mathbf{E}\left\{\zeta_t \varphi_t^T\right\}^{-1} \mathbf{E}\left\{\zeta_t v_t\right\}$$

where $\zeta \simeq \varphi$ in some sense.

How to choose ζ_t.

$$\mathbf{E}\left\{\zeta_t v_t\right\} = 0$$

$\mathbf{E}\left\{\zeta_t \varphi_t^T\right\}$ nonsingular

One (out of many) simple choiceis

$$\bar{\zeta}_t^T = \left(-\bar{y}_{t-1}, \, \dots, -\bar{y}_{t-n}, \, u_{t-1}, \, \dots, u_{t-n}\right)$$

where

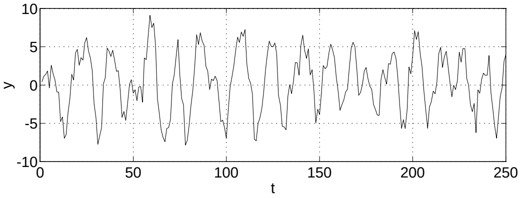
$$\bar{y}_t = \frac{B}{A} u_t$$

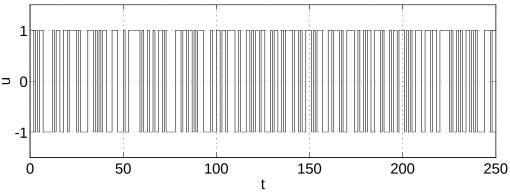
This can be compared with:

$$\varphi_t^T = \left(-y_{t-1}, \, \cdots, \, -y_{t-n}, \, u_{t-1}, \, \cdots, \, u_{t-n}\right)$$

IV method

$$y_t = q^{-1} \frac{0.90 + 0.085q^{-1}}{1 - 1.66q^{-1} + 0.83q^{-2}} u_t + e_t$$





```
>> ideks18
ns=[2 2 1];
th=iv4([y u],ns)
Discrete-time IDPOLY model: A(q)y(t) = B(q)u(t) + e(t)
A(q) = 1 - 1.666 q^-1 + 0.8336 q^-2

B(q) = 0.9233 q^-1 + 0.03516 q^-2

Estimated using IV4
Loss function 1.08084 and FPE 1.11599
Sampling interval: 1

[parm,p]=th2par(th);
estpres(parm,p);

estimate 99% cinf

   estimat      +/-       ll        ul
   -1.6663    0.0325   -1.6988   -1.6338
    0.8336    0.0299    0.8037    0.8634
    0.9233    0.1448    0.7785    1.0681
    0.0352    0.1722   -0.1370    0.2074
```

OE method

Consider a rearrangement of the OE structure:

$$
\begin{aligned}
y_t &= \frac{B}{A}u_t + e_t \\
&= \frac{B}{A}u_t - Bu_t + Bu_t + e_t \qquad\qquad \hat{y}_t = \frac{B}{A}u_t \\
&= (1-A)\hat{y}_t + Bu_t + e_t \\
&= -a_1\hat{y}_{t-1} - \ \dots\ - a_n\hat{y}_{t-n} + b_0u_t + \ \dots\ + b_nu_{t-n} + e_t \\
&= \varphi_t^T\theta + e_t \qquad \hat{y}_t = \varphi_t^T\theta
\end{aligned}
$$

where

$$
\begin{aligned}
\varphi_t^T &= [-\hat{y}_{t-1}, \cdots, -\hat{y}_{t-n}, u_{t-1}, \cdots, u_{t-n}] \\
\theta &= [\ a_1, \cdots, a_n, b_1, \cdots, b_n\ ]
\end{aligned}
$$

DTU

If we apply an PEM strategy, ie. we going to minimize

$$J = \frac{1}{2}\sum_{i=1}^{N}\varepsilon_i^2 \qquad \varepsilon_i = y_t - \varphi_t^T\hat{\theta} = y_t - \frac{B}{A}u_t$$

then the estimate **can** be found with the following iterations:

$$\varepsilon_i = y_i - \varphi_i^T\hat{\theta}_n$$

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \left[\sum_{i=1}^{N}\psi_i\psi_i^T\right]^{-1}\left[\sum_{i=1}^{N}\psi_i\varepsilon_i\right]$$

Here the gradient of the prediction is:

$$\psi_t^T = \frac{1}{A}[\ \dots\ -\hat{y}_{t-i}\ \dots\ u_{t-i}\ \dots\ ]$$

Analysis of the OE method

System $\quad y_t = \varphi_t^T \theta_0 + e_t$

Model $\quad y_t = \varphi_t^T \hat{\theta} + \varepsilon_t$

$$\hat{\theta} \rightarrow \theta_0 + \mathbf{E}\left\{\psi_t \psi_t^T\right\}^{-1} \mathbf{E}\left\{\psi_t e_t\right\}$$

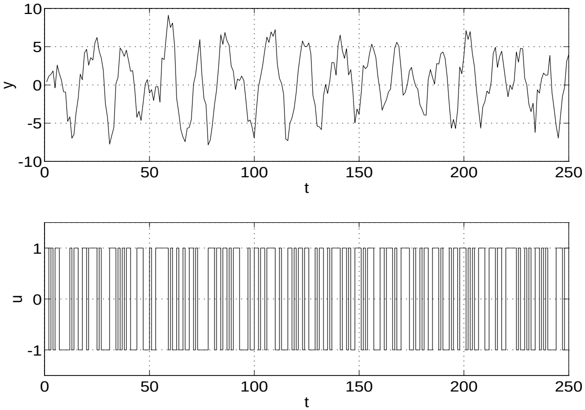
Consistent regardless the color of e_t.

$$\hat{\theta} - \theta_0 \in \mathbf{N}\left(0, \frac{1}{N}\mathbf{E}\left\{\psi_t \psi_t^T\right\}^{-1} \sigma^2\right)$$

$$\hat{P} = \left[\sum_{i=1}^{N} \psi_i \psi_i^T\right]^{-1} \hat{\sigma}^2 \qquad \hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N} \varepsilon_i^2$$

Example

$$y_t = q^{-1}\frac{0.90 + 0.085q^{-1}}{1 - 1.66q^{-1} + 0.83q^{-2}}u_t + e_t$$



```
>> ideks17
ns=[2 2 1];
th=oe([y u],ns)
Discrete-time IDPOLY model: y(t) = [B(q)/F(q)]u(t) + e(t)
B(q) = 0.8901 q^-1 + 0.08341 q^-2

F(q) = 1 - 1.668 q^-1 + 0.8358 q^-2

Estimated using OE
Loss function 0.853822 and FPE 0.882164
Sampling interval: 1

[parm,p]=th2par(th);
estpres(parm,p);

estimate 99% cinf

   estimate      +/-        ll        ul
    0.8901     0.1138     0.7763    1.0039
    0.0834     0.1358    -0.0524    0.2192
   -1.6679     0.0137    -1.6816   -1.6541
    0.8358     0.0116     0.8242    0.8473
```

Pseudo linear regression (PLR) methods

$$y_t = \varphi_t^T \theta + e_t$$

The ARMAX structure

Consider the ARMAX structure:

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$

where

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \\ B(q^{-1}) &= b_0 + b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \cdots + c_{n_c} q^{-n_c} \end{aligned}$$

This description can easily be into

$$\begin{aligned} y_t = & -a_1 y_{t-1} - \cdots - a_{n_a} y_{t-n_a} \\ & + b_0 u_t + b_1 u_{t-1} + \cdots + b_{n_b} u_{t-n_b} \\ & + c_1 e_{t-1} + \cdots + c_{n_c} e_{t-n_c} \\ & + e_t \end{aligned}$$

If the vectors:

$$\varphi_t = \left(-y_{t-1}, \cdots, -y_{t-n_a},\ u_t,\ u_{t-1},\ \cdots,\ u_{t-n_b},\ e_{t-1},\ \cdots,\ e_{t-n_c} \right)^T$$

$$\theta = (a_1, \cdots, a_{n_a}, b_0, b_1, \cdots, b_{n_b}, c_1, \cdots, c_{n_c})^T$$

are introduced then the system can be given is the regressors form:

$$y_t = \varphi_t^T \theta + e_t$$

ARMAX - PLR

Also denoted as Extended Least Squares (ELS).

The regressors e_{t-i} are unknown, but ε_{t-i} is a good estimate (if the parameter estimate is good).

$$\hat{\varphi}_t = (-y_{t-1}, \cdots, -y_{t-n_a}, \\ u_t, u_{t-1}, \cdots, u_{t-n_b}, \\ \varepsilon_{t-1}, \cdots, \varepsilon_{t-n_c})^T$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \left[\sum_{i=1}^{N} \hat{\varphi}_i \hat{\varphi}_i^T \right]^{-1} \times \sum_{i=1}^{N} \hat{\varphi}_i \varepsilon_i \qquad \varepsilon_i = y_i - \hat{\varphi}_i^T \hat{\theta}_k$$

$$\varepsilon_i = y_i - \hat{\varphi}_i^T \hat{\theta}_k = \frac{Ay_i - Bu_i}{C}$$

The L-structure

Consider the L-structure:

$$Ay_t = \frac{B}{F}u_t + \frac{C}{D}e_t + d \qquad \rightarrow \qquad y_t = \varphi_t^T\theta + e_t$$

$$\begin{aligned}
\frac{B}{F}u_t \quad &\equiv \quad w_t = \quad Bu_t + \frac{B}{F}u_t - Bu_t \\
&= \quad Bu_t + (1-F)\frac{B}{F}u_t \\
&= \quad Bu_t + (1-F)w_t
\end{aligned}$$

$$\begin{aligned}
\frac{C}{D}e_t \quad &\equiv \quad \eta_t = Ce_t + (1-D)\eta_t \\
&= \quad (C-1)e_t + (1-D)\eta_t \;\; + \;\; e_t
\end{aligned}$$

$$\begin{aligned}
Ay_t \quad &= \quad x_t \qquad = \frac{B}{F}u_t + \frac{C}{D}e_t + d \\
y_t \quad &= \quad (1-A)y_t + x_t
\end{aligned}$$

$$y_t = (1-A)y_t + Bu_t + (1-F)w_t + (C-1)e_t + (1-D)\eta_t + d \;\; + \;\; e_t$$

DTU

L-structure - PLR

$$y_t = (1-A)y_t + Bu_t + (1-F)w_t + (C-1)e_t + (1-D)\eta_t + d \quad + \quad e_t$$

$$y_t = \varphi_t^T \theta + e_t$$

$$\varphi_t = (-y_{t-1},\ \dots,u_t,\ \dots,\ \ -w_{t-1},\ \dots,e_{t-1},\ \dots,-\eta_{t-1},\ \dots,1)^T$$

$$\theta = (\quad a_1, \dots, \quad b_0, \dots, \quad f_1, \dots, \quad c_1, \dots, \quad d_1, \dots, \quad d)^T$$

$$\hat{w}_t = \frac{\hat{B}}{\hat{F}} u_t \qquad \hat{\eta}_t = \hat{A} y_t - \hat{w}_t - \hat{d} \qquad \varepsilon_t = \frac{\hat{D}}{\hat{C}} \hat{\eta}_t$$

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \left[\sum_{i=1}^{N} \hat{\varphi}_i \hat{\varphi}_i^T \right]^{-1} \times \sum_{i=1}^{N} \hat{\varphi}_i \varepsilon_i$$

$$P_N = \frac{1}{N} \sum_{i=1}^{N} \varepsilon_i^2 \times \left[\sum_{i=1}^{N} \varphi_i \varphi_i^T \right]^{-1}$$

DTU

Continuous time systems

Consider the system

$$y_t = G(p)u_t \qquad p = \frac{d}{dt} \qquad (t \in \mathbb{R})$$

$$G(s) = \frac{b_1 s^{n-1} + \ ... \ b_n}{s^n + a_1 s^{n-1} + \ ... \ a_n} \qquad \theta = (a_1,\ a_2,\ ...\ , a_n,\ b_1,\ ...\ b_n)$$

$$\frac{d^n}{dt^n} y_t + a_1 \frac{d^{n-1}}{dt^{n-1}} y_t + \ .... \ + a_n y_t = b_1 \frac{d^{n-1}}{dt^{n-1}} u_t + \ .... \ + b_n u_t + \xi_t$$

$$p^n y_t + a_1 p^{n-1} y_t + \ .... \ + a_n y_t = b_1 p^{n-1} u_t + \ .... \ + b_n u_t + \xi_t$$

State variable filter method

$$p^n y_t + a_1 p^{n-1} y_t + \;\ldots.\; + a_n y_t = b_1 p^{n-1} u_t + \;\ldots.\; + b_n u_t + \xi_t$$

Let

$$F(p) = p^n + \alpha_1 p^{n-1} + \;\ldots.\; + \alpha_n$$

then

$$\frac{p^n}{F(p)} y_t + a_1 \frac{p^{n-1}}{F(p)} y_t + \;\ldots.\; + a_n \frac{1}{F(p)} y_t = b_1 \frac{p^{n-1}}{F(p)} u_t + \;\ldots.\; + b_n \frac{1}{F(p)} u_t + e_t$$

or

$$z_t = \frac{p^n}{F(p)} y_t = \varphi_t^T \theta + e_t$$

$$\varphi_t^T = \left(\frac{p^{n-1}}{F(p)} y_t, \;\ldots, \frac{1}{F(p)} y_t, \; \frac{p^{n-1}}{F(p)} u_t, \;\ldots, \frac{1}{F(p)} u_t \right)$$

$$\theta^T = (a_1, \;\ldots, a_n, \; b_1, \;\ldots, b_n)$$

The rest is LS.

CD methods

Consider the system

$$y_t = G(p)u_t + H(p)e_t \qquad p = \frac{d}{dt_c}$$

$$G(s) = \frac{b_1 s^{n-1} + \ \dots b_n}{s^n + a_1 s^{n-1} + \ \dots a_n} \qquad H(s) = \frac{c_1 s^{n-1} + \ \dots c_n}{s^n + d_1 s^{n-1} + \ \dots d_n}$$

$$\theta = (a_1,\ a_2,\ \dots, a_n,\ b_1,\ \dots b_n,\ d_1,\ \dots, d_n,\ c_1,\ \dots, c_n)$$

In discrete time (T_s) the system is described by:

$$y_i = G_d(q)u_i + H_d(q)e_i$$

We can then minimize J where

$$\theta \ \rightarrow \ (G, H) \ \rightarrow \ (G_d, H_d) \ \rightarrow \ J = \sum_{i=0}^{N} \varepsilon_i^2$$

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

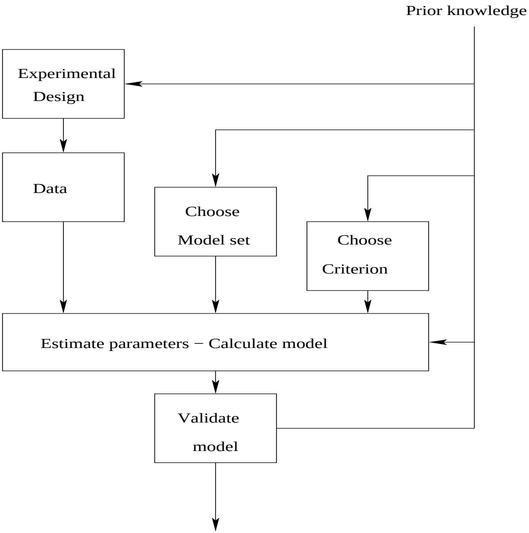
Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

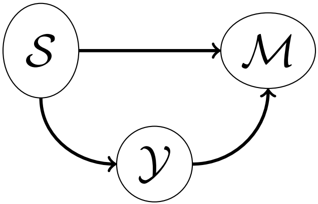
System identification (L19)

Learning objectives

- Multivariable systems (MARX, MARMAX, ...)
- SS estimation (D-time)
- C-time models (SS)


Prior knowledge
Experimental Design
Data
Choose Model set
Choose Criterion
Estimate parameters − Calculate model
Validate model

ID principle



$$\mathcal{D} = \|Y - \mathcal{M}(\theta)\|$$

ID principle

Prediction Error Methods (PEM)

$$J = \sum_{i=1}^{N} \frac{1}{2}\varepsilon_i^2 \qquad\qquad \varepsilon_i \quad \text{is some error}$$

Equation error

$$A(q^{-1})y_t = B(q^{-1})u_t + \varepsilon_t$$

Output error

$$y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \varepsilon_t$$

Prediction error

$$\varepsilon_t = y_t - \hat{y}(t|t-1,\theta) \qquad\qquad \varepsilon_t = A(q^{-1})y_t - B(q^{-1})u_t$$

DTU

Likelihood function

$$\mathcal{L}(\theta) = f(Y_N|\theta) \qquad \hat{\theta} = \ arg \ \max_{\theta} \ f(Y_N|\theta)$$

$$f(Y_N|\theta) = \prod_{i=1}^{N} f(y_i|Y_{i-1},\theta)$$

$$\hat{\theta}: \quad Min \ \sum_{i=1}^{N} -log\left(f(y_i|Y_{i-1},\theta)\right)$$

Mean value models

ARMAX:

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$

Box-Jenkins:

$$y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

L-structure:

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

$$y_t = \hat{y}_{t|t-1} + e_t$$

Linear regressian

$$y_t = \varphi_t^T \theta + e_t$$

Gaussian noise

In the Gaussian case

$$e_t \in \mathbf{N}\left(0, \sigma^2\right)$$

$$y_t | Y_{t-1}, \theta \in \mathbf{N}\left(\hat{y}_{t|t-1}, \sigma^2\right)$$

LS estimate

$$J = \sum_{i=1}^{N} \varepsilon_i^2 \qquad \varepsilon_t = \hat{e}_t$$

$$e_t = \frac{D(q^{-1})}{C(q^{-1})}\left[A(q^{-1})y_t - \frac{B(q^{-1})}{F(q^{-1})}u_t\right]$$

Linear regressian

$$\varepsilon_t = y_t - \varphi_t^T \hat{\theta}$$

PEM methods

For a given modelstructure establish the mapping:

$$\theta \quad \rightarrow \quad J = \sum_{i=1}^{N} \varepsilon_i^2$$

Use a numerical minimization procedure to find:

$$\hat{\theta} = arg \min_{\theta} \sum_{i=1}^{N} \varepsilon_i^2$$

E.g. the Gauss-Newton type iterations:

$$\hat{\theta}_{i+1} = \hat{\theta}_i + \Big[\sum_{i=1}^{N} \psi_i \psi_i^T\Big]^{-1} \sum_{i=1}^{N} \psi_i \varepsilon_i \qquad\qquad \psi_i = -\frac{d}{d\theta}\varepsilon_i$$

$$\hat{\theta} \in \mathbf{N}\left(\theta, \Big[\sum_{i=1}^{N} \psi_i \psi_i^T\Big]^{-1} \sigma^2\right)$$

ML

The Gaussian case:

$$J = \sum_{i=1}^{N} \varepsilon_i^2$$

The LaPlacian case:

$$J = \sum_{i=1}^{N} |\varepsilon_i|$$

The general case

$$J = \sum_{i=1}^{N} \rho(\varepsilon_i)$$

General ML method

$$\rho(\varepsilon_i) = -log\left(f(y_i|Y_{i-1}, \theta)\right)$$

Other methods and structure:

```
ARX, OE, IV4 and IV
```

Alternative method for handling the data:

```
dat=iddata(y,u,Ts);

  Data manipulation.
    iddata      - Construct a data object.
    detrend     - Remove trends from data sets.
    idfilt      - Filter data through Butterworth filters.
    idinput     - Generates input signals for identification.
    merge       - Merge several experiments.
    misdata     - Estimate and replace missing input and output data.
    resample    - Resamples data by decimation and interpolation.
```

Method for defining the structure:

```
ns=[na nb k]; % na(ny*ny), nb(ny*nu), k(ny,nu)
m=arx(dat,ns);
```

and alternative

```
mi=idpoly([1 -1.5 0.7],[1 0.5],1);
m=arx(dat,mi);
```

In general for transfer models

```
m=idpoly(A,B,C,D,F,s2,Ts) % Ts=0 means Continous time models
```

In general:

```
Model structure creation.
  idpoly     - Construct a model object from given polynomials.
  idss       - Construct a state space model object.
  idarx      - Construct a multivariable ARX model object.
  idgrey     - Construct a user-parameterized model object.
```

Multivariate Systems

MISO systems

$$A(q^{-1})y_t = \frac{B_1(q^{-1})}{F_1(q^{-1})}u_1(t-k_1) + \frac{B_2(q^{-1})}{F_2(q^{-1})}u_2(t-k_2) + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

$$th = pem(z, [n_a, n_b, n_c, n_d, n_f, n_k])$$

$$z = [y\ u_1\ u_2\ \dots\ u_m]$$

$$n_b = [n_{b_1}\ n_{b_2}\ \dots\ n_{b_m}] + 1$$

$$n_f = [n_{f_1}\ n_{f_2}\ \dots\ n_{f_m}]$$

$$n_k = [k_1\ k_2\ \dots\ k_m]$$

MARX - systems

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t \qquad y_t \in \mathbb{R}^{n_y}$$

$$A(q^{-1}) = I + A_1 q^{-1} + \ \dots \ + A_n q^{-n}$$

$$B(q^{-1}) = B_0 + B_1 q^{-1} + \ \dots \ + B_n q^{-n}$$

$$A = [eye(n_y),\ A_1,\ \dots,\ A_n]$$

$$B = [B_0,\ B_1,\ \dots,\ B_n]$$

$$thi = idarx(A, B, Ts)$$

$$th = arx([y\ u], thi)$$

MARX - systems

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t$$

$$A(q^{-1}) = I + A_1 q^{-1} + \ \dots \ + A_n q^{-n}$$

$$B(q^{-1}) = B_1 q^{-1} + \ \dots \ + B_n q^{-n}$$

$$\theta^T = \begin{bmatrix} A_1 & A_2 & \dots & A_n & B_1 & \dots & B_n \end{bmatrix} \qquad \varphi_t = \begin{bmatrix} -y_{t-1} \\ \vdots \\ -y_{t-n} \\ u_{t-1} \\ \vdots \\ u_{t-n} \end{bmatrix}$$

$$y_t = \theta^T \varphi_t + e_t$$

If R_e is a diagonal matrix, this can be regarded as p different regressions model with the same regressors.

MARMAX - systems

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$

Multivariate L - systems

$$A(q^{-1})y_t = F^{-1}(q^{-1})B(q^{-1})u_t + D(q^{-1})^{-1}C(q^{-1})e_t$$

Timevarying systems

Consider (for the sake of simplicity)

$$A_t(q^{-1})y_t = B_t(q^{-1})u_t + e_t \qquad t \in \mathbb{Z}$$

or even more simple

$$y_t = -ay_{t-1} + b_t u_{t-1} + e_t$$

Linear trend

$$b_t = b_0 + b_1 t$$

$$y_t = \varphi_t^T \theta + e_t \qquad \varphi_t^T = (-y_{t-1},\ u_{t-1},\ tu_{t-1}) \qquad \theta^T = (a,\ b_0,\ b_1)$$

Periodic trend

$$b_t = b_0 + b_{11} cos(\omega t) + b_{12} sin(\omega t) + b_{21} cos(2\omega t) + b_{22} sin(2\omega t) + \ \ldots$$

$$y_t = \varphi_t^T \theta + e_t$$

$$\varphi_t^T = (-y_{t-1},\ u_{t-1},\ u_{t-1} cos(\omega t),\ u_{t-1} sin(\omega t),\ u_{t-1} cos(2\omega t),\ u_{t-1} sin(2\omega t))$$

$$\theta^T = (a,\ b_0,\ b_{11},\ b_{12},\ b21,\ b_{22})$$

State space Systems

$$
\begin{aligned}
x_{t+1} &= Ax_t + Bu_t \qquad\qquad x_{t_0} = x_0 \\
y_t &= Cx_t + Du_t + e_t
\end{aligned}
$$

Parameters in

$$A = A(\theta) \quad B = B(\theta) \quad C = C(\theta) \quad D(\theta) \qquad x_0 = x_0(\theta)$$

Then the mapping from θ to

$$J(\theta) = \sum_{i=0}^{N} \varepsilon_i^2 \qquad\qquad \varepsilon_t = y_t - \hat{y}_t(\theta)$$

is to be minimized by means of some numerical methods.

Continuous time model

$$\dot{x}_t = A_c x_t + B_c u_t \qquad\qquad x_{t_0} = x_0$$

$$y_t = Cx_t + Du_t + \xi_t$$

Then there is a mapping from parameters in A_c to A or a numerical method for solving ODE can be applied.

SS-models

```
M = IDSS(A,B,C,D)
M = IDSS(A,B,C,D,K,X0,Ts)
M = IDSS(A,B,C,D,K,X0,Ts,'Property',Value,..)
M = IDSS(MOD)
    Where MOD is any SS, TF, or ZPK model or any IDMODEL object.

M: returned as a model structure object describing the discrete-
   time model

   x[k+1] = A x[k] + B u[k] + K e[k] ;      x[0] = X0
     y[k] = C x[k] + D u[k] + e[k]

A,B,C,D and K are the state-space matrices.  X0 is the initial
condition, if any, and Ts is the sampling time.  For Ts == 0, a
continuous-time model is constructed.
```

$$\dot{x} = \begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{bmatrix} x + \begin{bmatrix} \theta_3 \\ \theta_4 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 1 \end{bmatrix} x$$

```
A = [NaN, 0; 0, NaN]; B = [NaN;NaN]; C=[1, 1]; D = 0
Ts=0;
m0 = idss(A,B,C,D,Ts);
m = pem(data,m0);
```

SS - id

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \in \mathbb{F}_{iid}(0, R_1\sigma^2)$$

$$y_t = Cx_t + Du_t + e_t \qquad e_t \in \mathbb{F}_{iid}(0, R_2\sigma^2)$$

and

$$x_{t_0} \in \mathbb{F}(m_0, P_0)$$

Parameters:

$$A = A(\theta) \qquad B = B(\theta) \qquad C = C(\theta) \qquad D(\theta)$$

$$R_1 = R_1(\theta) \qquad R_2 = R_2(\theta) \qquad \sigma^2 = \sigma^2(\theta)$$

$$m_0 = m_0(\theta) \qquad P_0 = P_0(\theta)$$

Maximum likelihood:

$$\hat{\theta} = \ arg\ Max\ f(Y_N|\theta)$$

$$f(Y_N|\theta) = \prod_{i=1}^{N} f(y_i|Y_{i-1}, \theta)$$

$$\hat{\theta} : \quad Min \ \sum_{i=1}^{N} -log\,(f(y_i|Y_{i-1}, \theta))$$

The Kalman filter

Time update

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t$$

$$P_{t+1|t} = AP_{t|t}A^\top + R_1$$

Data updating

$$\begin{aligned} \hat{x}_{t+1|t+1} &= \hat{x}_{t+1|t} + \kappa_{t+1}\left[y_{t+1} - C\hat{x}_{t+1|t}\right] \\ \kappa_{t+1} &= P_{t+1|t}C^T[CP_{t+1|t}C^T + R_2]^{-1} \\ P_{t+1|t+1} &= [I - \kappa_{t+1}C]P_{t+1|t} \end{aligned}$$

$$x_t|Y_{t-1} \in \mathbf{N}\left(\hat{x}_{t|t-1}, P_{t|t-1}\right)$$

$$y_t = Cx_t + e_t$$

$$y_t|Y_{t-1} \in \mathbf{N}\left(C^\top\hat{x}_{t|t-1}, CP_{t|t-1}C^T + R_2\right)$$

$$\in \mathbf{N}\left(\hat{y}_{t|t-1}, Q_{t|t-1}\right)$$

$$y_t | Y_{t-1} \in \mathbf{N}\left(\hat{y}_{t|t-1}, Q_{t|t-1}\right)$$

$$\hat{\theta}: \quad Min \sum_{i=1}^{N} -log\left(f(y_i | Y_{i-1}, \theta)\right)$$

$$X \in \mathbf{N}(m, P)$$

$$f(x) = \frac{1}{\sqrt{Det(P)}\left(\sqrt{2\pi}\right)^n} \;\; exp\left(-\frac{1}{2}(x-m)^\top P^{-1}(x-m)\right)$$

$$J(\theta) = \sum_{t=1}^{N} \left(\frac{1}{2} log(Det[Q_{t|t-1}]) + \frac{1}{2} \epsilon_t^T Q_{t|t-1}^{-1} \epsilon_t \right)$$

DTU

Continuous time SS ID

$$\dot{x}_{t_c} = A(\theta)x_{t_c} + B(\theta)u_{t_c} + v_{t_c} \qquad v_{t_c} \in N(0, \Sigma_1(\theta))$$

$$y_t = C^\top(\theta)x_{tT_s} + e_t \qquad e_t \in N(0, \Sigma_2(\theta))$$

Sampling to:

$$A = e^{AT} \qquad B = \int_0^{T_s} e^{As} B ds$$

$$C = C \qquad R_2 = \Sigma_2$$

$$R_1 = \int_0^{T_s} e^{As}\Sigma_1(e^{A^\top s})ds$$

in

$$x_{t+1} = Ax_t + Bu_t + v_t \qquad v_t \in \mathbb{F}_{iid}(0, R_1)$$

$$y_t = C^T x_t + Du_t + e_t \qquad e_t \in \mathbb{F}_{iid}(0, R_2)$$

The EKF method

$$x_{t+1} = f(x_t, u_t, \theta, \zeta_t, t) \qquad \zeta_t \in \mathbb{F}(0, \Sigma_1)$$

$$y_t = g(x_t, u_t, \theta, \xi_t, t) \qquad \xi_t \in \mathbb{F}(0, \Sigma_2)$$

Parameter model:

$$\theta_{t+1} = \theta_t + \eta_t \qquad \eta_t \in \mathbb{F}(0, \Sigma_3)$$

Linearized description:

$$\left\{ \begin{array}{c} x_{t+1} \\ \theta_{t+1} \end{array} \right\} = \left[\begin{array}{cc} \nabla_x f & \nabla_\theta f \\ 0 & I \end{array} \right] \left\{ \begin{array}{c} x_t \\ \theta_t \end{array} \right\} + \left\{ \begin{array}{c} \nabla_u f \\ 0 \end{array} \right\} u_t$$

$$+ \left[\begin{array}{cc} \nabla_\zeta f & 0 \\ 0 & I \end{array} \right] \left\{ \begin{array}{c} \zeta_t \\ \eta_t \end{array} \right\}$$

$$y_t = \left(\begin{array}{cc} \nabla_x g^T & \nabla_\theta g^T \end{array} \right) \left\{ \begin{array}{c} x_t \\ \theta_t \end{array} \right\} + \nabla_u g^T u_t + \nabla_\xi g^T \xi_t$$

A Kalman (an extended Kalman) filter for this system will estimate x_t and θ.

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

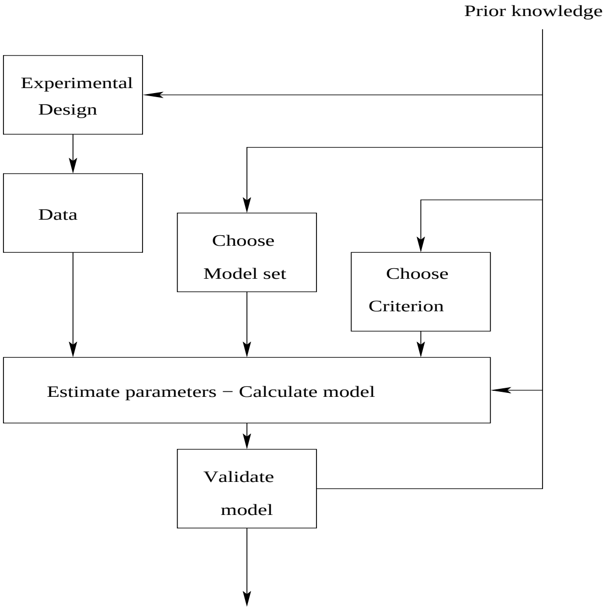
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Model Validation (L20)

Model validation



Model validation

Objective:

- Quality check on model
- Estimation of structural parameters
- Check quality of experiments

- Is the model to complex
- Is the model to simple

$$Measurements = Model + residual$$

$$y_t = \hat{y}_t + \varepsilon_t$$

- Model verification (model analysis)
- Residual analysis

Model analysis

Parameter significance

Results of estimation:

$$\hat{\theta}, \; P = Var\left\{\hat{\theta}\right\}$$

Marginally. Is $\theta_i = 0$?

$$|\hat{\theta}_i| < f^u_{1-\frac{\alpha}{2}}\sqrt{P_{ii}} \qquad H_0 : \quad \frac{\hat{\theta}_i}{\sqrt{P_{ii}}} \in N(0,1) \quad \Big(t(n-p)\Big)$$

Simultaneously. Are the parameters in the vector $\theta_b = 0$ (where $\theta_b \in \theta$)

$$z = \hat{\theta}_b^\top P_b^{-1}\hat{\theta}_b < f^F_{1-\alpha}(m, N-n)$$

N is the number of data, n is the number of parameters in θ and m is the number of parameters in θ_b.

```
>>present(th)
This matrix was created by the command ARX    on 10/25 2012 at 13:59
Loss fcn: 1.546   Akaike`s FPE: 1.648 Sampling interval 1
The polynomial coefficients and their standard deviations are
B =
         0    0.9914    1.0777    0.7099    0.1252
         0    0.0795    0.0963    0.1023    0.0951
A =
    1.0000   -0.5433   -0.3380   -0.0705    0.4161
         0    0.0545    0.0633    0.0633    0.0458
>>[parm,p]=th2par(th);
>>estpres(parm,p);
estimate 99% cinf
   estimat      +/-        ll        ul
   -0.5433    0.1403   -0.6836   -0.4030
   -0.3380    0.1631   -0.5011   -0.1749
   -0.0705    0.1631   -0.2336    0.0926
    0.4161    0.1180    0.2981    0.5341
    0.9914    0.2048    0.7866    1.1962
    1.0777    0.2481    0.8296    1.3257
    0.7099    0.2635    0.4464    0.9734
    0.1252    0.2451   -0.1199    0.3703
```

Pole-Zero cancellation

$$y_t = G(q)u_t + H(q)e_t \qquad G(q) = \frac{(q - z_1)(q - z_2)}{(q - p_1)(q - p_2)(q - p_3)}$$

Is $z_1 = p_3$ **?**

Poles (and zeroes) are determined from parameter (estimates), ie.

$$p_i = f(\theta)$$

then

$$\hat{p}_i = f(\hat{\theta}) = f(\theta) + \frac{\partial}{\partial \theta} f \, \tilde{\theta} + \, ...$$

but

$$\tilde{\theta} \in \mathbf{N}_a \, (0, P)$$

and then

$$\hat{p}_i \in \mathbf{N}_a \left(f(\theta), \frac{\partial}{\partial \theta} f \; P \frac{\partial}{\partial \theta} f^\top \right)$$

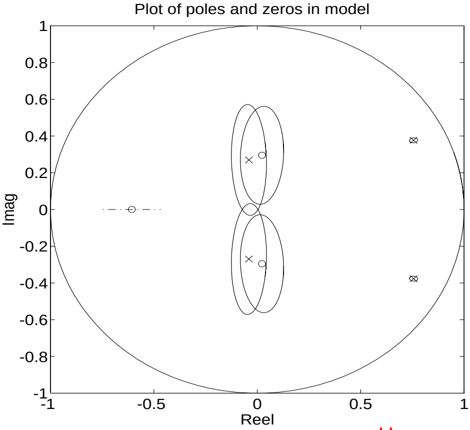
Example

System:

$$(1 - 1.5q^{-1} + 0.7q^{-2})y_t = q^{-1}(1 + 0.5q^{-1})u_t + e_t$$

Model:

$$(1 + a_1 q^{-1} + \cdots + a_4 q^{-4})y_t = q^{-1}(b_0 + \cdots + b_3 q^{-3})u_t + e_t$$



Use pzmap(model,'sd',2.5758);

Analysis of eigenvalues

Basic observation: If model is overparameterized the Hessian (and P) will be ill conditioned.
The same effect if system is not persistently excited.
This effect is less significant for higher noise levels.

$$P = \left(\sum_{i=0}^{N} \psi \psi^\top \right)^{-1} \hat{\sigma}^2$$

$$P v_i = \lambda_i v_i$$

```
[V,D]=eig(P)
```

$$PV = DV$$

Check also condition number.

Example

```
ns=[3 2 1];
th=arx([y u],ns);
[parm,p1]=th2par(th);
[v,d]=eig(p1);
v
v =
    0.4305    0.5618   -0.7028   -0.0074    0.0707
   -0.7241    0.6105    0.0124    0.0854   -0.3093
    0.3657    0.5582    0.6963   -0.0859    0.2498
    0.0004    0.0026    0.0194    0.9625    0.2707
    0.3958   -0.0102    0.1438    0.2428   -0.8739
d
d =
    0.0111         0         0         0         0
         0    0.0000         0         0         0
         0         0    0.0000         0         0
         0         0         0    0.0001         0
         0         0         0         0    0.0001
```

$$\theta_v \simeq 0.43\theta_1 - 0.72\theta_2 + 0.36\theta_3 + 0.39\theta_5$$

Residual analysis

Residual analysis

For correct model:

- $\varepsilon_t \in \mathbf{F}(0, \sigma^2)$
- ε_t have a symmetric distribution
- ε_t is white
- ε_t is uncorrelated with u_t

Correlation analysis

If ε_t is white, then:

$$r_\varepsilon(k) = E\{\varepsilon_{t+k}\varepsilon_t\} = \begin{cases} \sigma^2 & \text{for } \tau = 0 \\ 0 & \text{else} \end{cases}$$

If ε_t is uncorrelated with u_t, then:

$$r_{\varepsilon_t,u_t}(k) = E\{\varepsilon_{t+k}u_t\} = 0$$

Residual analysis

We have to estimate r

$$\hat{r}_\varepsilon(k) = \frac{1}{N-k}\sum_{t=1}^{N-k}\varepsilon_{t+k}\varepsilon_t \qquad \hat{r}_{\varepsilon,u}(k) = \frac{1}{N-k}\sum_{t=1}^{N-k}\varepsilon_{t+k}u_t$$

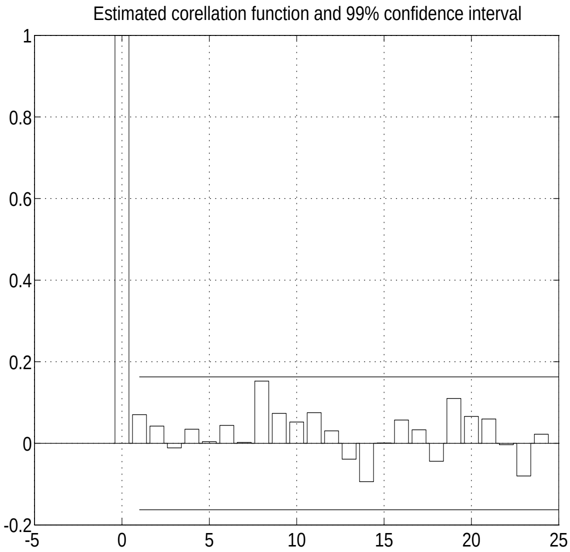
Furthermore:

$$\hat{\rho}_\varepsilon(k) = \frac{\hat{r}_\varepsilon(k)}{\hat{r}_\varepsilon(0)} \qquad \hat{\rho}_{\varepsilon,u}(k) = \frac{\hat{r}_{\varepsilon,u}(k)}{\sqrt{\hat{r}_\varepsilon(0)\hat{r}_u(0)}}$$

$$H_0: \quad \hat{\rho}_\varepsilon(k) \in \mathbf{N}\left(0, \frac{1}{N-k}\right) \qquad \hat{\rho}_{\varepsilon,u}(k) \in \mathbf{N}\left(0, \frac{1}{N-k}\right)$$

The marginal test (H_0 is rejected if):

$$\hat{\rho}_\varepsilon(k) < \frac{f^u_{\frac{\alpha}{2}}}{\sqrt{N-k}} \quad \text{or} \quad \hat{\rho}_\varepsilon(k) > \frac{f^u_{1-\frac{\alpha}{2}}}{\sqrt{N-k}}$$


Estimated corellation function and 99% confidence interval
1
0.8
0.6
0.4
0.2
0
-0.2
-5
0
5
10
15
20
25

The Portmanteau test

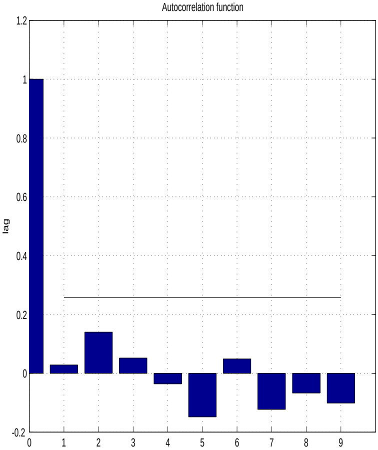
The simultaneous test (if the sequence is white) is based on

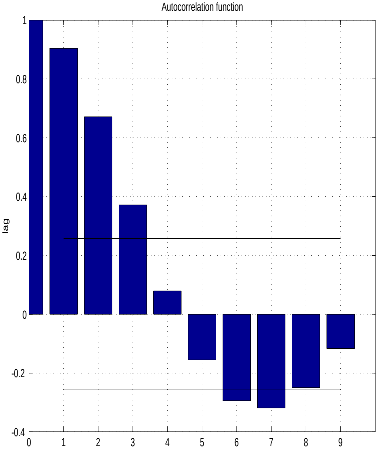
$$H_0 : \quad z = N \sum_{i=1}^{m} \hat{\rho}_\varepsilon(i)^2 \in \chi^2(m-d)$$

and H_0 (sequence is white) is rejected if

$$z = N \sum_{i=1}^{m} \hat{\rho}_\varepsilon(i)^2 > f^{\chi^2}_{1-\alpha}(m-d)$$

Succes and failure:





Spectral analysis

Let us do a FFT on the residuals:

$$X(k) = \sum_{i=0}^{N-1} \varepsilon_i * exp\left[-j\frac{2\pi k}{N} i\right] \qquad k = 0, \ , N-1$$

where $f_k = k\frac{1}{2T_s}$.

Then the estimate of the spectral density is:

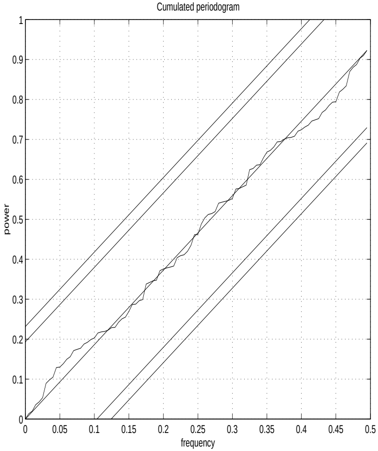
$$\hat{\phi}(\omega_k) = |X(k)|^2$$

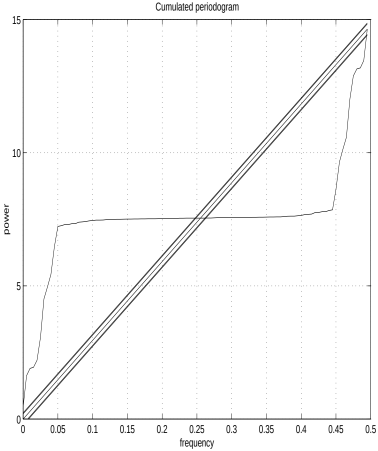
and

$$\sum_{k=0}^{N-1} \hat{\phi}(\omega_k) = \hat{\sigma}^2$$

DTU

Succes and failure:





Partial AutoCorrelation Function

Definition:

$$\phi_k = Cor\left\{y_{t+k}, y_t \middle| y_{t+k-1}, \; \dots, \; y_{t+1}\right\}$$

Estimation:

$$\hat{\phi}_k = -\hat{a}_k \qquad (1 + a_1 q^{-1} + \; \dots \; + a_k q^{-k}) y_t = e_t$$

Property: If y_t originate from an AR process of order n, then $\phi_k = 0$ for $k > n$.

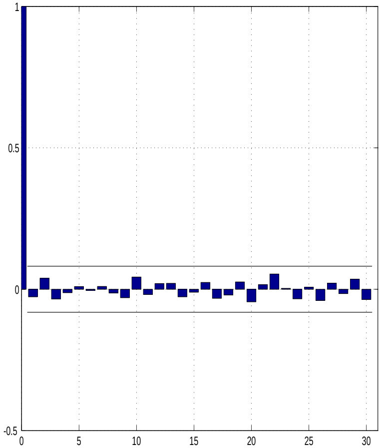
$$\hat{\phi}_k \in \mathbf{N}\left(0, \frac{1}{N}\right)$$

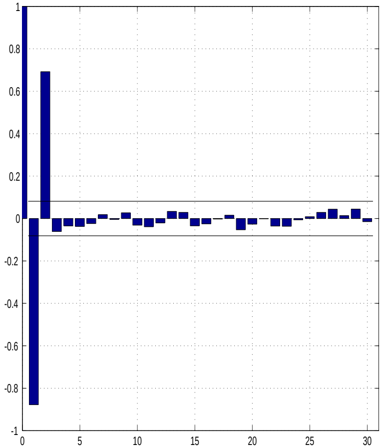
Test: H_0 is rejected if $|\hat{\phi}_k| > \frac{1}{\sqrt{N}} \; f_{1-\frac{1}{2}\alpha}$

Portmanteau:

$$N \sum_{i=1}^{m} \hat{\phi}_i^2 \in \chi^2(m-d)$$

Success and failure (for an AR(2) process):





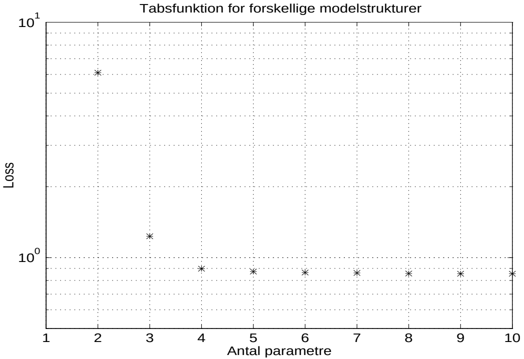
Test between models

Degree of explanation:

$$R^2 = \frac{J_0 - J(\hat{\theta})}{J_0} \qquad J_0 = \frac{1}{2}\sum_{i=1}^{N} y_i^2$$

where:

$$J(\hat{\theta}) = \frac{1}{2}\sum_{i=1}^{N} \varepsilon_i^2 \qquad W(\hat{\theta}) = \sum_{i=1}^{N} \varepsilon_i^2 \qquad W_N(\hat{\theta}) = \frac{1}{N}\sum_{i=1}^{N} \varepsilon_i^2$$



The F-test

which is based on

$$H_0: \quad z = \frac{J_1 - J_2}{J_2} \times \frac{N - d_2}{d_2 - d_1} \in F(d_2 - d_1, N - d_2)$$

ie $d_2 > d_1 \geq g_0$.

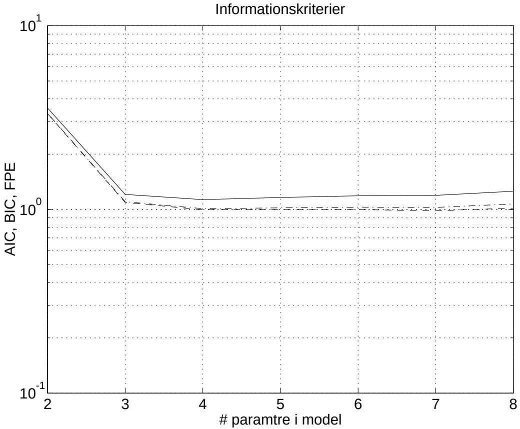
- Forward selction
- Backward selction

Use: [fskema,model,ic]=sfind(nmax,Ze,Zt,mode)

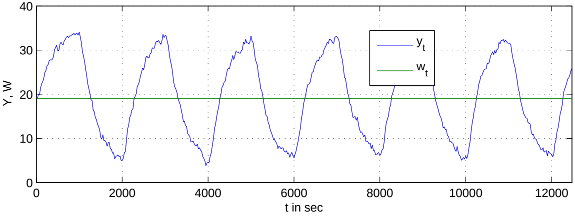
Information criterias

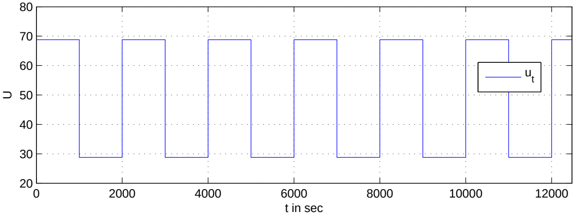
$$AIC = \left(1 + \frac{2d}{N}\right) W_N \qquad BIC = \left(1 + \frac{log(N)d}{N}\right) W_N$$

$$FPE = \frac{N+d}{N-d} W_N \simeq \left(1 + \frac{2d}{N} + \ ...\right) W_N$$

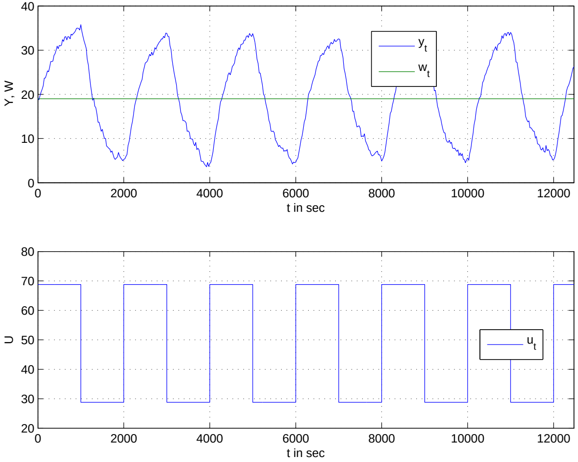


Estimation data





Test data



40
30
20
10
0
Y, W
0
2000
4000
6000
8000
10000
12000
t in sec
y_t
w_t
80
70
60
50
40
30
20
U
0
2000
4000
6000
8000
10000
12000
t in sec
u_t

```
function [nstruc,le,lt]=sfind(nmax,Ze,Zt,mode)
Usage: [nstruc,le,lt]=sfind(nmax,Ze,Zt,mode)

Input:

 nmax: Maximal order
 Ze: Estimation data set
 Zt: Validation (test) data set
 mode: Mode for printing
                (0) Nothing is printed,
                (1) Just the order is printed
                (2) estimation results are printed

Output:

 nstruc:  F-test table
 le:      loss function for estimation data
 lt:      loss function for test data (if exists)
```

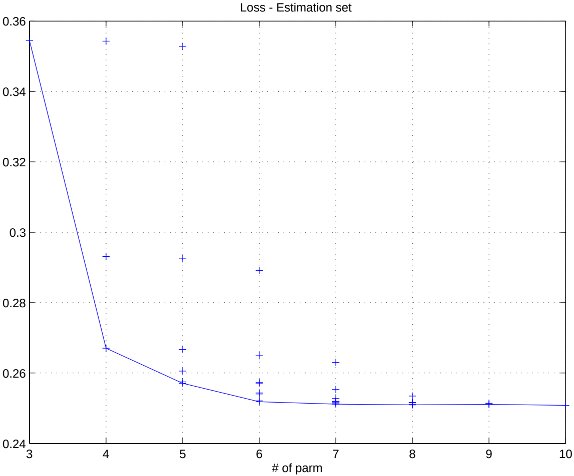
The search

1	1	1	0	0	1
1	1	2	0	0	1
1	1	3	0	0	1
1	2	1	0	0	1
1	2	2	0	0	1
1	2	3	0	0	1
1	3	1	0	0	1
1	3	2	0	0	1
1	3	3	0	0	1
1	4	1	0	0	1
1	4	2	0	0	1
1	4	3	0	0	1
2	1	1	0	0	1
2	1	2	0	0	1
2	1	3	0	0	1
2	2	1	0	0	1
2	2	2	0	0	1
2	2	3	0	0	1
2	3	1	0	0	1
2	3	2	0	0	1
2	3	3	0	0	1
2	4	1	0	0	1
2	4	2	0	0	1
2	4	3	0	0	1
3	1	1	0	0	1
3	1	2	0	0	1
3	1	3	0	0	1
3	2	1	0	0	1
3	2	2	0	0	1
3	2	3	0	0	1
3	3	1	0	0	1
3	3	2	0	0	1
3	3	3	0	0	1
3	4	1	0	0	1
3	4	2	0	0	1
3	4	3	0	0	1

Estimation losses

```
Estimations data

    Loss                 structure

    0.3545     1.0000     1.0000     1.0000          0          0     1.0000
    0.2671     2.0000     1.0000     1.0000          0          0     1.0000
    0.2571     2.0000     2.0000     1.0000          0          0     1.0000
    0.2518     2.0000     2.0000     2.0000          0          0     1.0000
    0.2512     3.0000     3.0000     1.0000          0          0     1.0000
    0.2509     3.0000     4.0000     1.0000          0          0     1.0000
    0.2511     3.0000     3.0000     3.0000          0          0     1.0000
    0.2508     3.0000     4.0000     3.0000          0          0     1.0000
```

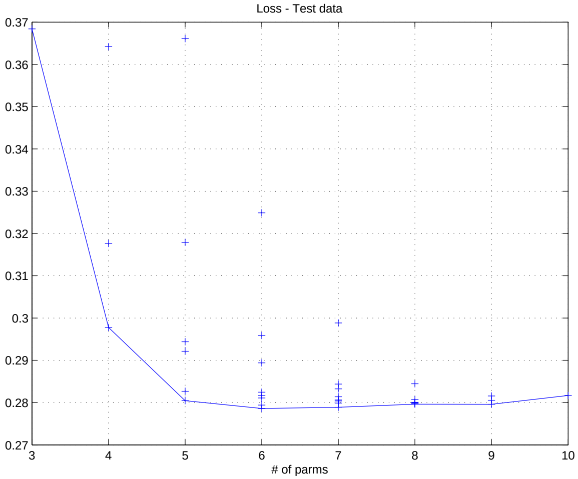
Loss - Estimation



Loss - Test data

Test data

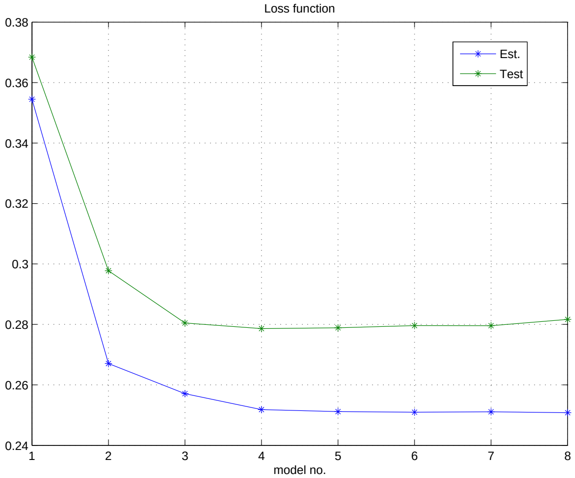
Loss		structure				
0.3684	1.0000	1.0000	1.0000	0	0	1.0000
0.2978	2.0000	1.0000	1.0000	0	0	1.0000
0.2805	2.0000	1.0000	2.0000	0	0	1.0000
0.2786	2.0000	1.0000	3.0000	0	0	1.0000
0.2789	3.0000	1.0000	3.0000	0	0	1.0000
0.2796	3.0000	4.0000	1.0000	0	0	1.0000
0.2796	3.0000	3.0000	3.0000	0	0	1.0000
0.2817	3.0000	4.0000	3.0000	0	0	1.0000


Loss - Test data
0.37
0.36
0.35
0.34
0.33
0.32
0.31
0.3
0.29
0.28
0.27
3
4
5
6
7
8
9
10
# of parms

Losses

Diff		Loss		#parm
NaN	NaN	0.3545	0.3684	3.0000
-0.0874	-0.0706	0.2671	0.2978	4.0000
-0.0100	-0.0173	0.2571	0.2805	5.0000
-0.0053	-0.0018	0.2518	0.2786	6.0000
-0.0007	0.0003	0.2512	0.2789	7.0000
-0.0002	0.0007	0.2509	0.2796	8.0000
0.0001	-0.0000	0.2511	0.2796	9.0000
-0.0003	0.0021	0.2508	0.2817	10.0000

Losses



Loss function
0.38
0.36
0.34
0.32
0.3
0.28
0.26
0.24
1
2
3
4
5
6
7
8
model no.
Est.
Test

The F-test

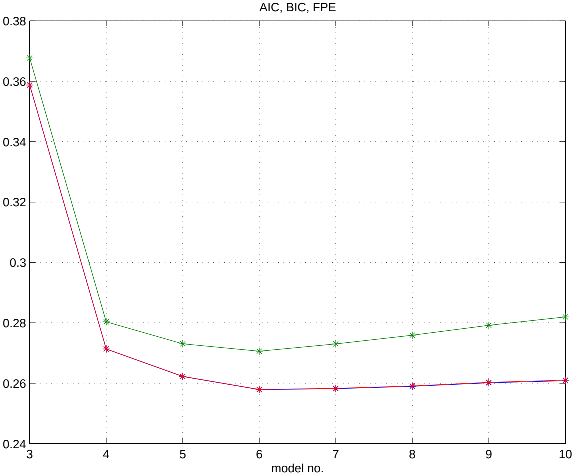
Models

3	1	1	1	0	0	1
4	2	1	1	0	0	1
5	2	2	1	0	0	1
6	2	2	2	0	0	1
7	3	3	1	0	0	1
8	3	4	1	0	0	1
9	3	3	3	0	0	1
10	3	4	3	0	0	1

F test

NaN	4	5	6	7	8	9	10
3	100	100	100	100	100	100	100
4	0	100	100	100	100	100	100
5	0	0	100	100	99	98	97
6	0	0	0	74	58	30	26
7	0	0	0	0	49	7	13
8	0	0	0	0	0	0	13
9	0	0	0	0	0	0	55

Information Criterias



```
AIC BIC FPE

     6     6     6
```

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

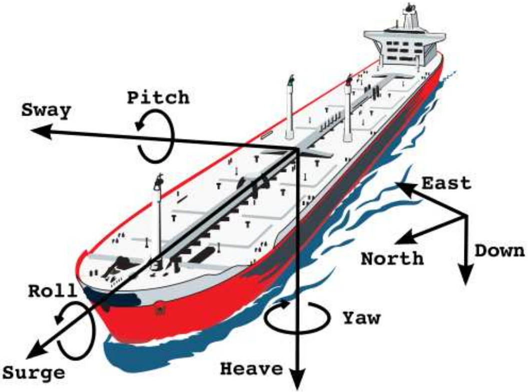
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

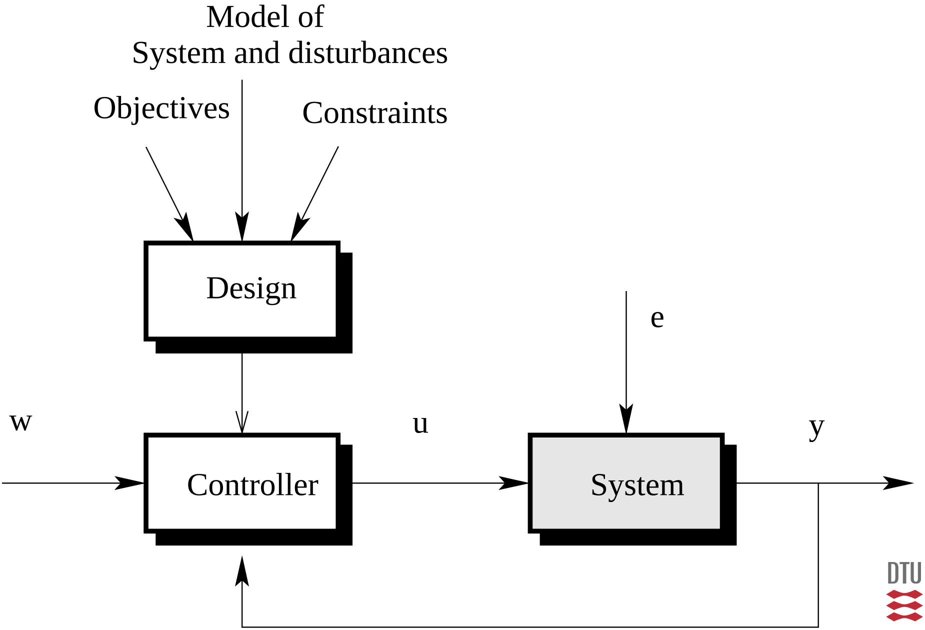
Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

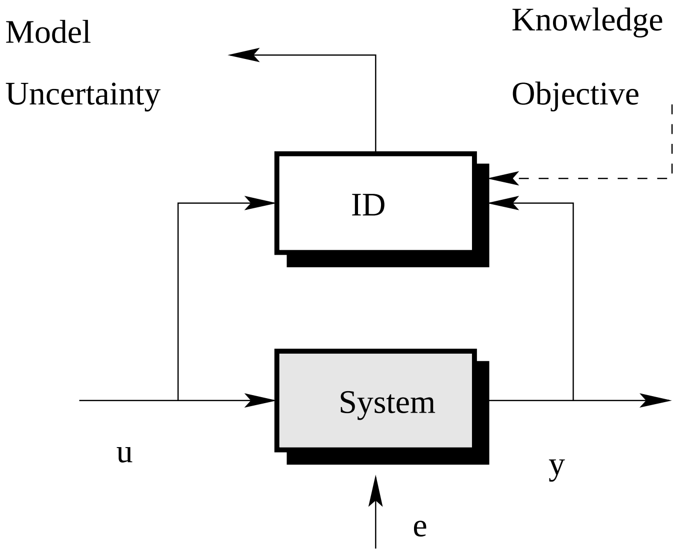
Recursive Estimation (L21)

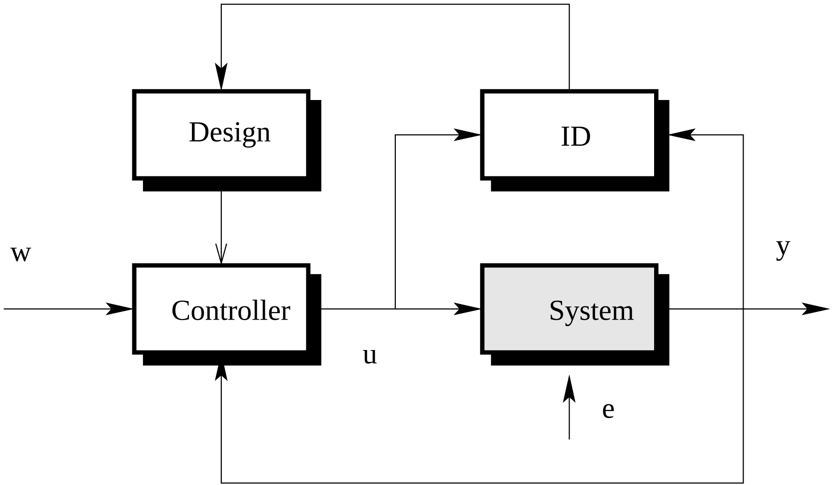
Surface Vessel



From: C. Holden, Roberto Galeazzi, C. Rodrguez, T. Perez, T. I. Fossen, M. Blanke, M. A. S. Neves. Nonlinear Container Ship Model for the Study of Parametric Roll Resonance Modeling, Identification and Control, 28, pp. 87-113, 2007.


Model of
System and disturbances
Objectives
Constraints
Design
e
w
u
y
Controller
System


Model
Uncertainty
Knowledge
Objective
ID
System
u
y
e


Design
ID
w
y
Controller
System
u
e

Recursive Estimation

Estimation

$$\hat{\theta}_t = funk\Big[Y_t\Big]$$

Recursive Estimation

$$\hat{\theta}_t = funk\Big[\hat{\theta}_{t-1}, y_t, X_t\Big]$$

$$X_{t+1} = f\Big(X_t, y_t\Big)$$

$$dim(X_t) < dim(Y_t)$$

advantage - disadvantage

Scalar Xmodels

$$y_t = G(q^{-1})u_t + v_t$$

$$v_t = H(q^{-1})e_t$$

$$e_t : f(e)$$

Parameterization:

$\theta \in G,\ H$ and f

Structures

ARX:

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t$$

ARMAX:

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$

Box-Jenkins:

$$y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

L-structure:

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

The Normal situation:

$$e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

The ARX structure

ARX

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t$$

$$y_t \quad + \quad a_1 y_{t-1} + \ ... \ a_n y_{t-n} = b_0 u_t + \quad b_n u_{t-n} + e_t$$

$$y_t = C_t^\top \theta + e_t$$

$$C_t = (-y_{t-1}, ..., -y_{t-n_a}, u_t, ..., u_{t-n_b})^T$$

$$\theta = (\quad a_1, \quad ... \quad , \quad a_{n_a}, \quad b_0, \quad ... \quad , b_{n_b})^T$$

$$J_t = \sum_{i=1}^{t} \frac{1}{2} \varepsilon_i^2 \qquad \varepsilon_i = y_i - C_i^T \hat{\theta}_t$$

Let $\hat{\theta}_t$ be an estimate based on Y_t:

$$\hat{\theta}_t : \quad \hat{\theta}_{t,n+1} = \hat{\theta}_{t,n} + \left[\sum_{i=1}^{t} C_i C_i^T\right]^{-1} \times \sum_{i=1}^{t} C_i \varepsilon_i$$

where:

$$\varepsilon_i = y_i - C_i^T \hat{\theta}_{t,n} \qquad G_t = \sum_{i=1}^{t} C_i \varepsilon_i \qquad H_t = \sum_{i=1}^{t} C_i C_i^T$$

Now, since $\hat{\theta}_t$ is a LS estimate:

$$\hat{\theta}_t : \sum_{i=1}^{t} C_i \varepsilon_i = 0$$

RARX or RLS

$$\hat{\theta}_t : \quad \hat{\theta}_{t,n+1} = \hat{\theta}_{t,n} + \left[\sum_{i=1}^{t} C_i C_i^T\right]^{-1} \times \sum_{i=1}^{t} C_i \varepsilon_i \qquad n = 1,\ 2,\ \ldots n_{stop}$$

$$\hat{\theta}_{t+1} : \quad \hat{\theta}_{t+1,n+1} = \hat{\theta}_{t+1,n} + \left[\sum_{i=1}^{t+1} C_i C_i^T\right]^{-1} \times \sum_{i=1}^{t+1} C_i \varepsilon_i$$

---

if $\hat{\theta}_{t+1,1} = \hat{\theta}_t$ then

$$G_{t+1} = \sum_{i=1}^{t+1} C_i \varepsilon_i = \sum_{i=1}^{t} C_i \varepsilon_i + C_{t+1}\varepsilon_{t+1} = C_{t+1}\varepsilon_{t+1}$$

An one-step procedure ($n_{stop} = 1, \hat{\theta}_{t+1} = \hat{\theta}_{t+1,1}$) gives:

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \left\{\sum_{i=1}^{t+1} C_i C_i^T\right\}^{-1} C_{t+1}\varepsilon_{t+1}$$

DTU

RARX or RLS

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \left\{ \sum_{i=1}^{t+1} C_i C_i^T \right\}^{-1} C_{t+1}\varepsilon_{t+1}$$

(just a copy)

$$H_{t+1} = \sum_{i=1}^{t+1} C_i C_i^T = H_t + C_{t+1}C_{t+1}^T$$

since $H_t = \sum_{i=1}^{t} C_i C_i^T$

$$H_{t+1} = P_{t+1}^{-1}$$

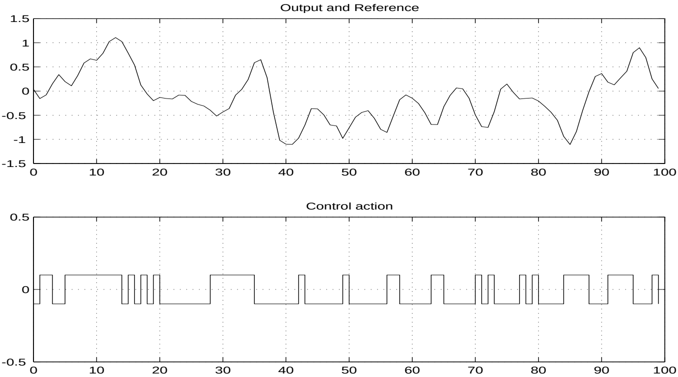
$$\begin{aligned} \varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\ P_{t+1}^{-1} &= P_t^{-1} + C_{t+1}C_{t+1}^T \\ \hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1}C_{t+1}\varepsilon_{t+1} \end{aligned}$$

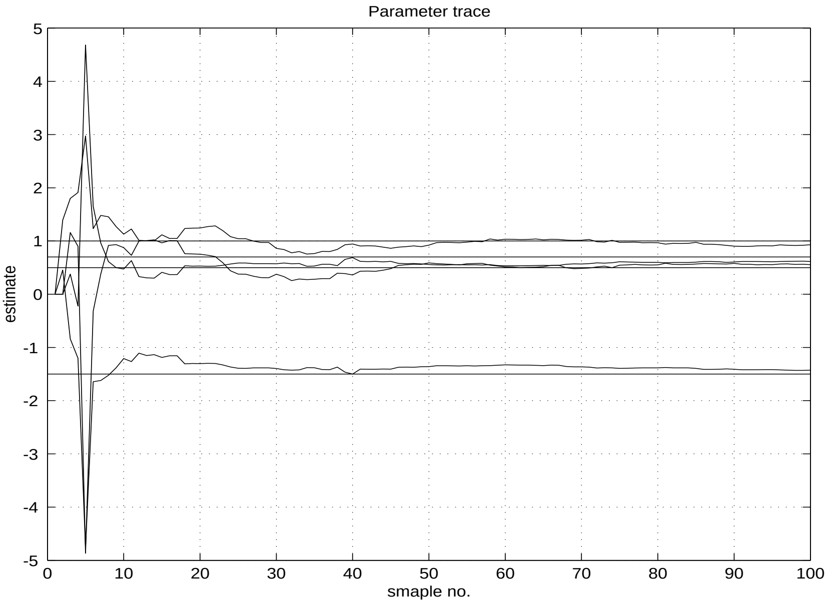
Example

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t \qquad e_t \in \mathbf{N}_{iid}(0, 0.1)$$

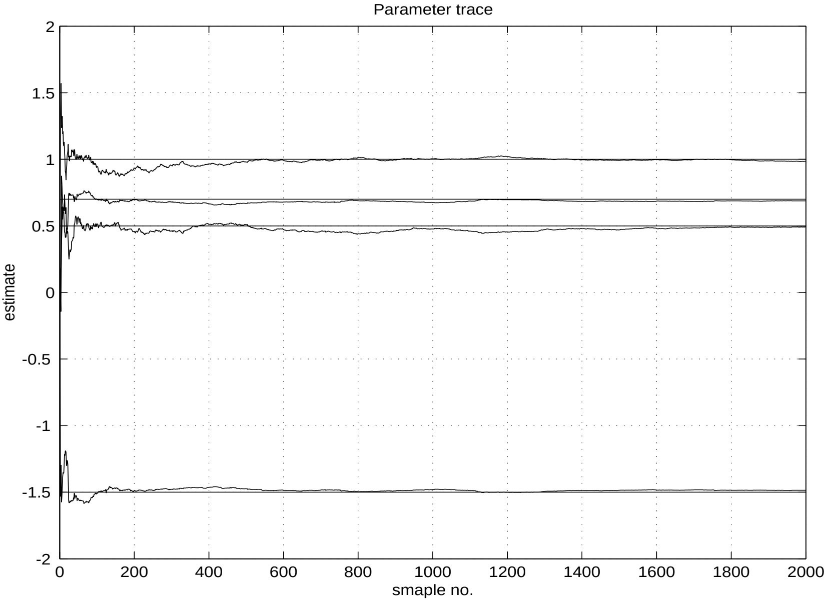
where

$$A(q^{-1}) = 1 - 1.5q^{-1} + 0.7q^{-2} \qquad\qquad B(q^{-1}) = q^{-1} + 0.5q^{-2}$$



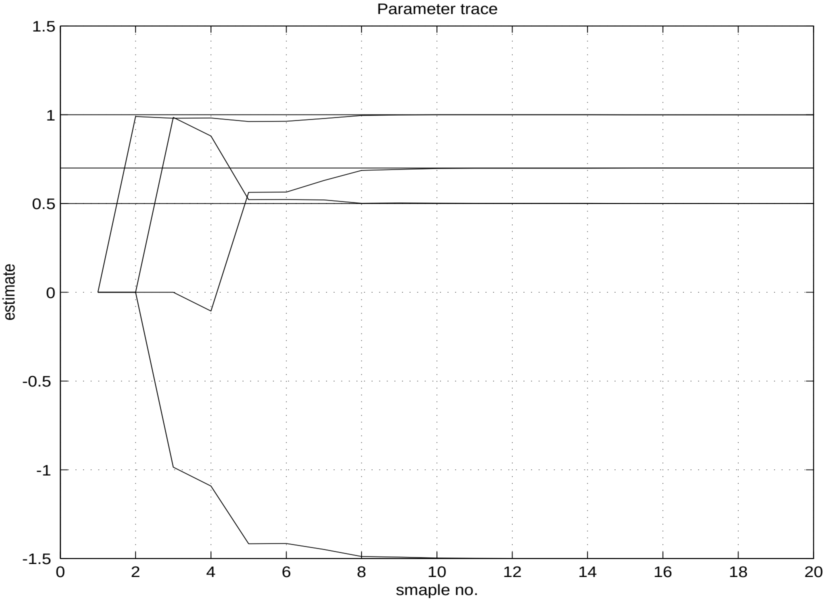

Parameter trace
5
4
3
2
1
0
-1
-2
-3
-4
-5
estimate
0
10
20
30
40
50
60
70
80
90
100
smaple no.

A longer run.



Parameter trace
estimate
smaple no.
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
200
400
600
800
1000
1200
1400
1600
1800
2000

and a short deterministic one



Hemes Invertion Lemma

$$P_{t+1}^{-1} = P_t^{-1} + C_{t+1}C_{t+1}^T \qquad P_{t+1} = \left[P_t^{-1} + C_{t+1}C_{t+1}^T\right]^{-1}$$

$$\left(A_{11}^{-1} + A_{12}A_{22}A_{21}\right)^{-1} = A_{11} - A_{11}A_{12}\left(A_{21}A_{11}A_{12} + A_{22}^{-1}\right)^{-1} A_{21}A_{11}$$

$$A_{11} = P_i \qquad A_{12} = C$$

$$A_{21} = C^T \qquad A_{22} = 1$$

$$P_{t+1} = P_t - \frac{P_t C_{t+1} C_{t+1}^T P_t}{1 + C_{t+1}^T P_t C_{t+1}}$$

RLS algorithm

$$\hat{\theta}_{t+1} = \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1}$$

$$P_{t+1} = P_t - \frac{P_t C_{t+1} C_{t+1}^T P_t}{1 + C_{t+1}^T P_t C_{t+1}}$$

$$\begin{aligned} K_{t+1} &= P_{t+1} C_{t+1} = P_t C_{t+1} - \frac{P_t C_{t+1} C_{t+1}^T P_t C_{t+1}}{1 + C_{t+1}^T P_t C_{t+1}} \\ &= \frac{P_t C_{t+1}}{1 + C_{t+1}^T P_t C_{t+1}} = \frac{P_t C_{t+1}}{s_{t+1}} \end{aligned}$$

$$P_{t+1} = P_t - \frac{P_t C_{t+1} C_{t+1}^T P_t}{1 + C_{t+1}^T P_t C_{t+1}} = P_t - \frac{P_t C_{t+1} C_{t+1}^T P_t}{1 + C_{t+1}^T P_t C_{t+1}}$$

RLS algorithm

$$P_{t+1} = P_t - \frac{P_t C_{t+1} C_{t+1}^T P_t}{1 + C_{t+1}^T P_t C_{t+1}} = P_t - \frac{P_t C_{t+1} C_{t+1}^T P_t}{1 + C_{t+1}^T P_t C_{t+1}}$$

LS/RARX-algorithm

$$\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
s_{t+1} &= 1 + C_{t+1}^T P_t C_{t+1} \\
K_{t+1} &= \frac{P_t C_{t+1}}{s_{t+1}} \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + K_{t+1} \varepsilon_{t+1} \\
P_{t+1} &= P_t - K_{t+1} s_{t+1} K_{t+1}^T
\end{aligned}$$

UDU, Square root algorithms

A skeleton:

```
[A,B,k,C]=sysinit(sflag); % Determine linear model
%----------------------------------------------------------------------
wt=prbs(nstp,15);
%----------------------------------------------------------------------
th=[A(2:end) B C(2:end)]';  th=diag((1+0.1*randn(size(th))))*th; th=th*0;
pil=[length(A)-1 length(B)];
pil=[0 cumsum(pil)]+1;
fi=zeros(size(th));
P=eye(size(th,1))*1e6;
%----------------------------------------------------------------------
measinit(sflag); % Initilialise the measurement system
for it=1:length(wt),
 y=meas;

 res=y-fi'*th;            % Prediction error
 s=1+fi'*P*fi;            % Auxillary variable
 K=P*fi/s;                % Gain
 P=(P-K*s*K');            % Covariance update
 th=th+K*res;             % Estimate update

 u=wt(it);                % Excitation

 fi(2:end)=fi(1:end-1);   % fi administration
 fi(pil(1:2))=[-y u];

 data=[data; t y u res th' trace(P)];
 act(u);  % Actuate control
end
```

The regressors

$$C_t^T = (-y_{t-1}, -y_{t-2}, -y_{t-3}, u_{t-1}, u_{t-2}, u_{t-3})$$

$$C_{t+1}^T = (-y_t, -y_{t-1}, -y_{t-2}, u_t, u_{t-1}, u_{t-2})$$

RLS algorithm

$$
\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
P_{t+1}^{-1} &= P_t^{-1} + C_{t+1} C_{t+1}^T \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1}
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
s_{t+1} &= 1 + C_{t+1}^T P_t C_{t+1} \\
K_{t+1} &= \frac{P_t C_{t+1}}{s_{t+1}} \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + K_{t+1} \varepsilon_{t+1} \\
P_{t+1} &= P_t - K_{t+1} s_{t+1} K_{t+1}^T
\end{aligned}
$$

friends

RLS algorithm

Alternative I:

$$\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^{\top}\hat{\theta}_t \\
P_{t+1}^{-1} &= P_t^{-1} + C_{t+1}C_{t+1}^T \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1}C_{t+1}\varepsilon_{t+1}
\end{aligned}$$

Alternative II:

$$\begin{aligned}
\hat{\theta}_{t+1} &= \hat{\theta}_t + \frac{P_t C_{t+1}}{1 + C_{t+1}^{\top} P_t C_{t+1}} \varepsilon_{t+1} \\
P_{t+1} &= P_t - \frac{P_t C_{t+1} C_{t+1}^{\top} P_t}{1 + C_{t+1}^{\top} P_t C_{t+1}}
\end{aligned}$$

or combination hereof.

$$K_{t+1} = P_{t+1}C_{t+1} = \frac{P_t C_{t+1}}{1 + C_{t+1}^T P_t C_{t+1}}$$

$$\begin{aligned}
P_{t+1} &= P_t - K_{t+1}\left[1 + C_{t+1}^T P_t C_{t+1}\right]K_{t+1}^T \\
&= \left[I - K_{t+1}C_{t+1}^T\right]P_t
\end{aligned}$$

In general:

$$\hat{\theta}_{t+1} = \hat{\theta}_t + K_{t+1}\varepsilon_{t+1}$$

$$\varepsilon_{t+1} = y_{t+1} - C_{t+1}^T \hat{\theta}_t$$

Introducing the error:

$$\tilde{\theta}_t = \theta_0 - \hat{\theta}_t$$

then:

$$\tilde{\theta}_{t+1} = \left[I - K_{t+1}C_{t+1}^T\right]\tilde{\theta}_t - K_{t+1}e_{t+1}$$

If (LS case)

$$K_{t+1} = \frac{P_t C_{t+1}}{1 + C_{t+1}^T P_t C_{t+1}}$$

$n-1$ eigenvalues in 1 and one eigenvalue in

$$\frac{1}{1 + C_{t+1}^T P_t C_{t+1}} \leq 1$$

DTU

Alternative III:

$$\begin{aligned}\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1} \\ P_{t+1}^{-1} &= P_t^{-1} + C_{t+1} C_{t+1}^T\end{aligned}$$

$$R_t = \frac{1}{t} P_t^{-1} = \frac{1}{t} \sum_{i=1}^{t} C_i C_i^T \qquad P_t = \frac{1}{t} R_t^{-1}$$

$$\begin{aligned}\hat{\theta}_{t+1} &= \hat{\theta}_t + \frac{1}{t+1} R_{t+1}^{-1} C_{t+1} \varepsilon_{t+1} \\ R_{t+1} &= R_t + \frac{1}{t+1} \left[C_{t+1} C_{t+1}^T - R_t \right]\end{aligned}$$

Analysis of Convergense:

$$\hat{\theta}_t \rightarrow \theta_0 + R^{-1} \mathbf{E}\left\{ C_t \varepsilon_t \right\}$$

Interpretation

Recursive minimization (PEM)

$$\hat{\theta}_t = arg\textit{Min} \sum_{i=0}^{t} \varepsilon_i^2$$

Bayes interpretation

$$\theta_t | \, Y_t \in \mathbf{N}\left(\hat{\theta}_t, \; P_t \sigma^2\right)$$

PLR

$$\hat{\theta}_t = \textit{Sol}\left(\sum_{i=0}^{t} C_i \varepsilon_i = 0\right)$$

Bayes method

Consider the system:

$$\theta_{t+1} = \theta_t$$

$$y_t = C_t^T \theta_t + e_t$$

Kalman filter

Time update:

$$\hat{\theta}_{t+1|t} = \hat{\theta}_{t|t}$$

$$P_{t+1|t} = P_{t|t}$$

Data update:

$$\begin{aligned}
\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1} \\
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
P_{t+1}^{-1} &= P_t^{-1} + C_{t+1} C_{t+1}^T
\end{aligned}$$

PLR

LS estimator:

$$\hat{\theta}_t = \left[\sum_{i=1}^{t} C_i C_i^T\right]^{-1} \left[\sum_{i=1}^{t} C_i y_i\right]$$

Let $\bar{\theta}_t$ be an arbitrary estimator

$$\varepsilon_i = y_i - C_i^T \bar{\theta}_t$$

$$\hat{\theta}_{t+1} = \bar{\theta}_t + \left[\sum_{i=1}^{t+1} C_i C_i^T\right]^{-1} \left[\sum_{i=1}^{t+1} C_i \varepsilon_i\right]$$

Linear regression:

$$\hat{\theta}_t : \sum_{i=1}^{t} C_i \varepsilon_i = 0$$

RLS and friends

RARX or RLS

$$\begin{aligned}\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1}C_{t+1}\varepsilon_{t+1}\\ P_{t+1}^{-1} &= P_t^{-1} + C_{t+1}C_{t+1}^\top\end{aligned}$$

STA-algorithm

$$\begin{aligned}\hat{\theta}_{t+1} &= \hat{\theta}_t + \frac{1}{T_{t+1}}C_{t+1}\varepsilon_{t+1}\\ T_{t+1} &= T_t + 1\end{aligned}$$

Gradient algorithm

$$\begin{aligned}\hat{\theta}_{t+1} &= \hat{\theta}_t + \frac{1}{T_{t+1}}C_{t+1}\varepsilon_{t+1}\\ T_{t+1} &= T_t + C_{t+1}^\top C_{t+1}\end{aligned}$$

RARX - RLS

$$\hat{\theta}_{t+1} = \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1}$$
$$P_{t+1}^{-1} = P_t^{-1} + C_{t+1} C_{t+1}^\top$$

Projection method

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{1}{T_{t+1}} C_{t+1} \varepsilon_{t+1}$$
$$T_{t+1} = C_{t+1}^\top C_{t+1}$$

Projection method: Given $\hat{\theta}_{t-1}$ find that $\hat{\theta}_t$ that minimize

$$J = \frac{1}{2} \|\hat{\theta}_t - \hat{\theta}_{t-1}\|^2$$

subject to

$$y_t = C^\top \theta_t$$

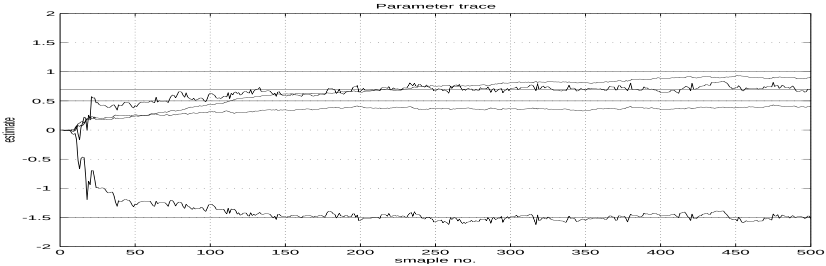
Projection algorithm

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{aC_{t+1}}{C_{t+1}^\top C_{t+1}} \varepsilon_{t+1}$$

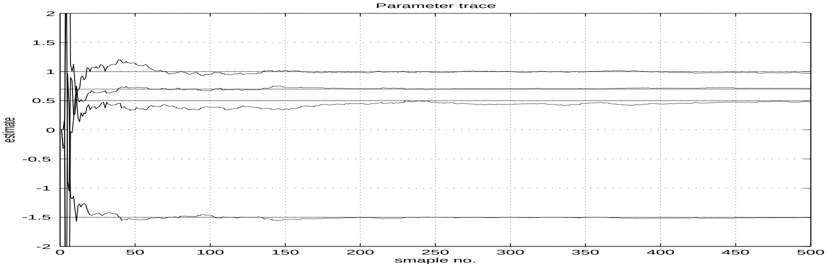
Normalized projection algorithm

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{aC_{t+1}}{c + C_{t+1}^\top C_{t+1}} \varepsilon_{t+1}$$

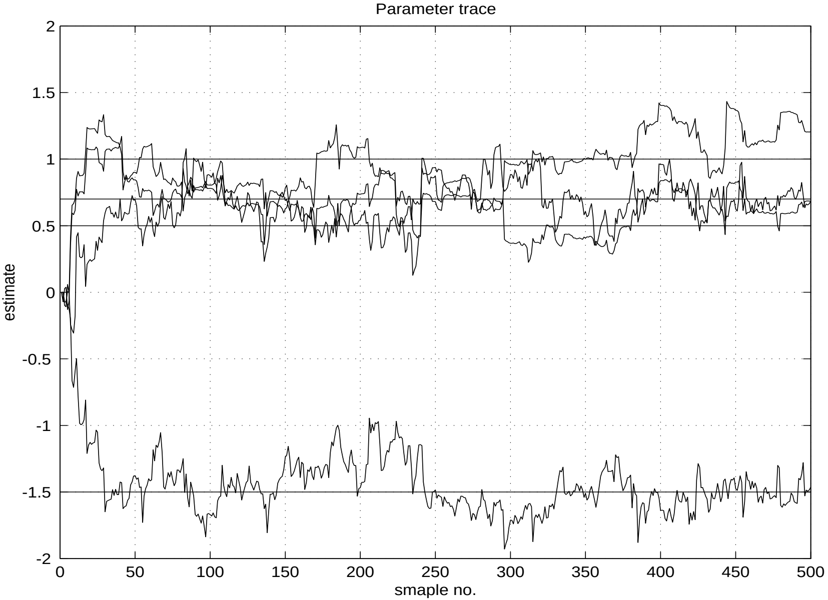
Unnormalized projection algorithm ($a = 0.5$).



RARX (RLS):



Normalized projection algorithm ($a = 0.5$):



ARMAX structure

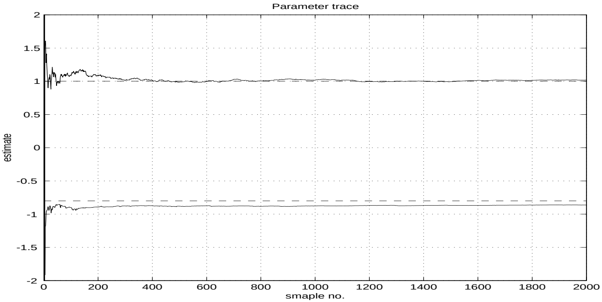
Example

System:

$$(1 - 0.8q^{-1})y_t = 1\ u_{t-1} + (1 + 0.7q^{-1})e_t$$

Model:

$$(1 + aq^{-1})y_t = bu_{t-1} + e_t$$



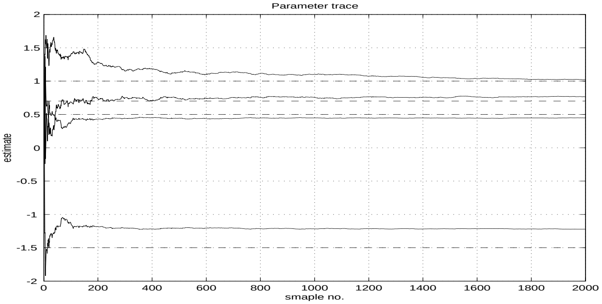
Example

System:

$$(1 - 1.5q^{-1}0.7q^{-2})y_t = (q^{-1} + 0.5q^{-2})u_t$$
$$+(1 - q^{-1} + 0.2q^{-2})e_t$$

Model:

$$(1 + a_1q^{-1} + a_2q^{-2})y_t = (b_1q^{-1} + b_2q^{-2})u_t + e_t$$



ARMAX structure

Consider a system given in the ARMAX structure:

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

Introducing the vectors:

$$C_t = (-y_{t-1}, \cdots, -y_{t-n_a}, \\ u_t,\ u_{t-1},\ \cdots,\ u_{t-n_b}, \\ e_{t-1},\ \cdots,\ e_{t-n_c}\ )^T$$

$$\theta = (a_1, \cdots, a_{n_a}, b_0, b_1, \cdots, b_{n_b}, c_1, \cdots, c_{n_c})^T$$

we can put the description into the regression form:

$$y_t = C_t^T \theta + e_t$$

where the prediction is $\hat{y}_{t|t-1} = C_t^T \theta$.

ARMAX structure

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t$$

$$y_t = C_t^\top \theta + e_t$$

$$C_t = \left(-y_{t-1}, \cdots, -y_{t-n_a}, \; u_t, \; u_{t-1}, \; \cdots, \; u_{t-n_b}, \; e_{t-1}, \; \cdots, \; e_{t-n_c} \right)^T$$

$$\theta = (\; a_1, \quad \cdots, \quad a_{n_a}, \quad b_0, \quad b_1, \quad \cdots, \quad b_{n_b}, \quad c_1, \quad \cdots, \quad c_{n_c})^T$$

$$\varepsilon_t = y_t - C_t^T \hat{\theta} = \frac{\hat{A}(q^{-1})y_t - \hat{B}(q^{-1})u_t}{\hat{C}(q^{-1})}$$

RELS algorithm

The innovation e_t is unknown but ε_t is a good estimate (if $\hat{\theta}_t$ is a good estimate).

$$\begin{aligned} \hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1} \\ P_{t+1}^{-1} &= P_t^{-1} + C_{t+1} C_{t+1}^\top \end{aligned}$$

Equivalent to RLS except for C_t which contains estimate of e_t and is extended in relation to RLS (or RARX).

RARMAX algorithm

Also denoted as the RML algorithm. If we apply the same arguments as for the RARX algorthm for minimizing:

$$J = \sum_{i=0}^{t} \frac{1}{2} \varepsilon_i^2$$

we obtain:

$$\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
P_{t+1}^{-1} &= P_t^{-1} + \psi_{t+1} \psi_{t+1}^\top \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} \psi_{t+1} \varepsilon_{t+1}
\end{aligned}$$

where

$$\psi_t = \frac{\partial}{\partial \theta} \hat{y}_{t|t-1} = -\frac{\partial}{\partial \theta} \varepsilon_t = \frac{1}{C} C_t$$

RARMAX algorithm

$$\begin{aligned} \varepsilon_t &= y_t - C_t^\top \hat{\theta} \\ &= \frac{A y_t - B u_t}{C} \end{aligned}$$

$$C\varepsilon_t = A y_t - B u_t$$

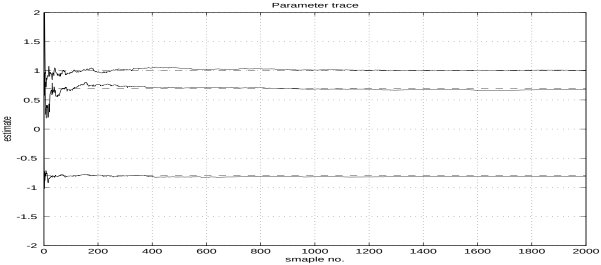
$$C \frac{\partial}{\partial a_i} \varepsilon_t = y_{t-i}$$

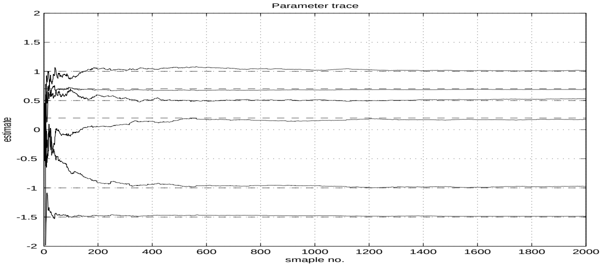
$$C \frac{\partial}{\partial b_i} \varepsilon_t = -u_{t-i}$$

$$\varepsilon_{t-i} + C \frac{\partial}{\partial c_i} \varepsilon_t = 0$$

$$\psi_t = \frac{1}{C} \begin{bmatrix} -y_{t-i} \\ u_{t-i} \\ \varepsilon_{t-i} \end{bmatrix} = \frac{1}{C} C_t$$

NB: stability check.


Parameter trace
estimate
smaple no.


Parameter trace
estimate
smaple no.

$$\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
P_{t+1}^{-1} &= P_t^{-1} + \psi_{t+1}\psi_{t+1}^\top \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1}\psi_{t+1}\varepsilon_{t+1}
\end{aligned}$$

$$\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
s_{t+1} &= 1 + \psi_{t+1}^T P_t \psi_{t+1} \\
K_{t+1} &= \frac{P_t \psi_{t+1}}{s_{t+1}} \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + K_{t+1}\varepsilon_{t+1} \\
P_{t+1} &= P_t - K_{t+1} s_{t+1} K_{t+1}^T
\end{aligned}$$

$$\psi_t = \frac{1}{\hat{C}(q^{-1})} C_t$$ *RARMAX*

$$\psi_t = C_t$$ *RELS*

L-structure

L-structure

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

$$y_t = C_t^\top \theta + e_t$$

$$C_t = (-y_{t-1},\ \dots, u_t,\ \dots,\ -w_{t-1},\ \dots, e_{t-1},\ \dots, -\eta_{t-1},\ \dots)^T$$

$$\theta = (a_1, \dots, b_0, \dots, f_1, \dots, c_1, \dots, d_1,\ \dots)^T$$

$$\varepsilon_t = y_t - C_t^T\hat{\theta} = \frac{\hat{D}(q^{-1})}{\hat{C}(q^{-1})}\left[\hat{A}(q^{-1})y_t - \frac{\hat{B}(q^{-1})}{\hat{F}(q^{-1})}u_t\right]$$

L-structure

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

$$y_t = C_t^\top \theta + e_t$$

pem $J_t = \frac{1}{2}\sum_{i=0}^{t} \varepsilon_i^2$ *plr* $\sum_{i=0}^{t} C_i \varepsilon_i = 0$

$$\begin{aligned} \hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1}\psi_{t+1}\varepsilon_{t+1} \\ P_{t+1}^{-1} &= P_t^{-1} + \psi_{t+1}\psi_{t+1}^\top \end{aligned}$$

$\psi_t = \mathit{Filt}(C_t)$ *pem*

$\psi_t = C_t$ *plr*

RPEM - RPLR

$$
\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
P_{t+1}^{-1} &= P_t^{-1} + \psi_{t+1} \psi_{t+1}^\top \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} \psi_{t+1} \varepsilon_{t+1}
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon_{t+1} &= y_{t+1} - C_{t+1}^T \hat{\theta}_t \\
s_{t+1} &= 1 + \psi_{t+1}^T P_t \psi_{t+1} \\
K_{t+1} &= \frac{P_t \psi_{t+1}}{s_{t+1}} \\
\hat{\theta}_{t+1} &= \hat{\theta}_t + K_{t+1} \varepsilon_{t+1} \\
P_{t+1} &= P_t - K_{t+1} s_{t+1} K_{t+1}^T
\end{aligned}
$$

$\psi_t = \frac{1}{\hat{C}(q^{-1})} C_t$ *RPEM* $\psi_t = C_t$ *RPLR*

- rarx
- rarmax
- roe
- rbj
- rpem
- rplr

Time variation

Time variation

- Problem with RLS (Examples)
- Forgetting methods
- Model based estimation

RLS algorithm

$$J_t = \sum_{i=0}^{t} \varepsilon_{t-i}^2$$

$$J_{t+1} = J_t + \varepsilon_{t+1}^2$$

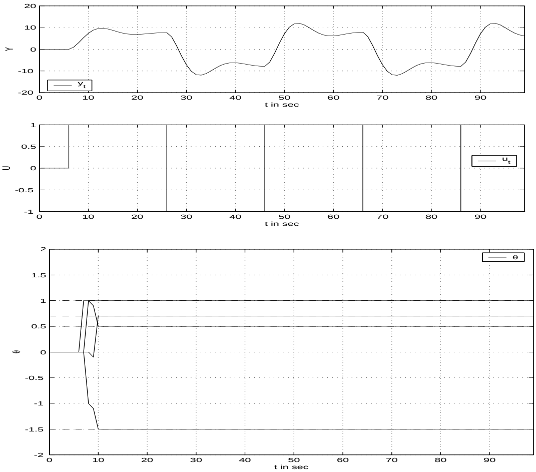
Recursions

$$\begin{aligned}\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1} \\ P_{t+1}^{-1} &= P_t^{-1} + C_{t+1} C_{t+1}^\top\end{aligned}$$

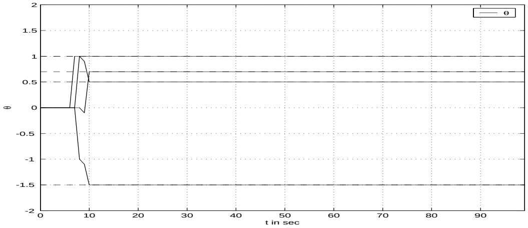
$$\begin{aligned}\hat{\theta}_{t+1} &= \hat{\theta}_t + \frac{P_t C_{t+1}}{1 + C_{t+1}^\top P_t C_{t+1}} \varepsilon_{t+1} \\ P_{t+1} &= \left[P_t - \frac{P_t C_{t+1} C_{t+1}^\top P_t}{1 + C_{t+1}^\top P_t C_{t+1}} \right]\end{aligned}$$

Example

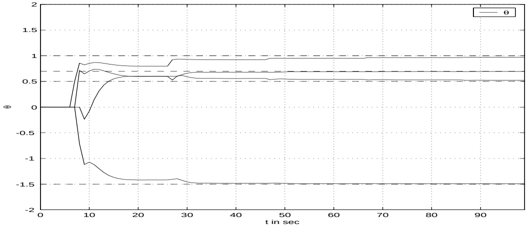
$$P_0 = 10^9 \; I \qquad \sigma^2 = 0$$



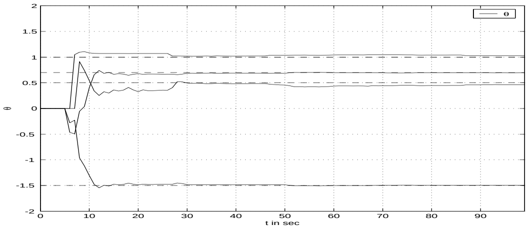
$P_0 = 10^9\ I$


θ
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
10
20
30
40
50
60
70
80
90
t in sec
θ

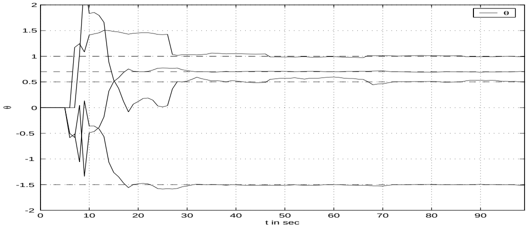
$P_0 = 1\ I$


θ
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
10
20
30
40
50
60
70
80
90
t in sec
θ

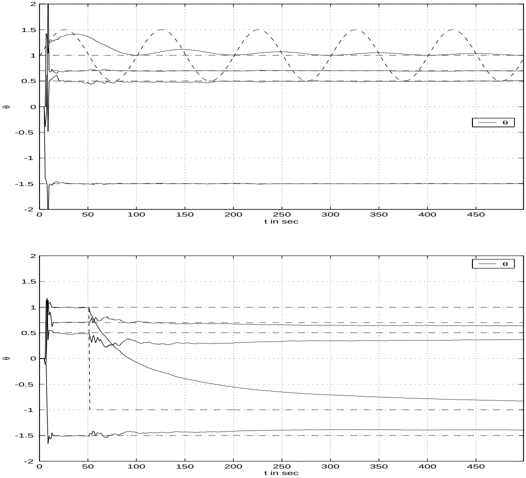
$P_0 = 10^9\ I \qquad \sigma^2 = 0.01$



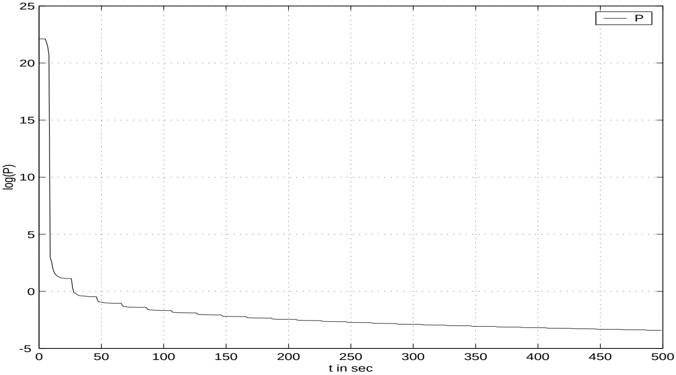
$P_0 = 10^9\ I \qquad \sigma^2 = 0.1$



$P_0 = 10^9\ I \qquad \sigma^2 = 0.01 \qquad w_t$: PRBS



($\sigma^2 = 0.01$, $P_0 = 10^9\ I$, sqwave as excitation)



Forgetting methods

The problem is $K_t \to 0$. How to avoid this.

Model based methods

Model for the variation.

Resetting method

$$\hat{\theta}_{t_i} = \theta_i \qquad P_{t_i} = P_i$$

For example:

$$\theta_i = \hat{\theta}_{t_i} \qquad P_{t_i} = P_1 \qquad t_i = i * N$$

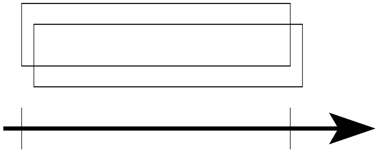
Good in connection to batch processes.

Rectangular forgetting

$$J_t = \sum_{i=0}^{N} \varepsilon_{t-i}^2$$

$$J_t = J_{t-1} + \varepsilon_t^2 - \varepsilon_{t-N-1}^2$$

Large storage demand - Osc. horizon



Exponential forgetting

The most popular method so far.

$$J_t = \sum_{i=0}^{t} \lambda^i \varepsilon_{t-i}^2 \qquad \lambda \leq 1 \qquad \text{For } \lambda = 1 \text{ we have RLS}$$

$$J_{t+1} = \lambda J_t + \varepsilon_{t+1}^2$$

Recursions

$$\begin{aligned} \hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1} \\ P_{t+1}^{-1} &= \lambda\, P_t^{-1} + C_{t+1} C_{t+1}^\top \end{aligned}$$

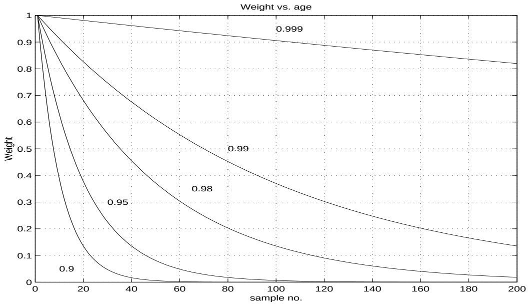
$$\begin{aligned} \hat{\theta}_{t+1} &= \hat{\theta}_t + \frac{P_t C_{t+1}}{\lambda + C_{t+1}^\top P_t C_{t+1}} \varepsilon_{t+1} \\ P_{t+1} &= \left[P_t - \frac{P_t C_{t+1} C_{t+1}^\top P_t}{\lambda + C_{t+1}^\top P_t C_{t+1}} \right] \frac{1}{\lambda} \end{aligned}$$

Equivalent horizon

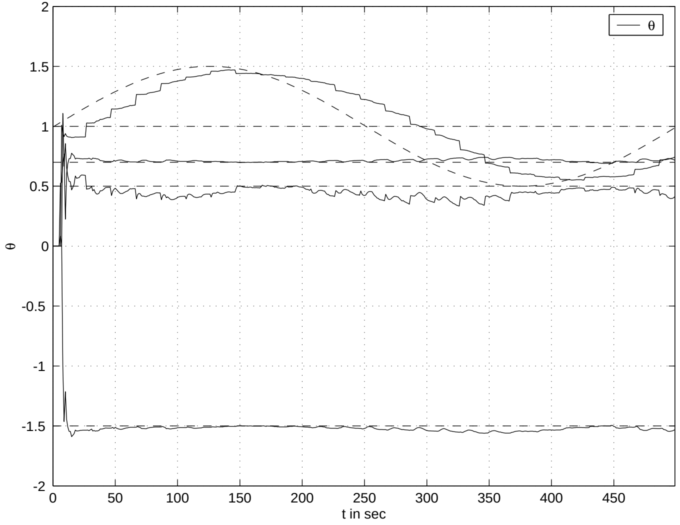
$$N = \sum_{i=0}^{\infty} \lambda^i = \frac{1}{1-\lambda}$$

$$\lambda = 1 - \frac{1}{N}$$

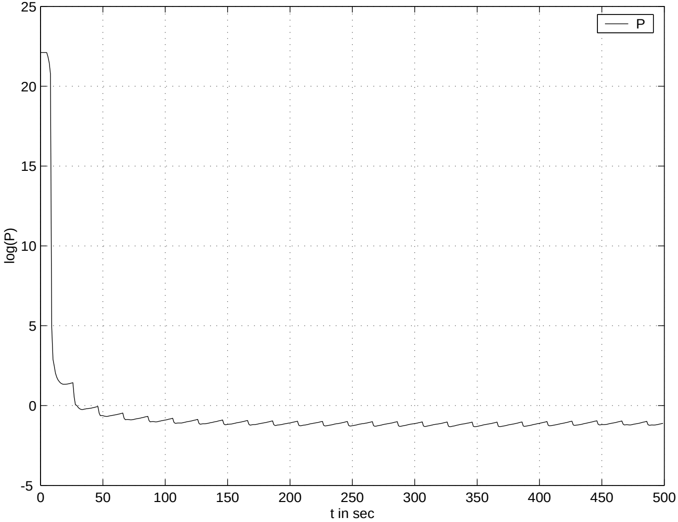
λ is called a forgetting factor, but is actually a memory factor.



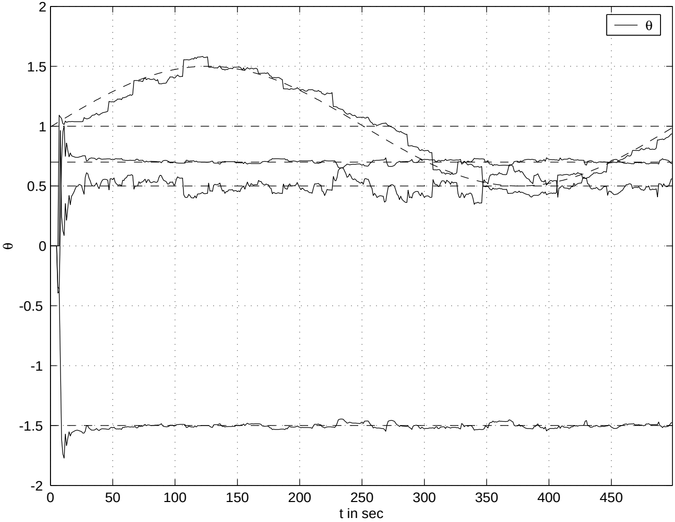
$\lambda = 0.98$



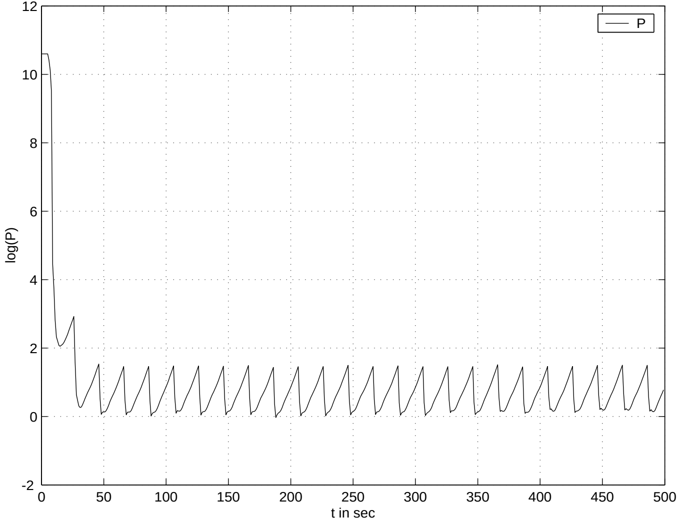
$\lambda = 0.98$


P
log(P)
t in sec
25
20
15
10
5
0
-5
0
50
100
150
200
250
300
350
400
450
500

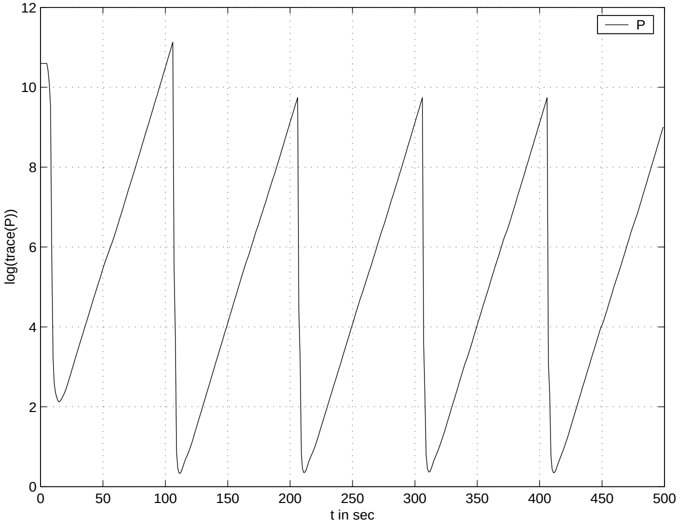
$\lambda = 0.90$



$\lambda = 0.90$



12
10
8
6
4
2
0
-2
log(P)
0
50
100
150
200
250
300
350
400
450
500
t in sec
P


P
log(trace(P))
t in sec
0
2
4
6
8
10
12
0
50
100
150
200
250
300
350
400
450
500

DTU

Fortescue

Original proposed as a method for avoiding covariance blow-up.

$$J_{t+1} = \lambda_{t+1} J_t + \varepsilon_{t+1}^2 \equiv N_\infty \times \sigma^2$$

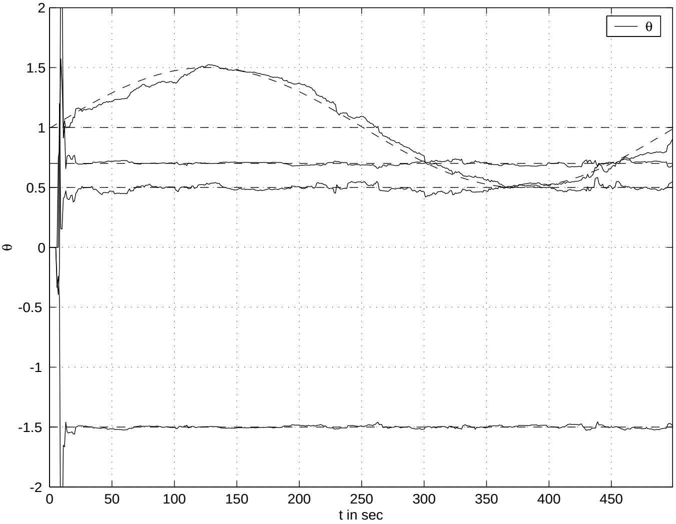
$$\lambda_{t+1} = 1 - \frac{1}{N_\infty} \times \frac{\varepsilon_{t+1}^2}{\sigma^2(1 + C_{t+1}^\top P_t C_{t+1})}$$

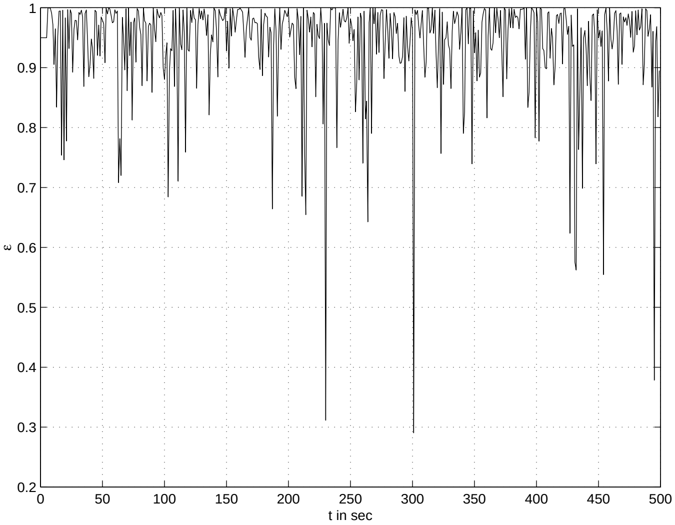
Recursion

$$\begin{aligned} \hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1} \\ P_{t+1}^{-1} &= \lambda_{t+1} \, P_t^{-1} + C_{t+1} C_{t+1}^\top \end{aligned}$$

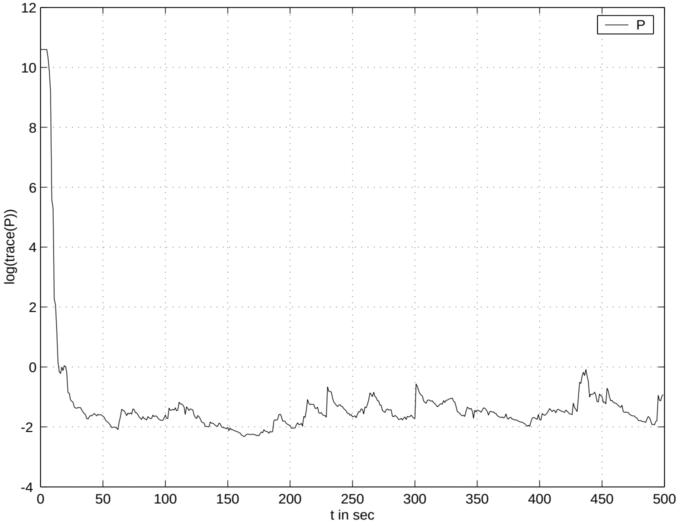
DTU

$N = 20$

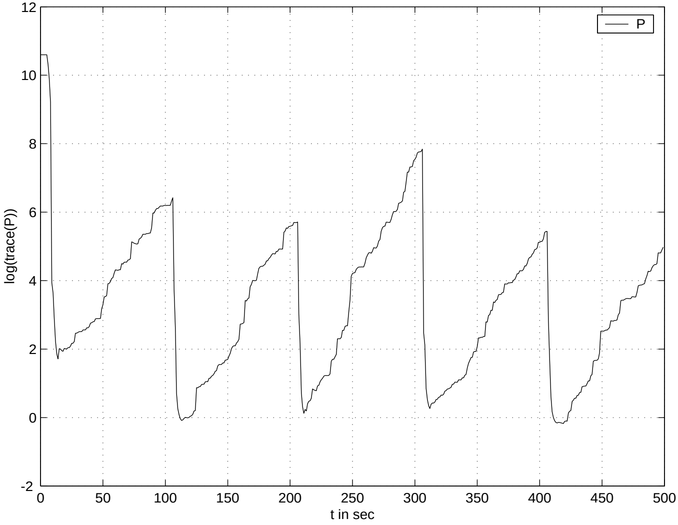



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
ε
0
50
100
150
200
250
300
350
400
450
500
t in sec

PRBS as exitation



sqwave as exitation



Exponential Forgetting with resettling algorihm
Originally proposed by Goodwin.

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{P_t C_{t+1}}{1 + C_{t+1}^\top P_t C_{t+1}} \varepsilon_{t+1}$$

$$P_{t+1} = \left[P_t - \frac{P_t C_{t+1} C_{t+1}^\top P_t}{\lambda + C_{t+1}^\top P_t C_{t+1}} \right] \frac{1}{\lambda} - \alpha P_t^2$$

SF1

Originally proposed by J.E. Parkum

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{P_t C_{t+1}}{1 + C_{t+1}^\top P_t C_{t+1}} \varepsilon_{t+1}$$

$$P_{t+1} = \frac{\alpha_1 - \alpha_0}{\alpha_1} \left[P_t - \frac{P_t C_{t+1} C_{t+1}^\top P_t}{1 + C_{t+1}^\top P_t C_{t+1}} \right] + \alpha_0 I$$

$$\alpha_0 \leq \lambda(P_t) \leq \alpha_1$$

Forgetting methods - in summary

$$
\begin{aligned}
\hat{\theta}_{t+1} &= \hat{\theta}_t + P_{t+1} C_{t+1} \varepsilon_{t+1} \\
P_{t+1}^{-1} &= \lambda_1 P_t^{-1} + \lambda_2 C_{t+1} C_{t+1}^{\top}
\end{aligned}
$$

In general: $\lambda_i(\varepsilon_t, P_t, C_t) \;\; i = 1, 2.$

ε : *Degree of explanations*

P : *Information book keeping*

C : *incomming information*

Bounded trace
Constant trace
Hägglund

Model methods

Kalman filter

Also denoted as linear forgetting.

$$\begin{aligned} \theta_{t+1} &= \theta_t + v_t & = Ax_t + v_t \\ y_t &= C_t^\top \theta_t + e_t & = Cx_t + e_t \end{aligned}$$

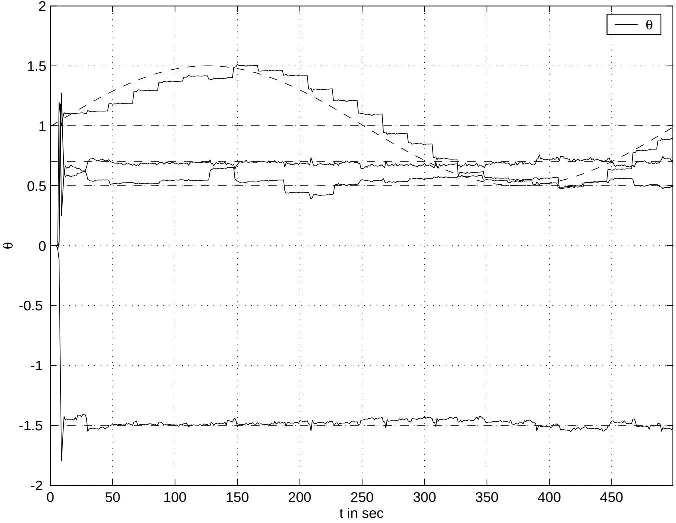
Data update

$$\begin{aligned} \hat{\theta}_{t|t-1} &= \hat{\theta}_{t-1|t-1} + P_t C_t \varepsilon_t \\ P_{t|t-1}^{-1} &= P_{t-1|t-1}^{-1} + C_t C_t^\top \end{aligned}$$

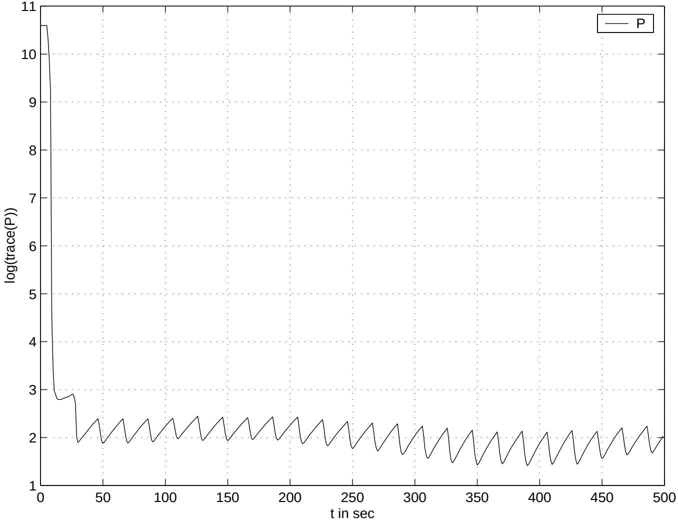
Time update

$$\begin{aligned} \hat{\theta}_{t+1|t} &= \hat{\theta}_{t|t} \\ P_{t+1|t} &= P_{t|t} + R_1 \end{aligned}$$

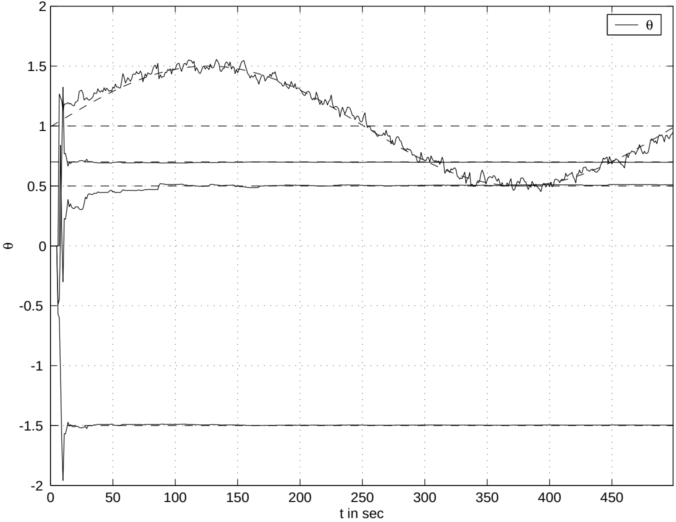
sqwave as exitation



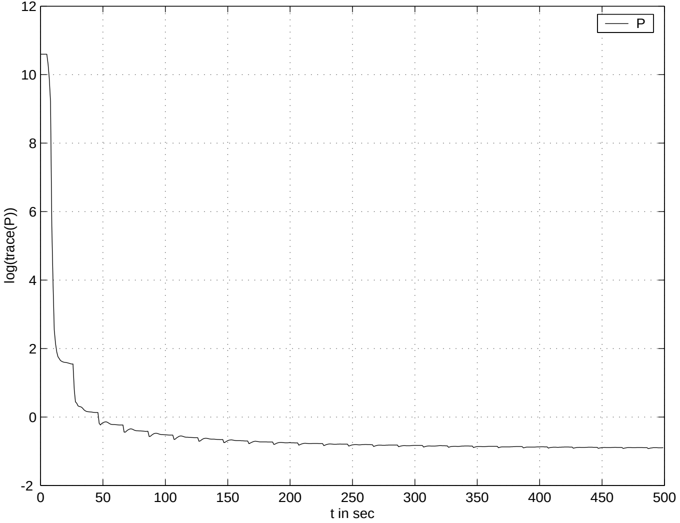
sqwave as exitation



PRBS as exitation



PRBS as exitation



Linear growth

$$y_t = C_t^\top \theta_t + e_t$$

$$\theta_t = \theta_0 + \alpha\, t$$

$$\begin{aligned} y_t &= C_t^\top \left[\theta_0 + t\,\alpha\right] + e_t \\ &= \begin{bmatrix} C_t^\top & tC_t^\top \end{bmatrix} \begin{bmatrix} \theta_0 \\ \alpha \end{bmatrix} + e_t \end{aligned}$$

DTU

Piecewise linear growth

$$y_t = C_t^\top \theta_t + e_t$$

$$\theta_t = \theta_{T_i} + (t - T_i)\alpha_i \qquad T_i \le t < T_{i+1}$$

$$\begin{aligned} y_t &= C_t^\top \left[\theta_{T_i} + (t - T_i)\alpha_i\right] + e_t \\ &= \begin{bmatrix} C_t^\top & (t - T_i)C_t^\top \end{bmatrix} \begin{bmatrix} \theta_{T_i} \\ \alpha_i \end{bmatrix} + e_t \end{aligned}$$

Example

$$y_t + a y_{t-1} = b u_{t-1} + e_t$$

$$b = b_0 + b_{11} cos(\omega t) + b_{12} sin(\omega t) + ...$$

$$\begin{aligned} y_t &= -a y_{t-1} + b_0 u_{t-1} + b_{11} cos(\omega t) u_{t-1} + b_{12} sin(\omega t) u_{t-1} + e_t \\ &= C_t^\top \theta + e_t \end{aligned}$$

$$\theta = \begin{bmatrix} a \\ b_0 \\ b_{11} \\ b_{12} \end{bmatrix} \quad C = \begin{bmatrix} -y_{t-1} \\ u_{t-1} \\ u_{t-1}\ cos(\omega t) \\ u_{t-1}\ sin(\omega t) \end{bmatrix}$$

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Adaptive systems (L22)

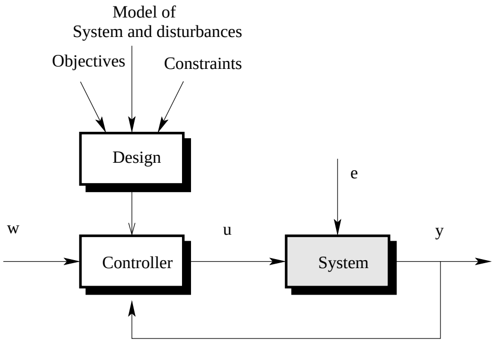
Wikipedia: Adaptive behavior is a type of behavior that is used to adjust to another type of behavior or situation.

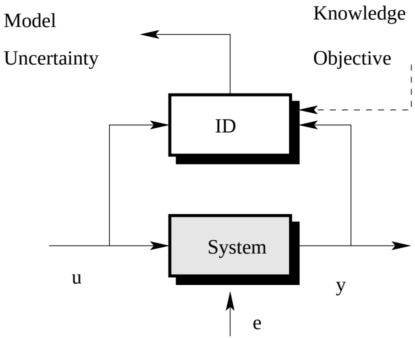
Here: device, algorithm or method with the ability ajust itself (or its behavior) to the actual system.

Prediction, Filtering and smoothing

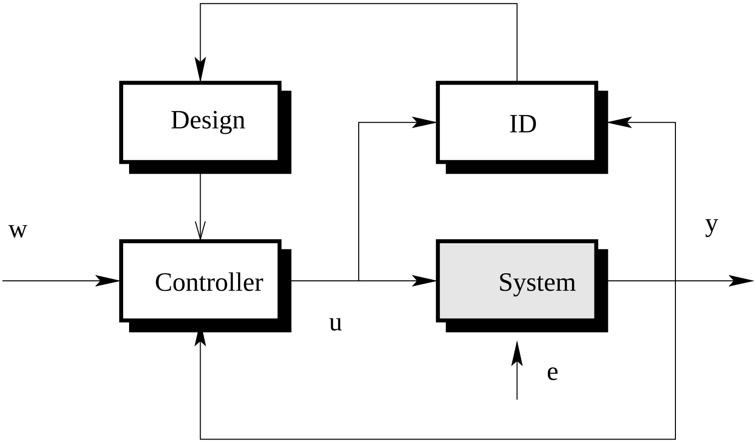
Detection, isolation and fault estimation

Control


Model of
System and disturbances
Objectives
Constraints
Design
e
w
u
y
Controller
System

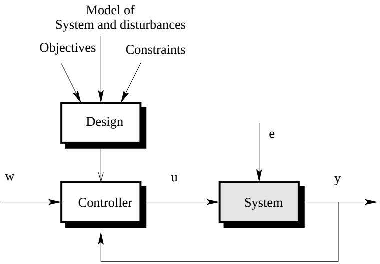

Model
Uncertainty
Knowledge
Objective
ID
System
u
y
e

Self Tuning Controller (STC)

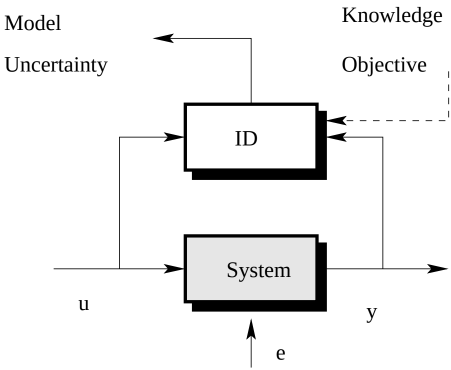


PID: Process information around w_c. Robust. Not necessarily optimal (rarely optimal wrt. disturbances).

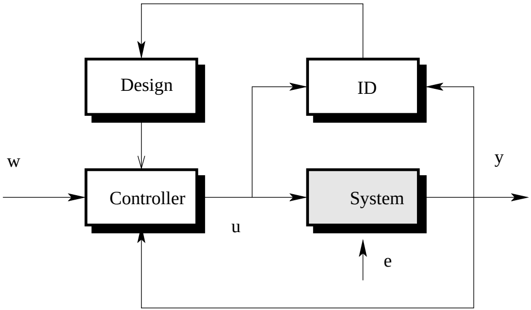
Stochastic Control: requires a precise model (also of the disturbances).



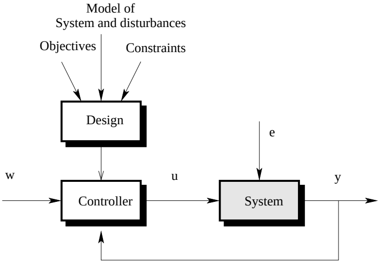
System Identification: Parameter estimation. Validation.

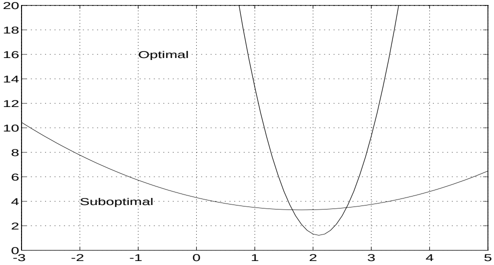


Adaptive Control: Parameter changes (time variations, nonlinear system).



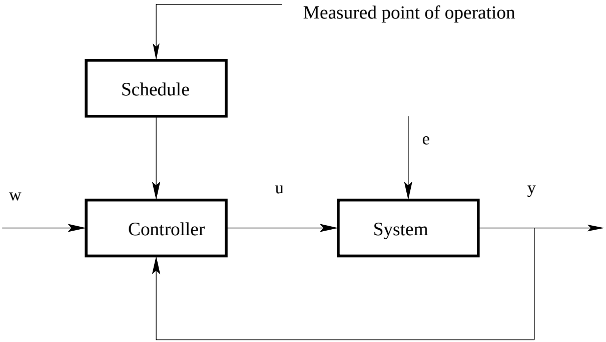
Other approach: Robust Control





→ hhn et al.

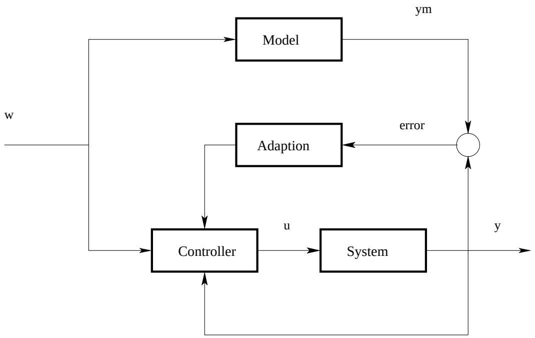
Gain Scheduling (GS), Linear Parameter Variation (LPV)

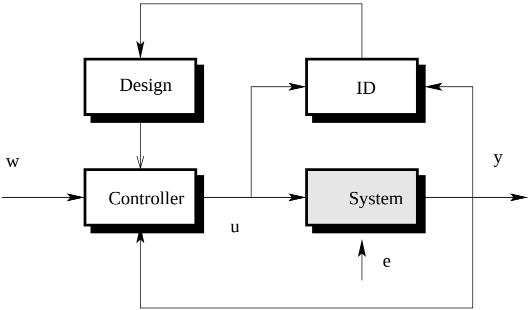


Examples: Tank system. Wind turbine. Ship. Aircraft.

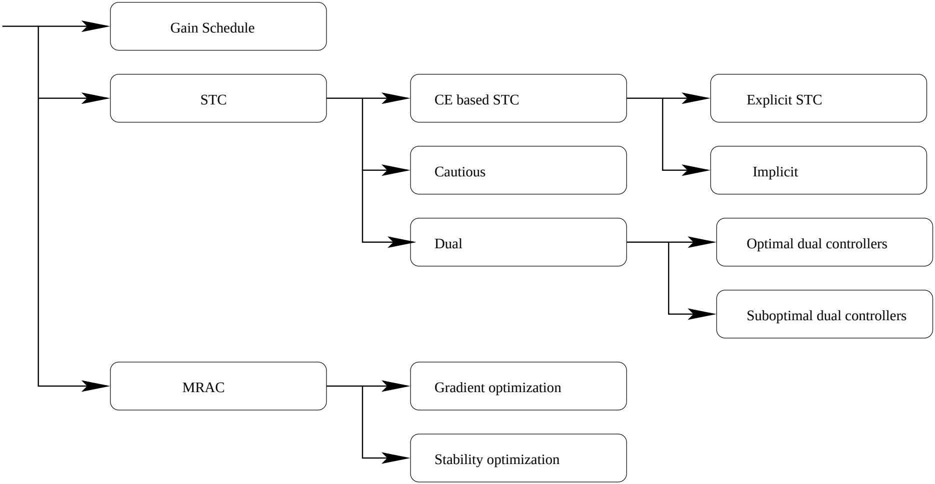
To be considered as an open loop adaptive controller or a methods for dealing with nonlinear systems.

Model Reference Adaptive Control




Design
ID
w
y
Controller
System
u
e

Classification of adaptive controllers



Recap: Polynomials and vectors

Let us consider the result of a polynomium operating on a signal

$$\begin{aligned} S(q^{-1})y_t &= s_0 y_t + s_1 y_{t-1} + \dots + s_n y_{t-n} \\ &= \gamma^T \vartheta \end{aligned}$$

where

$$\gamma^T = \begin{bmatrix} y_t & y_{t-1} & \dots & y_{t-n} \end{bmatrix}$$

$$\vartheta^T = \begin{bmatrix} s_0 & s_1 & \dots & s_n \end{bmatrix}$$

DTU

Explicit version of Basic self Tuner

The Basic Self Tuner (Explicit version)

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t$$

ID:

$$y_t = C_t^\top \theta + e_t$$

$$\hat{\theta}_t = arg\ Min \sum_{i=0}^{t} \varepsilon_i^2$$ pem or plr

---

Control: MV.

$$u_t = arg\ Min\ \mathbf{E}\left\{ y_{t+k}^2 \right\}$$

ID: The Basic Self Tuner - (Explicit version)

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t$$

ID:

$$\begin{aligned} \theta^\top &= (\dots\ a_i,\ \dots\ b_i,\ \dots\ c_i\ \dots) \\ C_t^\top &= (\dots\ -y_{t-i},\ \dots\ u_{t-i},\ \dots\ e_{t-i}\ \dots) \end{aligned}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{\bar{P}_t \psi_t}{1 + \psi_t^\top \bar{P}_t \psi_t}\,\varepsilon_t \qquad \varepsilon_t = y_t - C_t^\top \hat{\theta}_t$$

$$P_t = \bar{P}_t - \frac{\bar{P}_t \psi \psi_t^\top \bar{P}_t}{1 + \psi_t^\top \bar{P}_t \psi_t}$$

$$\bar{P}_t = funk\left(P_{t-1}, \varepsilon_t, \hat{C}_t, \psi_t\right) \qquad \text{Forgetting}$$

$$\psi_t = \frac{1}{\hat{C}_{t-1}} \hat{C}_t$$

$$J = \mathbf{E}\left\{ y_{t+k}^2 \right\}$$

Minimum variance control

$$\hat{C} = \hat{A}G + q^{-k}S$$

$$y_{t+k} = \frac{1}{\hat{C}}\left[\hat{B}Gu_t + Sy_t\right] + Ge_{t+k}$$

$$Ru_t = -Sy_t \qquad R = \hat{B}G$$

Controller:

$$Ru_t + Sy_t = 0$$

$$\gamma_t = (\, u_t,\ u_{t-1},\ \ldots\ y_t,\ y_{t-1},\ \ldots\,)^\top$$

$$\vartheta = (\, r_0,\ ,r_1,\ \ldots\ s_0,\ s_1,\ \ldots\,)^\top$$

$$u_t = arg\ Sol\left(\ \gamma_t^T \vartheta = 0\ \right)$$

Example

System:

$$y_t - 0.9\, y_{t-1} = 1\, u_{t-2} + e_t$$

NB: an ARX system

Model:

$$y_t + \hat{a}\, y_{t-1} = \hat{b}\, u_{t-2} + \varepsilon_t$$

Estimation:

$$\varepsilon_t = y_t + \hat{a}_{t-1} y_{t-1} - \hat{b}_{t-1} u_{t-2}$$

$$C_t^T = \begin{bmatrix} -y_{t-1} & u_{t-2} \end{bmatrix} \qquad \theta^T = \begin{bmatrix} a & b \end{bmatrix}$$

$$\begin{aligned} \hat{\theta}_t &= \hat{\theta}_{t-1} + \frac{\bar{P}_t C_t}{1 + C_t^\top \bar{P}_t C_t} \varepsilon_t \\ P_t &= \bar{P}_t - \frac{\bar{P}_t C_t C_t^\top \bar{P}_t}{1 + C_t^\top \bar{P}_t C_t} \end{aligned}$$

$$\bar{P}_t = funk\left(P_{t-1}, \varepsilon_t, \hat{C}_t, \psi_t\right)$$

$$J_e = \sum_{i=1}^{t} \varepsilon_i^2 \simeq t\sigma^2 \qquad \varepsilon_t = e_t \quad \text{for correct parameters}$$

Design:

$$1 = (1 + \hat{a}_t q^{-1})(1 + g q^{-1}) + q^{-2} s$$

$$G = 1 - \hat{a}_t q^{-1} \qquad S = \hat{a}_t^2$$

Controller:

$$\hat{b}_t (1 - \hat{a}_t q^{-1}) u_t = -\hat{a}_t^2 y_t$$

or

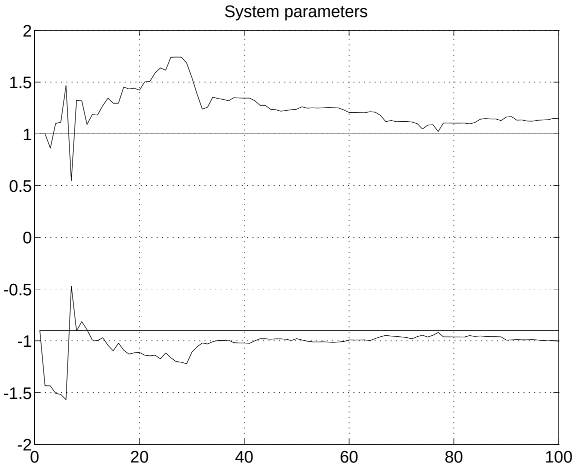
$$u_t = \hat{a}_t u_{t-1} - \frac{\hat{a}_t^2}{\hat{b}_t} y_t$$

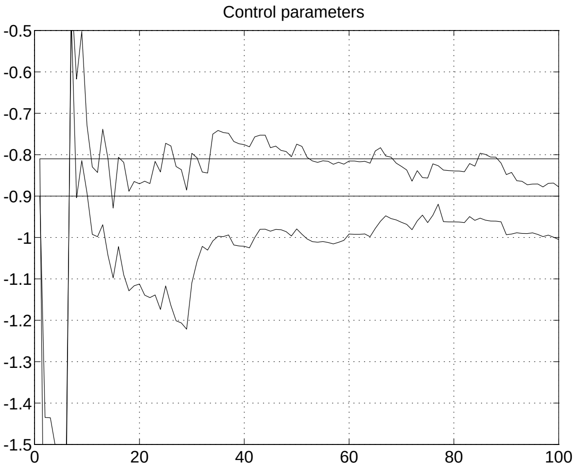
$$J_c = \sum_{i=0}^{t} y_i^2 \simeq \mathbf{E}\left\{ y_t^2 \right\} t \simeq 1.81 \sigma^2 \; t$$

In closed loop (for correct parameters)

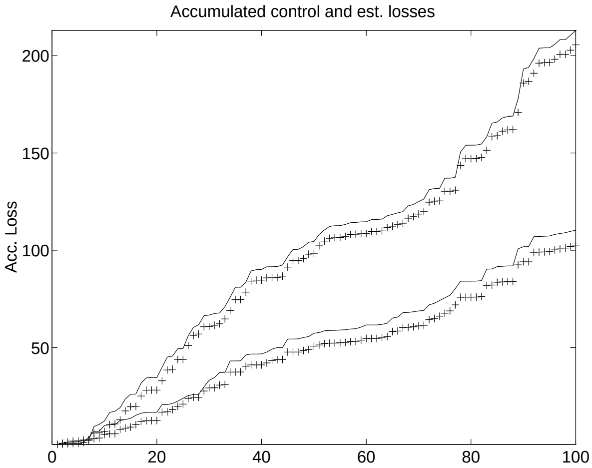
$$y_t = (1 + 0.9 q^{-1}) e_t$$

Example



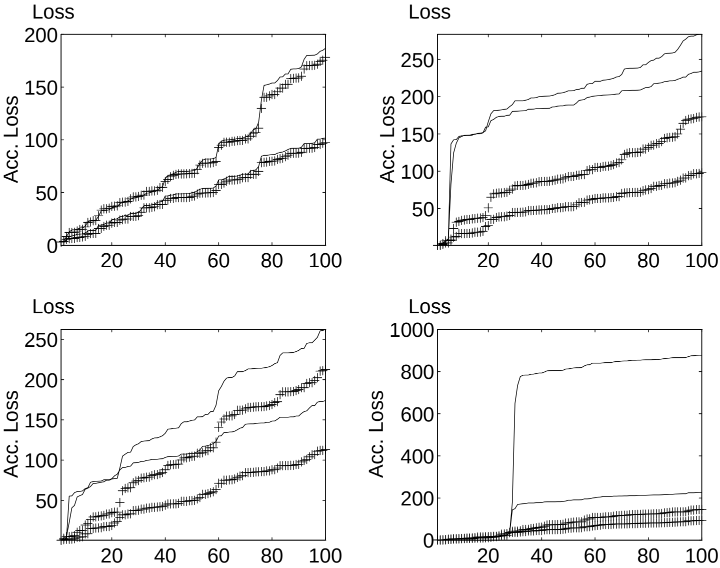

Control parameters
-0.5
-0.6
-0.7
-0.8
-0.9
-1
-1.1
-1.2
-1.3
-1.4
-1.5
0
20
40
60
80
100

Example



Accumulated control and est. losses
Acc. Loss
200
150
100
50
0
20
40
60
80
100

Example


Loss
Acc. Loss
200
150
100
50
0
20
40
60
80
100
Loss
Acc. Loss
250
200
150
100
50
20
40
60
80
100
Loss
Acc. Loss
250
200
150
100
50
20
40
60
80
100
Loss
Acc. Loss
1000
800
600
400
200
0
20
40
60
80
100

Control: The Basic Self Tuner - Explicit version

$$J = \mathbf{E}\left\{ y_{t+k}^2 \right\} \qquad A(q^{-1})y_t = B(q^{-1})u_{t-k} + C(q^{-1})e_t$$

$$\hat{C} = \hat{A}G + q^{-k}S$$

$$y_{t+k} = \frac{1}{\hat{C}}\left[\hat{B}Gu_t + Sy_t\right] + Ge_{t+k}$$

$$Ru_t = -Sy_t \qquad R = \hat{B}G$$

Controller:

$$Ru_t + Sy_t = 0$$

$$\begin{aligned} \gamma_t &= \left(u_t,\ u_{t-1},\ \ldots\ y_t,\ y_{t-1},\ \ldots \right)^\top \\ \vartheta &= \left(r_0,\ ,r_1,\ \ldots\ s_0,\ s_1,\ \ldots \right)^\top \end{aligned}$$

$$u_t = arg\ Sol\left(\gamma_t^T \vartheta = 0 \right)$$

DTU

$$J = \mathbf{E}\left\{ y_{t+1}^2 \right\} \qquad A(q^{-1})y_t = B(q^{-1})u_{t-1} + e_t$$

$$1 = \hat{A} + q^{-1}S \qquad S = q(1 - \hat{A}) = \tilde{A}$$

$$y_{t+1} = \left[\hat{B}u_t + Sy_t\right] + e_{t+1}$$

$$Ru_t = -Sy_t \qquad R = \hat{B}$$

Controller:

$$Ru_t + Sy_t = 0$$

$$\begin{aligned} \gamma_t &= (\, u_t,\ u_{t-1},\ \dots\ y_t,\ y_{t-1},\ \dots\,)^\top \\ \vartheta &= (\, r_0,\ ,r_1,\ \dots\ s_0,\ s_1,\ \dots\,)^\top \end{aligned}$$

$$u_t = arg\ Sol\left(\ \gamma_t^T\vartheta = 0\ \right)$$

Implicit version of Basic self Tuner

The Basic Self Tuner II (Implicit version)

$$A(q^{-1})y_t = B(q^{-1})u_{t-1} + e_t \qquad C = 1 \qquad k = 1$$

MV: $J = \mathbf{E}\left\{y_{t+1}^2\right\}$ $\qquad C = AG + q^{-k}S$

$$y_{t+k} = \frac{1}{C}\left[BGu_t + Sy_t\right] + Ge_{t+k}$$

$$Ru_t = -Sy_t \qquad R = BG$$

$$1 = A * 1 + q^{-1}S \qquad G = 1 \qquad S = q(1-A) = -a_1 - a_2q^{-1} - \ \dots\ - a_nq^{1-n}$$

$$y_{t+1} = [Bu_t + Sy_t] + e_{t+1} \qquad Bu_t = -Sy_t$$

Notice the simple design

$$y_{t+1} = [Sy_t + Ru_t] + e_{t+1} = C_{t+1}^\top\theta + e_{t+1}$$

$$C_{t+1}^\top = \begin{bmatrix} y_t & y_{t-1} & \dots & u_t & \dots \end{bmatrix}$$

$$\theta^T = \begin{bmatrix} s_0 & s_1 & \dots & r_0 & \dots \end{bmatrix}$$

$$\hat{\theta}: \qquad y_t = C_t^\top\hat{\theta}_{t-1} + \varepsilon_t; \qquad J_e = \frac{1}{t}\sum_{i=0}^{t}\varepsilon_i^2$$

$$u_t: \qquad C_{t+1}^\top\hat{\theta}_t = 0 \qquad J_c = \mathbf{E}\left\{y_{t+1}^2\right\}$$

The Basic Self Tuner III

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + e_t$$

$$C = 1$$

MV: $J = \mathbf{E}\left\{ y_{t+k}^2 \right\}$ $\qquad BGu_t = -Sy_t$

$$1 = AG + q^{-k}S$$

$$y_{t+k} = [Sy_t + BGu_t] + Ge_{t+k}$$

$$\begin{aligned} y_{t+k} &= [Sy_t + Ru_t] + Ge_{t+k} \\ &= C_{t+k}^\top \theta + Ge_{t+k} \end{aligned}$$

$\hat{\theta}$: $\quad y_t = C_t^\top \hat{\theta}_{t-1} + \varepsilon_t; \qquad J_e = \frac{1}{t}\sum_{i=0}^{t} \varepsilon_i^2$

u_t : $\quad C_{t+k}^\top \hat{\theta}_t = 0 \qquad J_c = \mathbf{E}\left\{ y_{t+k}^2 \right\}$

The Basic Self Tuner IV

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t$$

MV: $J = \mathbf{E}\left\{y_{t+k}^2\right\}$ $\qquad BGu_t = -Sy_t$

$$C = AG + q^{-k}S$$

$$y_{t+k} = \frac{1}{C}\left[Sy_t + BGu_t\right] + Ge_{t+k}$$

RLS:
$$\begin{aligned} y_{t+k} &= [Sy_t + Ru_t] + Ge_{t+k} \\ &= C_{t+k}^\top \theta + Ge_{t+k} \end{aligned}$$

Notice the missing C

$\hat{\theta}$: $\qquad y_t = C_t^\top \hat{\theta}_{t-1} + \varepsilon_t; \qquad J_e = \frac{1}{N}\sum_{i=0}^{t} \varepsilon_i^2$

u_t : $\qquad C_{t+k}^\top \hat{\theta}_t = 0 \qquad J_c = \mathbf{E}\left\{y_{t+k}^2\right\}$

Implicite STC

Advantage:

- Design simple,
- RLS (even if $C \neq 1$),
- $J_e \simeq J_c$

Disadvantage:

- More parameters ($k >> 1$),
- Not all strategies can be transformed into an implicitte strategy.

Convergence analysis

θ_0 a possible convergence point.

$$\mathcal{S} \quad : \qquad y_t = \frac{1}{C} C_t^\top \theta_0 + Ge_t$$

$$\mathcal{M} \quad : \qquad y_t = \quad C_t^\top \hat{\theta} + \varepsilon_t$$

$$\mathbf{E}\Big\{C_t \varepsilon_t\Big\} = 0$$

$$\begin{aligned} \varepsilon_t &= y_t - C_t^\top \hat{\theta} = \frac{1}{C} C_t^\top \theta_0 + Ge_t - C_t^\top \hat{\theta} \\ &= Ge_t \ + \frac{1-C}{C} C_t^\top \theta_0 \ \text{ for } \ \hat{\theta} = \theta_0 \end{aligned}$$

DTU

Implicite STC.

Synergy

Control: $J_c = \mathbf{E}\left\{y_{t+k}^2\right\} = \mathbf{E}\left\{\varepsilon_{t+k}^2\right\}$

Estimation: $J_e = \dfrac{1}{N}\sum \varepsilon_i^2$

Fixation

Controller:

$$u_t = -\frac{S}{R} y_t$$

Model:

$$y_{t+k} = [Ru_t + Sy_t] + Ge_{t+k}$$

$$\frac{1}{2} < \frac{r_0}{b_0}$$

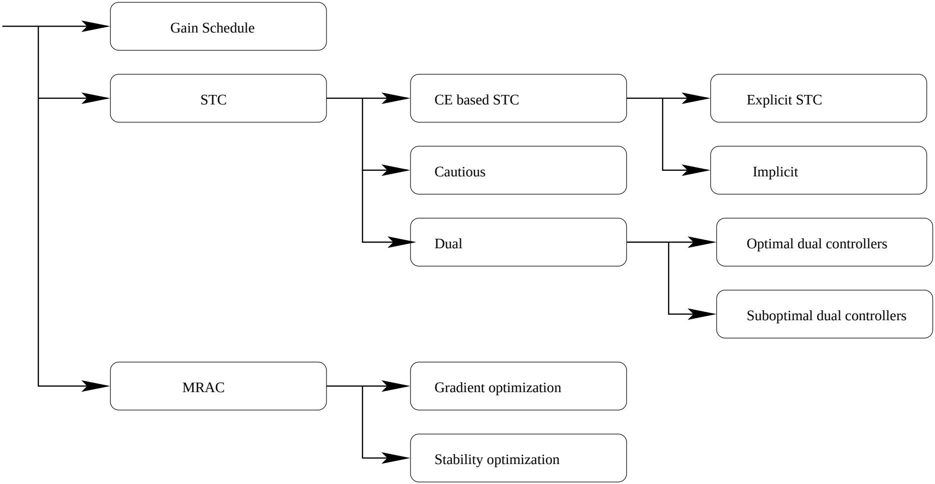
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

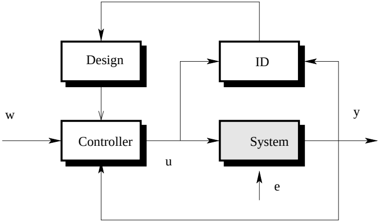
Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Adaptive Control II (L23)

Classification of adaptive controllers



Explicit STC



Also denoted as indirect adaptive methods.

- Minimalvariance, MV_0
- PZ, GSP
- GMV (MV_i)
- GPC (or MPC)
- LQG
- Deadbeat, PID ao.

- Kalmanfilter/observer
- Polplacement controller
- LQG
- Robust

Explicit Adaptive Control (STC) $=$ Indirect Adaptive Control (MRAC)
The system parameters are estimated and transformed to a set of control parameter in the Design block. The adaptation mechanisme (estimation procedure) is applied indirectly (ie. not directly) on the control parameters.

Implicit Adaptive Control (STC) $=$ Direct Adaptive Control (MRAC)
The model is rephrased in terms of the control parameters, which are estimated. The adaptation mechanism (estimation procedure) is working on the control parameters directly.

Explicit MV$_0$ control

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

Cost:

$$J = \mathbf{E}\left\{(y_{t+k} - w_t)^2\right\}$$

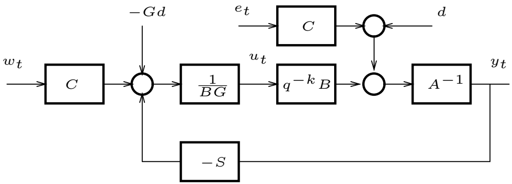
Controller:

$$BGu_t = Cw_t - Sy_t - Gd$$

Design:

$$C = AG + q^{-k}S$$

$$G(0) = 1 \qquad ord(G) = k - 1 \qquad ord(S) = max(n_a - 1, n_c - k)$$



Main loop

```
measinit;                                % Initilialise the measurement system
for it=1:nstp,
 w=wt(it);
 [y,t]=meas;                             % Measure output

 u=...

 act(u);                                  % Actuate control
end
```

Cannonical realization

$$Ru_t = Qw_t - Sy_t - \delta$$

$$\begin{aligned} x^r_{t+1} &= A^r x^r_t + B^r \begin{bmatrix} w_t \\ y_t \end{bmatrix} \\ u_t &= C^r x^R_t + D^r \begin{bmatrix} w_t \\ y_t \end{bmatrix} + u_0 \end{aligned}$$

DTU

Cannonical realization

```
%-------------------------------------------------------------------------
[A,B,k,C,s2]=sysinit(dets); % Determine linear model (ie. get system)
%-------------------------------------------------------------------------
[Q,R,S,G]=dsnmv0(A,B,k,C);
%-------------------------------------------------------------------------
[Ar,Br,Cr,Dr]=armax2ss(R,Q,0,-S);
nr=length(Ar); Xr=zeros(nr,1);
%-------------------------------------------------------------------------

measinit;         % Initilialise the measurement system
for it=1:nstp,
 w=wt(it); wf=wft(it);
 [y,t]=meas;

 u=Cr*Xr+Dr*[wf;y]+u0;

 act(u);  % Actuate control
 Xr=Ar*Xr+Br*[wf;y];
end
```

Direct realization

$$Ru_t = Qw_t - Sy_t - \delta \qquad\qquad Ru_t - Qw_t + Sy_t + \delta = 0$$

$$\underline{C}_t^T = \left[\, u_t,\ u_{t-1},\ \ldots\ -w_t,\ -w_{t-1},\ \ldots\ y_t,\ y_{t-1},\ \ldots\ 1 \right] \qquad\qquad \underline{C} = C_r$$

$$\underline{\theta}^T = \quad \left[\, r_0,\ r_1,\ \ldots \qquad q_0, \qquad q_1,\ \ldots \qquad s_0,\ \ s_1,\ \ldots\ \delta \right]$$

$$u_t = -\frac{\underline{\tilde{C}}_t^T \underline{\tilde{\theta}}}{r_0} \qquad\qquad \underline{\tilde{\theta}} = \underline{\theta}(2 : end);$$

$$\underline{C}_{t-1}^T = \left[\, u_{t-1},\ u_{t-2},\ \ldots\ -w_{t-1},\ -w_{t-2},\ \ldots\ y_{t-1},\ y_{t-2},\ \ldots\ 1 \right]$$

Direct realization

```
%------------------------------------------------------------------------
[A,B,k,C,d,s2]=sysinit(dets); % Determine linear model (ie. get system)
%------------------------------------------------------------------------
[Q,R,S,G]=dsnmv0(A,B,k,C);
%------------------------------------------------------------------------
nr=length([R Q S])+1; fir=zeros(nr,1); thr=[R Q S G(1)*d]';
pil=1+[0 length(R) length([R Q])];
%------------------------------------------------------------------------
measinit;          % Initilialise the measurement system
for it=1:nstp,
 w=wt(it);
 [y,t]=meas;                          % Measure output

 % Ru=Qw-Sy-Gd
 fir(2:end)=fir(1:end-1);
 fir(pil)=[0 -w y];
 u=-fir'*thr/thr(1);
 fir(1)=u;

 act(u);  % Actuate control
end
```

ID (Explicit version)

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

ID:

$$\begin{aligned} \theta^\top &= (\ldots\ a_i,\ \ldots\ b_i,\ \ldots\ c_i\ \ldots\ d) \\ C_t^\top &= (\ldots\ -y_{t-i},\ \ldots\ u_{t-i},\ \ldots\ e_{t-i}\ \ldots\ 1) \end{aligned}$$

$$\begin{aligned} \hat{\theta}_t &= \hat{\theta}_{t-1} + \frac{\bar{P}_t \psi_t}{1 + \psi_t^\top \bar{P}_t \psi_t}\, \varepsilon_t \qquad \varepsilon_t = y_t - C_t^\top \hat{\theta}_t \\ P_t &= \bar{P}_t - \frac{\bar{P}_t \psi \psi_t^\top \bar{P}_t}{1 + \psi_t^\top \bar{P}_t \psi_t} \end{aligned}$$

$$\bar{P}_t = funk\left(P_{t-1}, \varepsilon_t, \hat{C}_t, \psi_t\right)$$ Forgetting

$$\psi_t = \frac{1}{\hat{C}_{t-1}} \hat{C}_t$$

Explicit MV_0 controller

```
%-------------------------------------------------------------------------
[A,B,k,C,d,s2]=sysinit(dets); % Determine linear model (ie. get system)
%-------------------------------------------------------------------------
na=length(A)-1; mb=length(B); nc=length(C)-1;
th=[A(2:end) B C(2:end) d]'; th0=th;
th=th*0;
pil=[na mb];
pil=[0 cumsum(pil)]+1;
fi=zeros(size(th));
p0=10000;
P=eye(size(th,1))*p0;
%-------------------------------------------------------------------------
[Q,R,S,G]=dsnmv0(A,B,k,C); % just for the structure
%-------------------------------------------------------------------------
nr=length([R Q S]); fir=zeros(nr,1); fir=[fir; 1]; thr=[R Q S G(1)*d]';
pilr=1+[0 length(R) length([R Q])];
%-------------------------------------------------------------------------
```

Explicit MV_0 controller

Main loop

```
measinit;          % Initilialise the measurement system
for it=1:nstp,
 w=wt(it);
 [y,t]=meas;                          % Measure output

 % ID block
 res=y-fi'*th;
 K=P*fi/(1+fi'*P*fi);
 P=P-K*fi'*P;
 th=th+K*res;

 % Design block
 A=[1 th(1:na)']; B=th(pil(2):pil(3)-1)';
 C=[1 th(pil(3):end)']; d=th(end);
 [Q,R,S,G]=dsnmv0(A,B,k,C);
 thr=[R Q S G(1)*d]';
```

Explicit MV_0 controller

```
 % Ru=Qw-Sy                          Controller
 fir(2:end)=fir(1:end-1);
 fir(pilr)=[0 -w y];
 u=-fir'*thr/thr(1);
 fir(1)=u;

 fi(2:end)=fi(1:end-1);
 fi(pil(1:2))=[-y u]';

 act(u);  % Actuate control
end
```

System:

$$Ay_t = q^{-k}Bu_t + Ce_t + d$$

Cost:

$$J = \mathbf{E}\left\{(A_m y_{t+k} - B_m w_t)^2\right\}$$

Controller:

$$BGu_t = B_m C w_t - S y_t - Gd$$

Design:

$$A_m C = AG + q^{-k} S$$

$$G(0) = 1 \qquad ord(G) = k-1 \qquad ord(S) = max(n_a - 1, n_c + n_{a_m} - k)$$

Explicit PZ controller

```
%-------------------------------------------------------------------------
[A,B,k,C,d,s2]=sysinit(dets); % Determine linear model (ie. get system)
%-------------------------------------------------------------------------
na=length(A)-1; mb=length(B); nc=length(C)-1;
th=[A(2:end) B C(2:end)d]'; th0=th;
th=th*0;
pil=[na mb];
pil=[0 cumsum(pil)]+1;
fi=zeros(size(th));
p0=10000;
P=eye(size(th,1))*p0;
%-------------------------------------------------------------------------
Am=[1 -0.6];
Bm=sum(Am);
[Q,R,S,G]=dsnpz(A,B,k,C,Am,Bm); % just for the structure
%-------------------------------------------------------------------------
nr=length([R Q S]); fir=zeros(nr,1); fir=[fir; 1]; thr=[R Q S G(1)*d]';
pilr=1+[0 length(R) length([R Q])];
%-------------------------------------------------------------------------
```

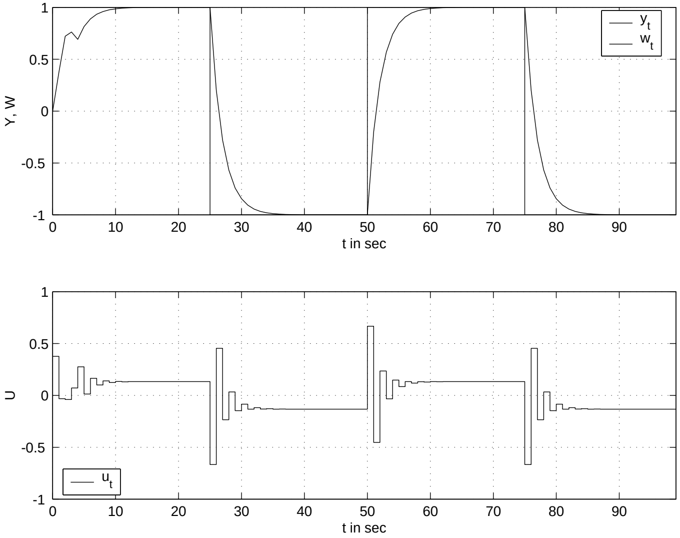
Main loop

```
measinit;          % Initilialise the measurement system
for it=1:nstp,
 w=wt(it);
 [y,t]=meas;                    % Measure output

 % ID block
 res=y-fi'*th;
 K=P*fi/(1+fi'*P*fi);
 P=P-K*fi'*P;
 th=th+K*res;

 % Design block
 A=[1 th(1:na)']; B=th(pil(2):pil(3)-1)';
 C=[1 th(pil(3):end)']; d=th(end);
 [Q,R,S,G]=dsnpz(A,B,k,C,Am,Bm);
 thr=[R Q S G(1)*d]';
```

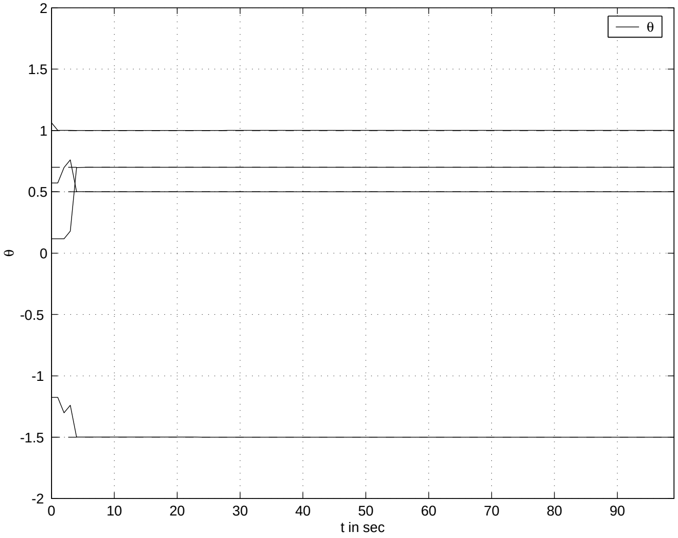
Explicit PZ-control

```
 % Ru=Qw-Sy                          Controller
 fir(2:end)=fir(1:end-1);
 fir(pilr)=[0 -w y];
 u=-fir'*thr/thr(1);
 fir(1)=u;

 fi(2:end)=fi(1:end-1);
 fi(pil(1:2))=[-y u]';

 act(u);  % Actuate control
end
```

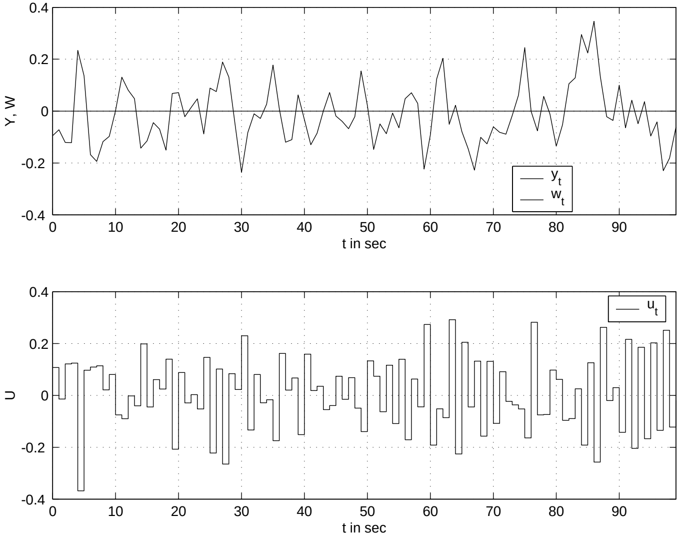
Explicit PZ-control



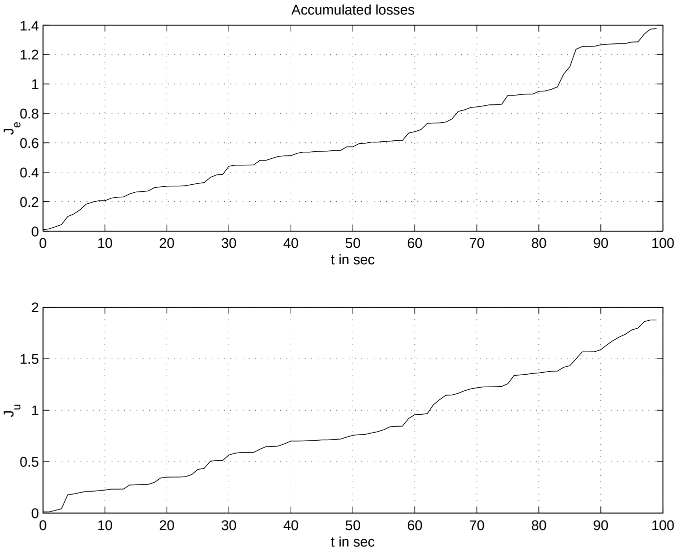
Explicit PZ-control



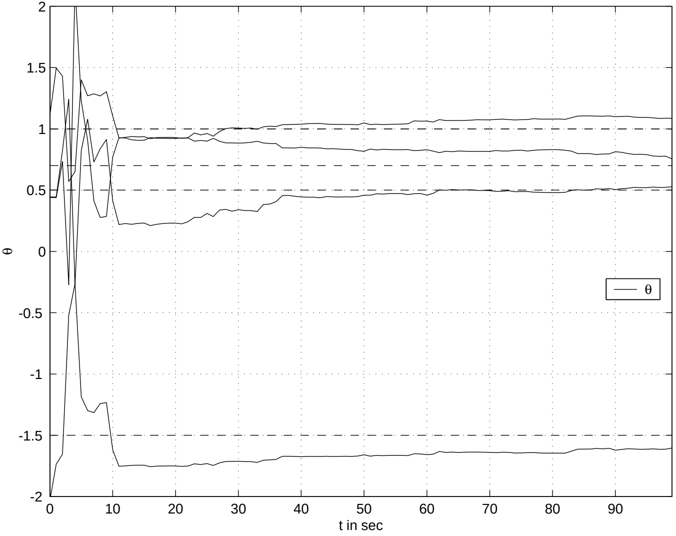
Explicit PZ-control

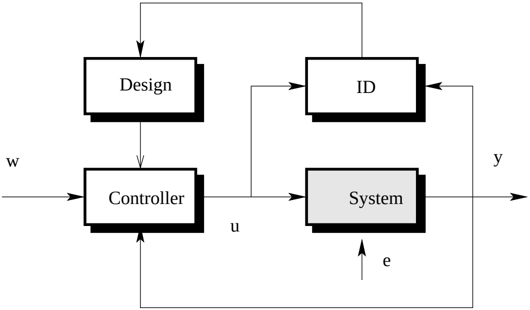


Explicit PZ-control



Explicit PZ-control





Implicit Adaptive Control (STC) = Direct Adaptive Control (MRAC)
The model is rephrased in terms of the control parameters, which are estimated. The adaptation mechanism (estimation procedure) is working on the control parameters directly.

$$J = \mathbf{E}\left\{(y_{t+k} - w_t)^2\right\}$$

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d$$

$$C = AG + q^{-k}S \qquad R = BG \qquad \delta = G(1)d$$

$$\begin{aligned}
\xi_{t+k} &= y_{t+k} - w_t \\
&= \frac{1}{C}\left[Sy_t + Ru_t - Cw_t + \delta\right] + Ge_{t+k} \\
&= \frac{1}{C}C_{t+k}^{\top}\theta + \bar{e}_{t+k}
\end{aligned}$$

Model:
$$\begin{aligned}
\xi_t &= y_t - w_{t-k} \\
&= C_t^{\top}\hat{\theta}_{t-1} + \varepsilon_t
\end{aligned}$$

Control:
$$u_t = argSol\left(C_{t+k}^{\top}\hat{\theta}_t = 0.\right)$$

Implicit MV$_0$

RLS-algorithm
Computer burden (design)
Active (estimation) and passive (control) have similar cost function.
Has to know k.
The min phase problem.
The b_0 problem.

Implicit MV_0

$$J = \mathbf{E}\Big\{(y_{t+k} - w_t)^2\Big\}$$

1 Measure y_t.
2 Create $\xi_t = y_t - w_{t-k}$.
3 Create $C_t = (y_{t-k},\ \ldots,u_{t-k},\ \ldots,-w_{t-k},\ 1)^T$ and C_{t+k}.
4 Update the estimates:

$$\epsilon_t = \xi_t - C_t\hat{\theta}_{t-1} \tag{12}$$

$$P_t^{-1} = P_{t-1}^{-1} + C_t C_t^T \tag{13}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + P_t C_t \epsilon_t \tag{14}$$

5 Determine u_t such that:

$$C_{t+k}^T \hat{\theta}_t = 0 \tag{15}$$

6 Actuate the control.

Implicit PZ-regulator

$$J = \mathbf{E}\Big\{(A^m(q^{-1})y_{t+k} - B^m(q^{-1})w_t)^2\Big\}$$

$$A^m C = AG + q^{-k}S \qquad R = BG \quad Q = B^m C \quad \delta = Gd$$

$$\begin{aligned} \xi_{t+k} &= A^m y_{t+k} - B^m w_t \\ &= \frac{1}{C}\left[Sy_t + Ru_t - Qw_t + \delta\right] + Ge_{t+k} \\ &= \frac{1}{C} C_{t+k}^\top \theta + \bar{e}_{t+k} \end{aligned}$$

Model: $$\begin{aligned} \xi_t &= A^m y_t - B^m w_{t-k} \\ &= C_t^\top \hat{\theta}_{t-1} + \varepsilon_t \end{aligned}$$

Control: $$u_t = argSol\left(C_{t+k}^\top \hat{\theta}_t = 0.\right)$$

$$J = \mathbf{E}\Big\{(A^m(q^{-1})y_{t+k} - B^m(q^{-1})w_t)^2\Big\}$$

1. Measure y_t.
2. Create $\xi_t = A_m(q^{-1})y_t - B_m(q^{-1})w_{t-k}$.
3. Create $C_t = (y_{t-k},\ \dots, u_{t-k},\ \dots, -w_{t-k},\ 1)^T$ og C_{t+k}^T.
4. Update the estimates:

$$\epsilon_t = \xi_t - C_t\hat{\theta}_{t-1} \tag{16}$$

$$P_t^{-1} = P_{t-1}^{-1} + C_t C_t^T \tag{17}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + P_t C_t \epsilon_t \tag{18}$$

5. Determine u_t such that: $C_{t+k}^T\hat{\theta}_t = 0$.

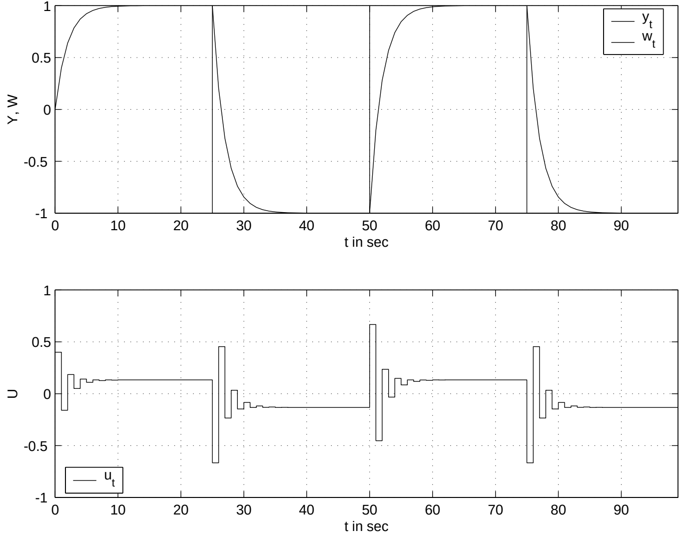
```
%-----------------------------------------------------------------------
[A,B,k,C,d,s2]=sysinit(dets); % Determine linear model (ie. get system)
%-----------------------------------------------------------------------
Am=[1 -0.6];
Bm=sum(Am);
[Ax,Bx,Cx,Dx]=armax2ss(1,Bm,k,Am);
nx=length(Ax); Xm=zeros(nx,1);
%-----------------------------------------------------------------------
[Q,R,S,G]=dsnpz(A,B,k,C,Am,Bm);
nr=length([R Q S d]); fi=zeros(nr,1);
th=[R Q S G(1)*d]'; th0=th; % th=th*0;
pil=1+[0 length(R) length([R Q])];
p0=10000;
P=eye(nr)*p0;

th(pil(2))=Q(1);  % First coefficient in Q=C*Bm is known
P(pil(2),pil(2))=0;
%th(pil(1))=R(1); % Fixed b0
%P(pil(1),pil(1))=0;
%-----------------------------------------------------------------------
```

Implicit PZ-regulator

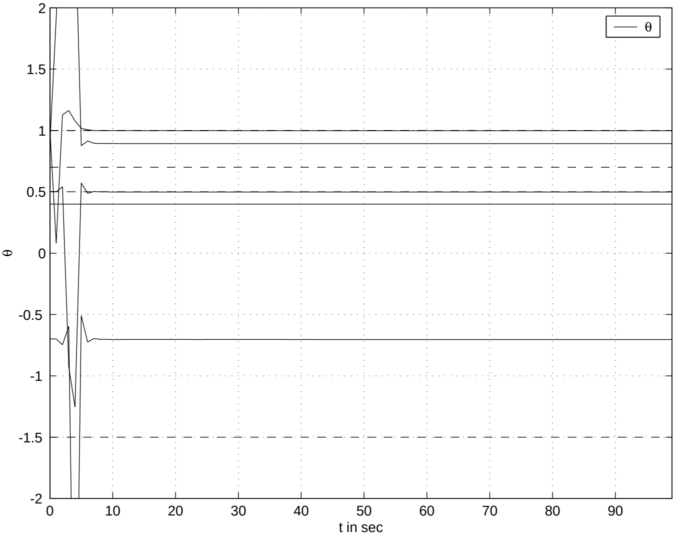
```
measinit;           % Initilialise the measurement system
for it=1:nstp,
 w=wt(it);
 [y,t]=meas;                            % Measure output
 xi=Cx*Xm+Dx*[-w;y];

 % ID block
 res=xi-fi'*th;
 K=P*fi/(1+fi'*P*fi);
 P=P-K*fi'*P;
 th=th+K*res; % th'

 fi(2:end)=fi(1:end-1);
 fi(pil)=[0 -w y]';
 u=-fi'*th/th(1);
 fi(pil(1))=u;

 act(u);  % Actuate control
 Xm=Ax*Xm+Bx*[-w;y];
end
```

Implicit PZ-regulator

Deterministic:



Implicit PZ-regulator



Implicit GMV-regulator

$$J = \mathbf{E}\left\{ \left(\frac{B_y}{A_y} y_{t+k} - \frac{B_w}{A_w} w_t \right)^2 + \rho \left(\frac{B_u}{A_u} u_t \right)^2 \right\}$$

$$A^m C = B_y AG + q^{-k} S$$

$$R = A_u BG + \alpha B_u C \qquad \delta = Gd$$

$$\begin{aligned} \xi_{t+k} &= \tilde{y}_{t+k} - \tilde{w}_t + \alpha \tilde{u}_t \\ &= \frac{1}{C}\left[S\check{y}_t + R\check{u}_t - Q\check{w}_t + \delta \right] + Ge_{t+k} \\ &= \frac{1}{C} C_{t+k}^{\top} \theta + \bar{e}_{t+k} \end{aligned}$$

$$\check{y} = \frac{1}{A_y} y \qquad \check{w} = \tilde{w} = \frac{B_w}{A_w} w \qquad \check{u} = \frac{1}{A_u} u$$

Implicit GMV-regulator

Model:

$$\xi_t = \tilde{y}_t - \tilde{w}_{t-k} + \alpha \tilde{u}_{t-k}$$
$$= C_t^\top \hat{\theta}_{t-1} + \varepsilon_t$$

Control:

$$u_t = argSol\left(C_{t+k}^\top \hat{\theta}_t = 0. \right)$$

Implicit GMV-regulator

$$J = \mathbf{E}\Big\{ \Big(\frac{B_y}{A_y} y_{t+k} - \frac{B_w}{A_w} w_t \Big)^2 + \rho \Big(\frac{B_u}{A_u} u_t \Big)^2 \Big\}$$

1. Measure y_t.
2. Create $\xi_t = \tilde{\boldsymbol{y}}_t - \tilde{\boldsymbol{w}}_{t-k} + \alpha \tilde{\boldsymbol{u}}_{t-k} \quad \left[\alpha = \frac{\rho}{b_0}\right]$.
3. Create $C_t = (\breve{y}_{t-k}, ..., \breve{u}_{t-k}, ..., -\breve{w}_{t-k}, ..., 1)^T$ og C_{t+k}.
4. Update the estimates:

$$\epsilon_t = \xi_t - C_t \hat{\theta}_{t-1} \tag{19}$$

$$P_t^{-1} = P_{t-1}^{-1} + C_t C_t^T \tag{20}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + P_t C_t \epsilon_t \tag{21}$$

5. Determine $\breve{u}_t$ such that: $C_{t+k}^T \hat{\theta}_t = 0$.
6. Determine $u_t = A_u \breve{u}_t$.

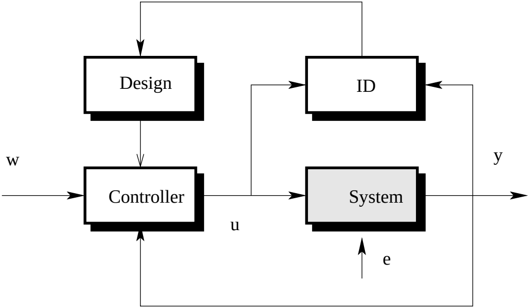
Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

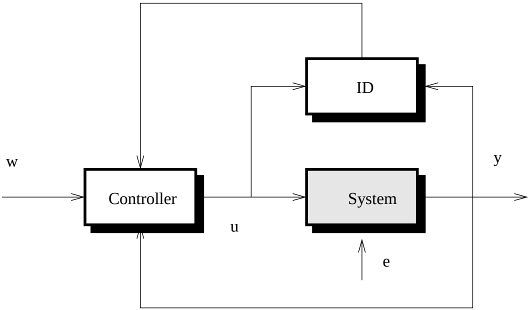
Adaptive Control (L24)

CE adaptive control

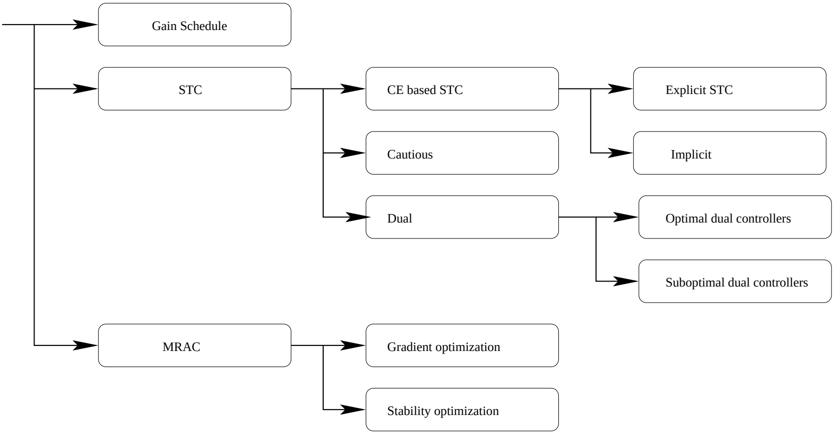


- **CE explicitte adaptive control:** Use estimate instead of known parameters

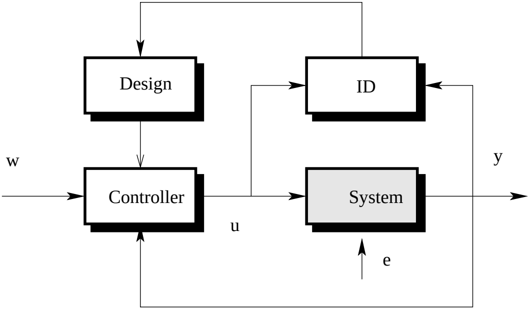
CE adaptive control



- **CE implicitte adaptive control:** If possible use the predictive nature of the controller to rephrase the (prediction) model in the control parameters and use these as if the were correct.


Gain Schedule
STC
CE based STC
Explicit STC
Cautious
Implicit
Dual
Optimal dual controllers
Suboptimal dual controllers
MRAC
Gradient optimization
Stability optimization

Dual Adaptive Control



Contradiction in goals:

- Control objective: small signals
- Estimation: large signals.

Plan

1 Known system
2 CE control
3 Cautious Control
4 Optimal Dual Control
5 Suboptimal Dual Control

Known system

Known system

System:

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

Objective, Cost function:

$$J = \mathbf{E}\left\{ \sum_{i=1}^{N} y_{t+i}^2 \right\} \qquad N = 1, \ \ldots$$

Admissible strategies:

$$u_t = func\Big(Y_t, \theta\Big)$$

Eventually $k = 1$ and $C = 1$.

Known System

N=1, Minimal variance Control

$$J = \mathbf{E}\left\{ y_{t+k}^2 \right\}$$

$$\min_{u_t(Y_t)} \mathbf{E}\left\{ y_{t+k}^2 \right\} = \mathbf{E}\left\{ \min_{u_t(Y_t)} \mathbf{E}\left\{ y_{t+k}^2 | Y_t \right\} \right\}$$

$$C = AG + q^{-k} S \qquad y_{t+k} = \frac{1}{C}\left[BGu_t + Sy_t \right] + Ge_{t+k}$$

$$u_t = -\frac{S}{BG} y_t$$

Closed loop:

$$y_t = G(q^{-1})e_t = \left(1 + g_1 q^{-1} + \ \dots g_{k-1} q^{1-k}\right) e_t$$

$$y_t = e_t \qquad \text{for } k = 1$$

Example

System:

$$y_t - y_{t-1} = bu_{t-1} + e_t$$

$$C = AG + q^{-k}S \qquad 1 = (1 - q^{-1})1 + q^{-1}s_0$$

$$u_t = -\frac{S}{BG}y_t \qquad u_t = -\frac{1}{b}y_t$$

$$y_t = e_t \qquad J^* = \sigma^2$$

CE adaptive control

ARX and unit delay ($k = 1$)

$$A(q^{-1})y_t = B(q^{-1})u_{t-1} + e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$$y_t + a_1 y_{t-1} + \ \dots\ + a_n y_{t-n}$$

$$= b_0 u_{t-1} + b_1 u_{t-1} + \ \dots\ + b_n u_{t-n} + e_t$$

$$y_t = C_t^T \theta + e_t$$

$$\theta = \left[\ a_1, \quad a_2, \dots \quad b_0, \quad b_1, \dots \right]^\top$$

$$C_t = \left[-y_{t-1}, -y_{t-2}, \dots u_{t-1}, u_{t-2}, \dots \right]^\top$$

ID: RLS (and not RML o.a.)

Minimal Variance Control

$$J = \mathbf{E}\left\{ y_{t+1}^2 \right\}$$

$$y_{t+1} = C_{t+1}^T \theta + e_{t+1} \qquad = b_0 u_t + \phi_{t+1}^T \vartheta + e_{t+1}$$

$$u_t : C_{t+1}^T \theta = 0 \qquad u_t = -\frac{\phi_{t+1}^T \vartheta}{b_0}$$

$$\theta = \left[\ a_1, \quad a_2, \ \ldots \ , a_n, \quad b_0, \quad b_1, \ \ldots \right]^\top \qquad \vartheta = \left[\ a_1, \quad a_2, \ \ldots \ , a_n, \quad b_1, \ \ldots \right]^\top$$

$$C_{t+1} = \left[-y_t, \ -y_{t-1}, \ \ldots \ y_{t-n}, \ u_t, \ u_{t-1}, \ \ldots \right]^\top \qquad \phi_{t+1} = \left[-y_t, \ -y_{t-1}, \ \ldots \ y_{t-n}, \ u_{t-1}, \ \ldots \right.$$

Connection to general formulation:

$$1 = A * 1 + q^{-1} S$$

$$S = q\left[1 - A(q^{-1})\right] \qquad S = -a_1 - a_2 q^{-1} - \ .. \ - a_n q^{-n}$$

$$B u_t = -S y_t \qquad S y_t + B u_t = 0$$

$$-a_1 y_t - a_2 y_{t-1} - \ \ldots \ + b_0 u_t + b_1 u_{t-1} + \ldots \ = 0$$

$$C_{t+1}^T \theta = 0$$

CE control

Certainty equivalence principle. Estimate the parameter in the model:

$$y_t = C_t^T \theta + e_t$$

$$\varepsilon_t = y_t - C_t^T \hat{\theta}_{t-1}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t \varepsilon_t$$

$$K_t = \frac{P_{t-1} C_t}{1 + C_t^T P_{t-1} C_t}$$

$$P_t = P_{t-1} - K \left[1 + C_t^T P_{t-1} C_t \right] K_t^\top$$

or:

$$K_t = P_t C_t$$
$$P_t^{-1} = P_{t-1}^{-1} + C_t C_t^\top$$

and use the control (as if the estimate are correct):

$$u_t : C_{t+1}^\top \hat{\theta}_t = 0 \qquad u_t = -\frac{\phi_{t+1}^\top \hat{\vartheta}_t}{\hat{b}_0}$$

Have to ensure $\hat{b}_0 \neq 0$.

Example

The system:

$$y_t - y_{t-1} = bu_{t-1} + e_t$$

is controlled by:

$$u_t = -\frac{1}{\hat{b}_t} y_t$$

where

$$\varepsilon_t = y_t - y_{t-1} - \hat{b}_{t-1} u_{t-1} \qquad C_t = u_{t-1}$$

$$\hat{b}_t = \hat{b}_{t-1} + K_t \varepsilon_t$$

$$K_t = P_t u_{t-1} \qquad P_t = \left(\sum_{i=1}^{t} u_{i-1}^2 \right)^{-1}$$

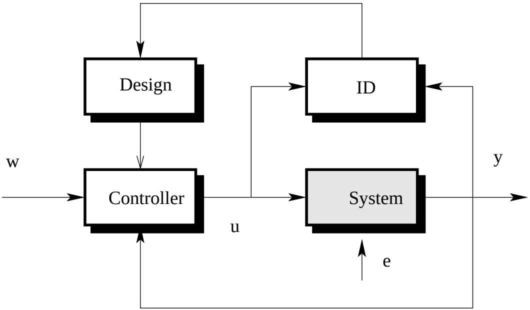
$$y_{t+1} = y_t + bu_t + e_{t+1} = y_t - \frac{b}{\hat{b}} y_t + e_{t+1} = \frac{\hat{b} - b}{\hat{b}} y_t + e_{t+1}$$

$$\mathbf{E}\left\{ y_{t+1}^2 | Y_t \right\} = \frac{P_b}{\hat{b}_t^2} y_t^2 + \sigma^2 \qquad \text{cmp.} \qquad \sigma^2 \qquad \left(P_b = P_t \sigma^2 \right)$$

Cautious adaptive control

Cautious Control

A strategy in which the uncertainty of θ play a role.



$$u_t = func(Y_t,\ \hat{\theta}_t,\ P_t)$$

Example

$$y_t - y_{t-1} = bu_{t-1} + e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$$\min_{u_t(Y_t)} \mathbf{E}\left\{y_{t+1}^2\right\} = \mathbf{E}\left\{\min_{u_t(Y_t)} \mathbf{E}\left\{y_{t+1}^2 | Y_t\right\}\right\} = \mathbf{E}\left\{\min_{u_t(Y_t)} V\right\}$$

Estimation gives:

$$b|Y_t \in \mathbf{N}\left(\hat{b}_t, P_b(t)\right)$$

$$\begin{aligned} V &= \mathbf{E}\left\{y_{t+1}^2 | Y_t\right\} \\ &= \mathbf{E}\left\{(y_t + bu_t + e_{t+1})^2 | Y_t\right\} && \text{insert } y_{t+1} \\ &= (y_t + \hat{b}_t u_t)^2 + P_b(t) u_t^2 + \sigma^2 && \text{use: } \mathbf{E}\left\{x^2\right\} = m^2 + \sigma^2 \end{aligned}$$

$$u_t = -\frac{1}{\hat{b}_t} \frac{\hat{b}_t^2}{\hat{b}_t^2 + P_b(t)} y_t$$

$$V^* = \frac{P_b}{\hat{b}_t^2 + P_b} y_t^2 + \sigma^2$$

Example: Cautius adaptive control

$$u_t = -\frac{\hat{b}_t^2}{\hat{b}_t^2 + P_b(t)} \frac{1}{\hat{b}_t} y_t \qquad \text{cmp.} \qquad u_t = -\frac{1}{\hat{b}_t} y_t \qquad \text{cmp.} \qquad u_t = -\frac{1}{b_0} y_t$$

$$V^* = \frac{P_b}{\hat{b}_t^2 + P_b} y_t^2 + \sigma^2 \qquad \text{cmp.} \qquad \frac{P_b}{\hat{b}_t^2} y_t^2 + \sigma^2 \qquad \text{compare with} \qquad \sigma^2$$

- If $\hat{b} \to b_0$ then $Ca \neq Ce = K$ (unless $P \to 0$).
- If $P_b \to 0$ then $Ca = Ce \neq K$ (unless $\hat{b} \to b_0$).
- Both approached a situation with known parameter if $\hat{b} \to b_0$ and $P_b \to 0$.

- Notice for Ca: $u_t \to 0$ for $P_b(t) \to \infty$.

Cautious Control - ARX, $k = 1$

One step control ($N = 1$). System (ARX (C=1) with $k = 1$):

$$y_t = C_t^\top \theta + e_t \quad = \quad \phi_t^T \vartheta + b_0 u_{t-1} + e_t$$

Cost:

$$J = \mathbf{E}\left\{ y_{t+1}^2 \right\}$$

Estimation results in

$$\theta | Y_t \in \mathbf{N}\left(\hat{\theta}_t, P(t)\right) \qquad \theta = \begin{bmatrix} \vartheta \\ b \end{bmatrix} \qquad C_t = \begin{bmatrix} \phi \\ u_{t-1} \end{bmatrix} \qquad P = \begin{bmatrix} P_c & P_x \\ P_x^\top & P_b \end{bmatrix}$$

$$\begin{aligned}
\mathbf{E}\left\{ y_{t+1}^2 | Y_t \right\} &= \mathbf{E}\left\{ (C_{t+1}^\top \theta + e_{t+1})^2 | Y_t \right\} \\
&= (C_{t+1}^\top \hat{\theta}_t)^2 + C_{t+1}^\top P(t) C_{t+1} + \sigma^2 \\
&= (\phi_{t+1}^\top \hat{\vartheta}_t + \hat{b}_0 u_t)^2 + \phi_{t+1}^\top P_c \phi_{t+1} + 2\phi_{t+1}^\top P_x u_t + P_b u_t^2 + \sigma^2
\end{aligned}$$

$$u_t = -\frac{\phi_{t+1}^\top \hat{\vartheta}_t}{\hat{b}_0} \frac{\hat{b}_0^2 + \phi_{t+1}^T P_x \hat{b}_0}{\hat{b}_0^2 + P_b} \quad \rightarrow \quad -\frac{\phi_{t+1}^\top \hat{\vartheta}}{\hat{b}_0} \text{ (Ce) for } P_b \rightarrow 0$$

$$V^* = \sigma^2 + (\phi^\top \hat{\vartheta})^2 + \phi^\top P_c \phi^\top - \frac{(\phi^\top [\hat{b}_0 \hat{\vartheta}_t + P_x])^2}{\hat{b}_0^2 + P_b}$$

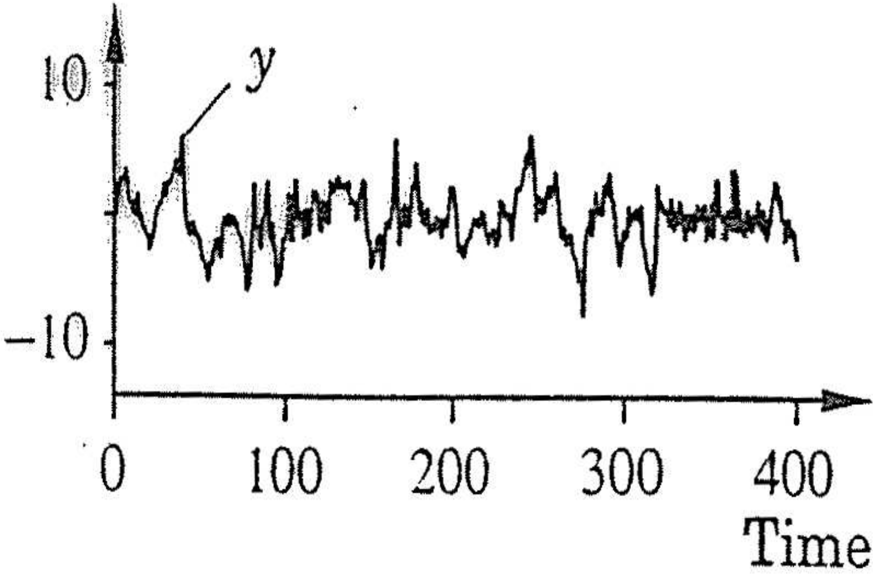
Notice: P depend on previous signals. Both y_t and u_t.

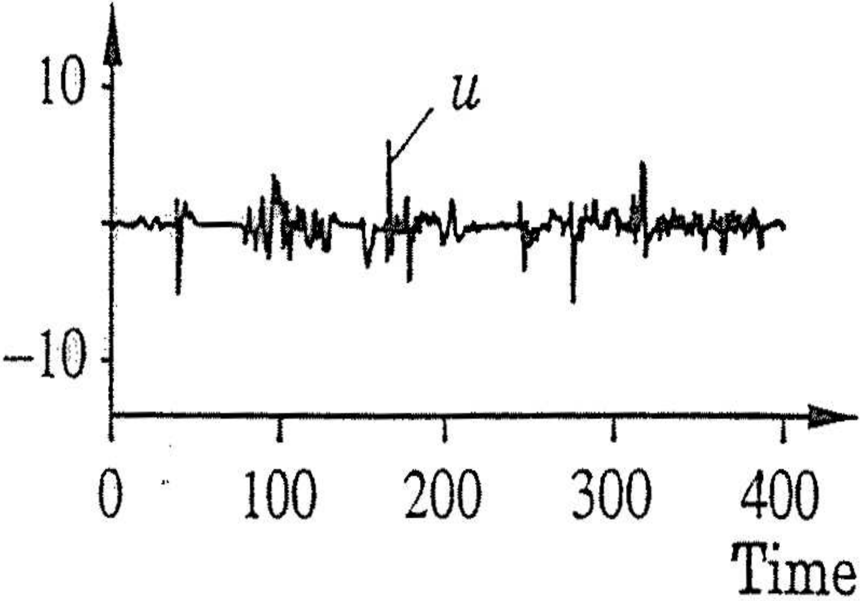
Cautious control

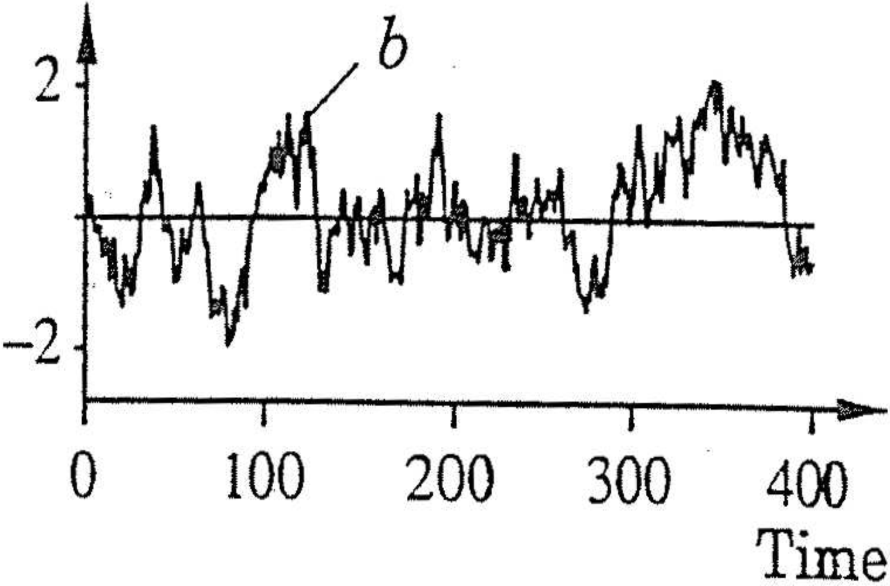
$$y_t - y_{t-1} = b_t u_{t-1} + e_t$$

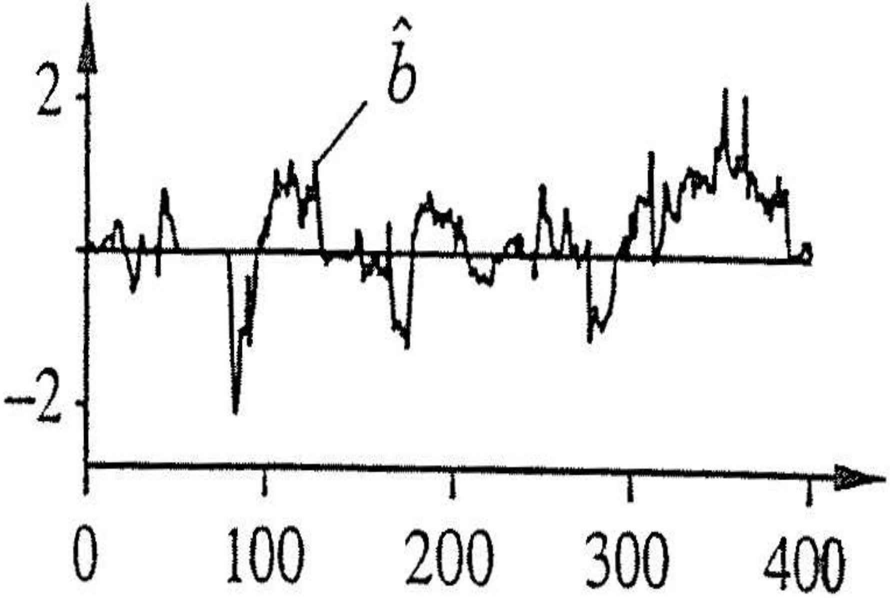
NB. Turn-off occurs when the the control signal is small.

Cautious control

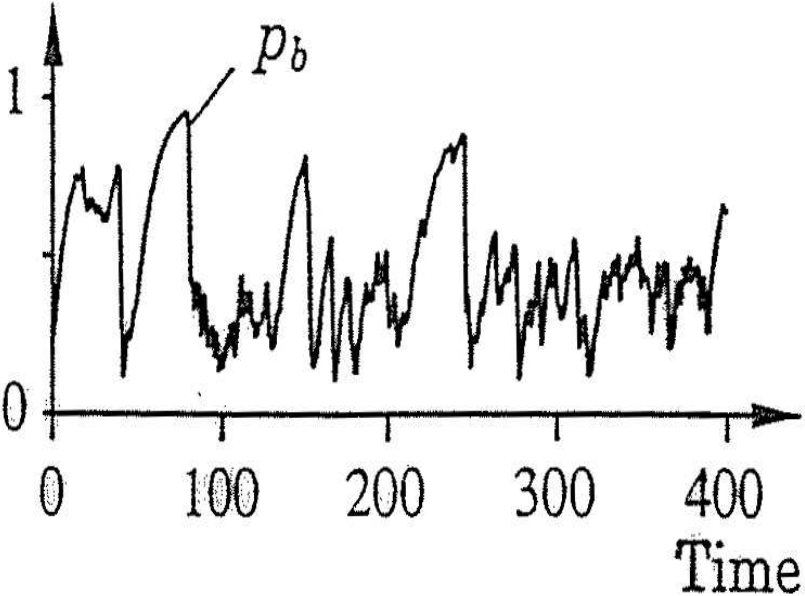



10
−10
u
0
100
200
300
400
Time

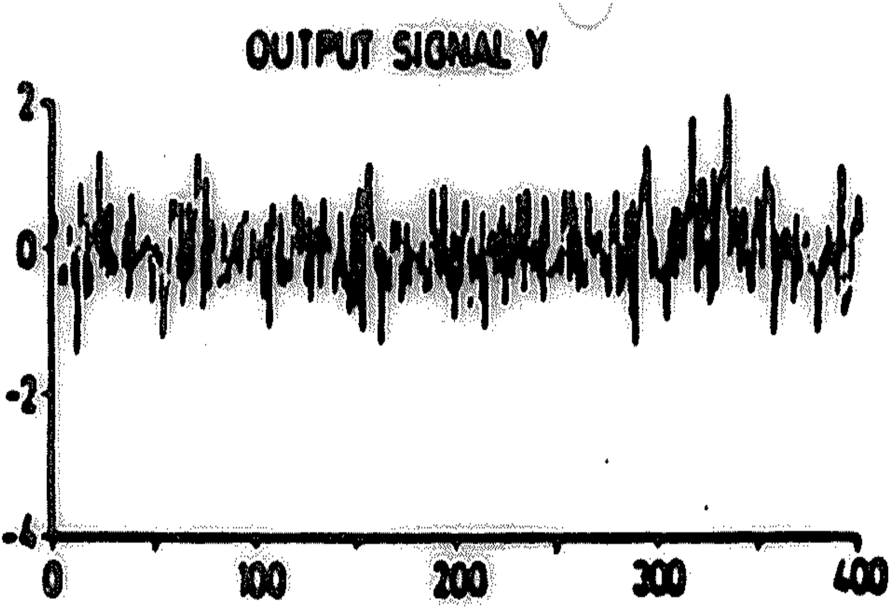

b
2
−2
0
100
200
300
400
Time

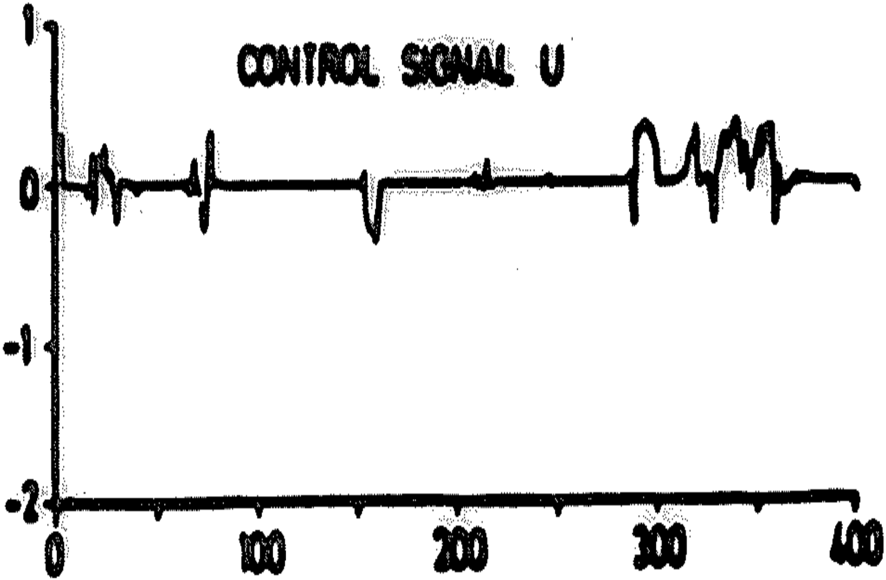


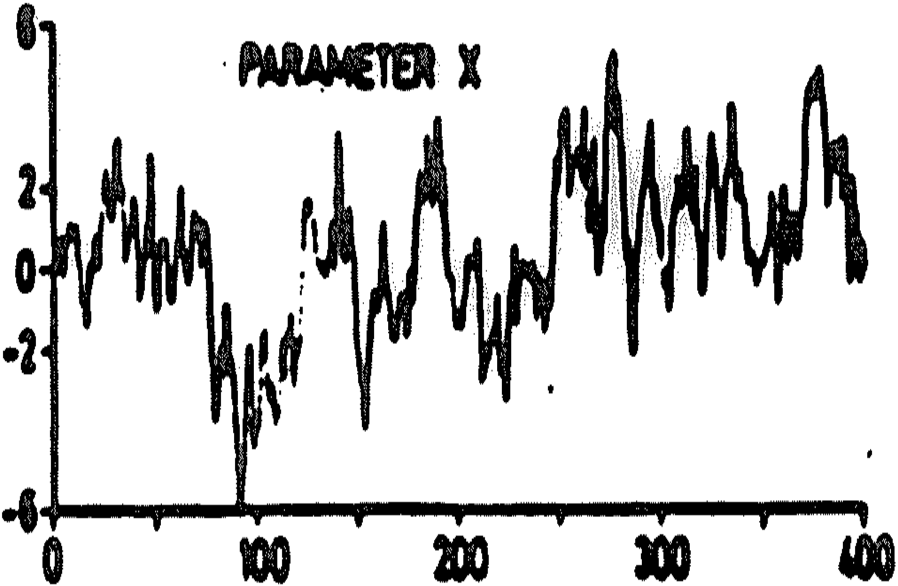
$\hat{b}$
2
−2
0
100
200
300
400

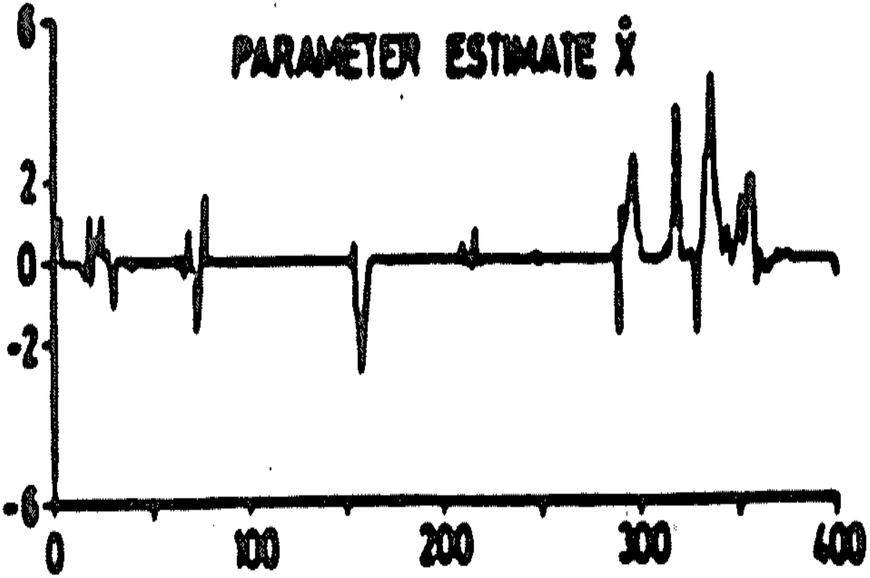


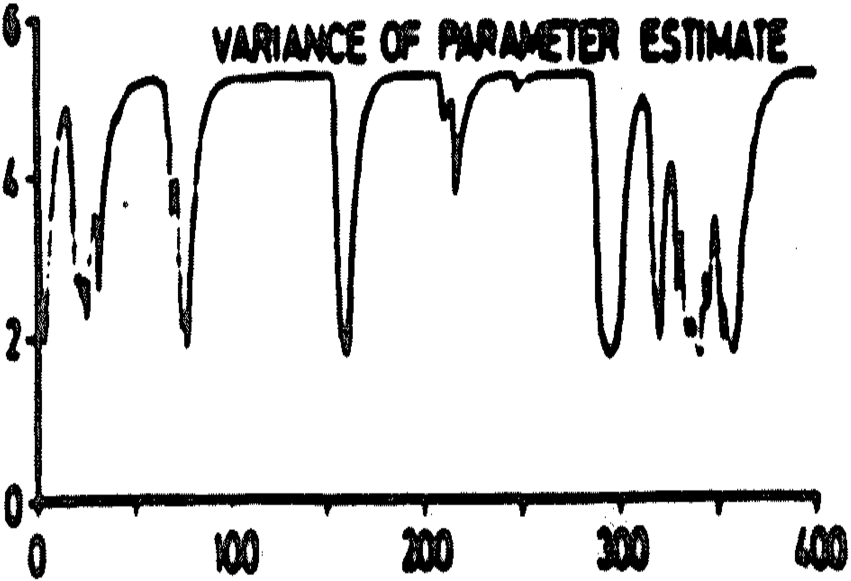
p_b
1
0
0
100
200
300
400
Time


OUTPUT SIGNAL Y
2
0
-2
-4
0
100
200
300
400


CONTROL SIGNAL U
1
0
-1
-2
0
100
200
300
400


PARAMETER X
6
2
0
-2
-6
0
100
200
300
400


PARAMETER ESTIMATE $\hat{x}$
6
2
0
-2
-6
0
100
200
300
400


VARIANCE OF PARAMETER ESTIMATE
0
2
0
100
200
300
400

Optimal Dual adaptive control

Optimal Dual Control, $N > 1$

Objective, Cost function:

$$J = \mathbf{E}\Big\{ \sum_{i=t}^{N} y_i^2 \Big\}$$

Bellman

$$x_{t+1} = f(x_t, u_t, v_t)$$

$$J_t = \min_{\{u_i\}_t^N} \mathbf{E}\Big\{ \sum_{i=t}^{t+N} I_i(x_i, u_i) \Big\}$$

$$u_t = func(Y_t)$$

$$V(Y_t) = \min_{\{u_i\}_t^N} \mathbf{E}\Big\{ \sum_{i=t}^{N} I_i \Big| Y_t \Big\}$$

$$V(Y_t) = \min_{u_t} \mathbf{E}\Big\{ I_t + V(Y_{t+1}) \mid Y_t \Big\}$$

System (output):

$$y_t = C_t^\top \theta_t + e_t$$

System (process):

$$\theta_{t+1} = \theta_t + v_t \qquad v_t \in \mathbf{N}_{iid}(0, R_1)$$

$$C_{t+1} = \begin{bmatrix} -\theta_t^\top \\ S \end{bmatrix} C_t + \begin{bmatrix} \mathbf{0} \\ 1 \\ \mathbf{0} \end{bmatrix} u_t + \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} e_t$$

Estimation (Kalman filter) gives $\theta | Y_t \in \mathbf{N}\left(\hat{\theta}_t, P_t\right)$

Hyper states: $x_t = \left(C_t, \hat{\theta}_t, P_t\right)$.

At the end point ($t + N$) the situation is as the Cautious controller and $V_N(C_{t+N}, \hat{\theta}_N, P_{t+N})$ is known, but the functional recursion (Bellman equation) is hard to solve:

$$V_t(x_t) = \min_{u_t} \mathbf{E}\left\{y_{t+1}^2 + V_{t+1}(x_{t+1}) | Y_t\right\}$$

Resort to numerical methods which involve:

1. Discretization of V in the hyper states
2. Using a quadrature formula for determining
$$\int V_t(C_t, \hat{\theta}_t, P_t) f(x_t) dx_t$$
3. Minimization over u_t for combination of the discretized hyper space.

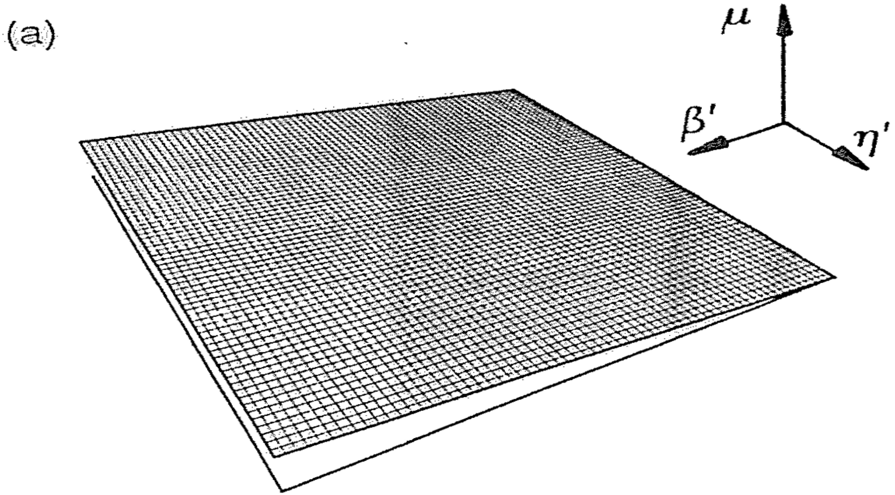
Dual control

$$\eta = \frac{y_t}{\sqrt{R_2}} \qquad \beta = \frac{\hat{b}}{\sqrt{P}} \qquad \mu = -\frac{\hat{b}}{y_t} u_t$$

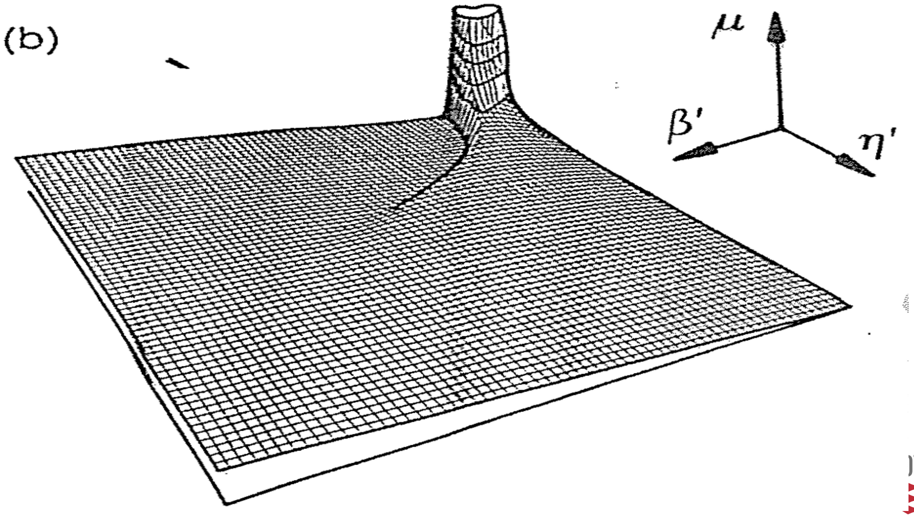
$$\eta' = \frac{\eta}{1+\eta} \qquad \beta' = \frac{\beta^2}{1+\beta^2}$$

Figures from Äström, Wittenmark (1995):Adaptive control

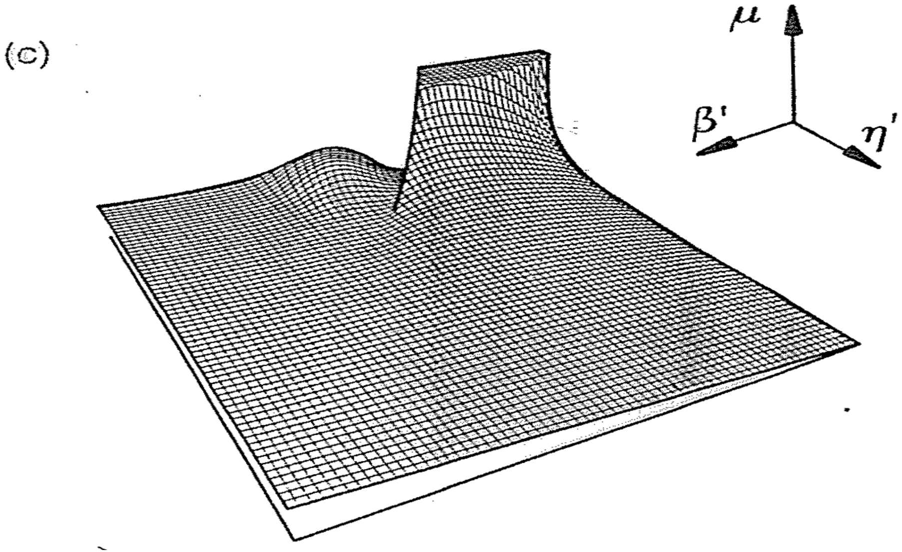
N=1 (Cautious control)



N=3

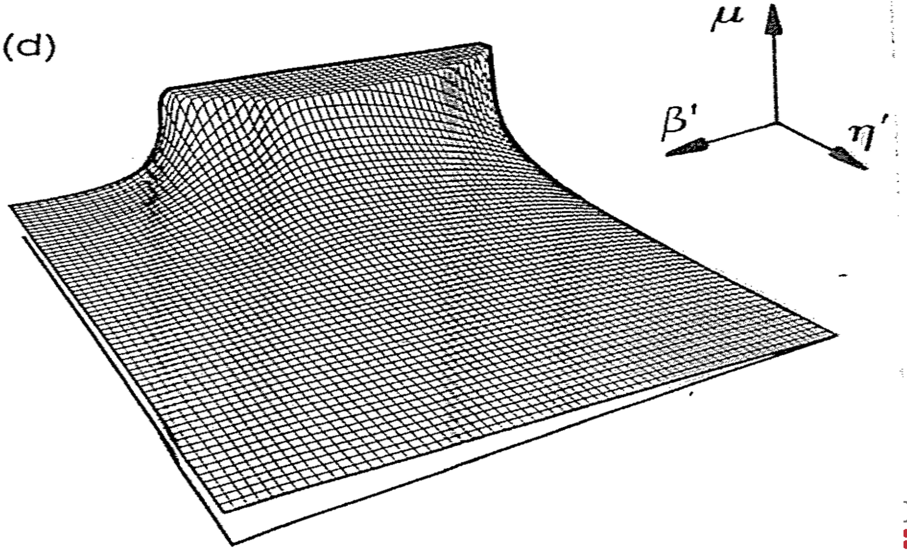


N=6

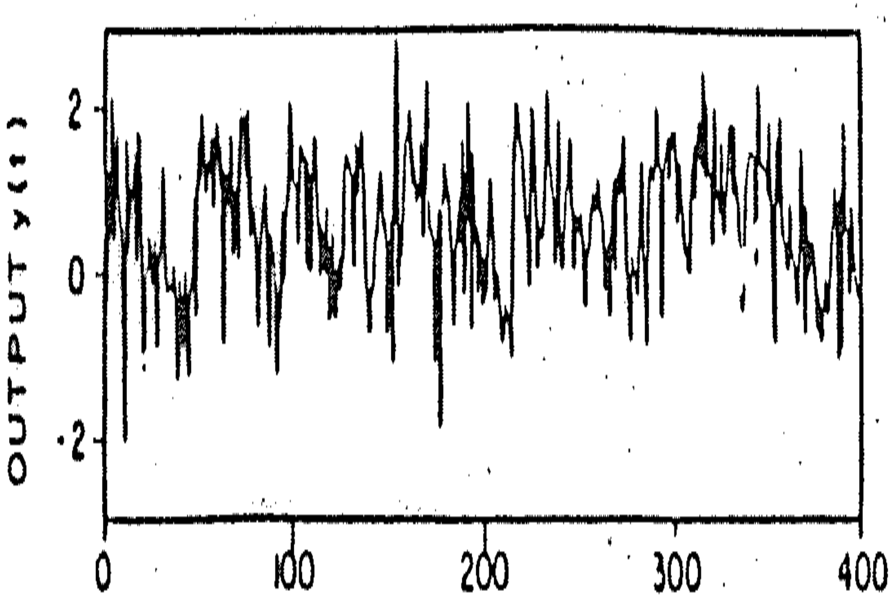


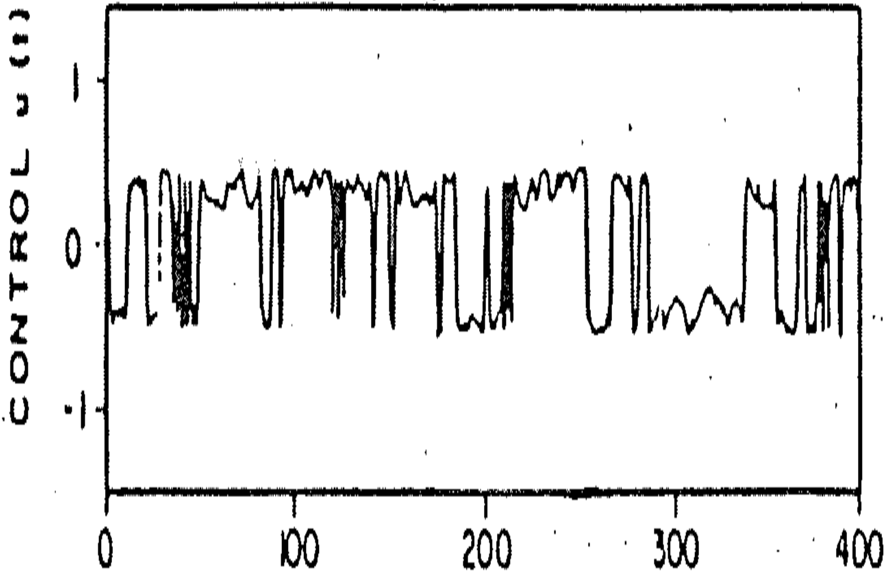
Dual control

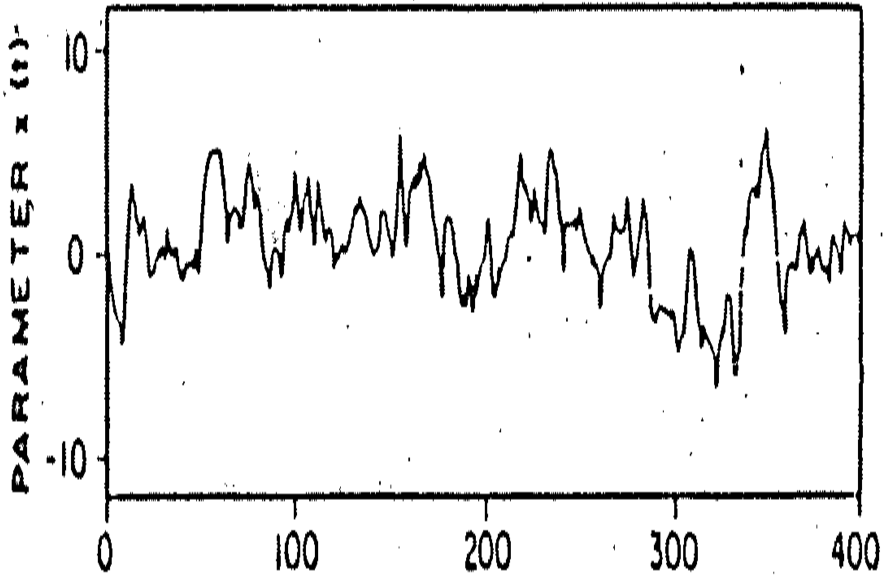
N=31

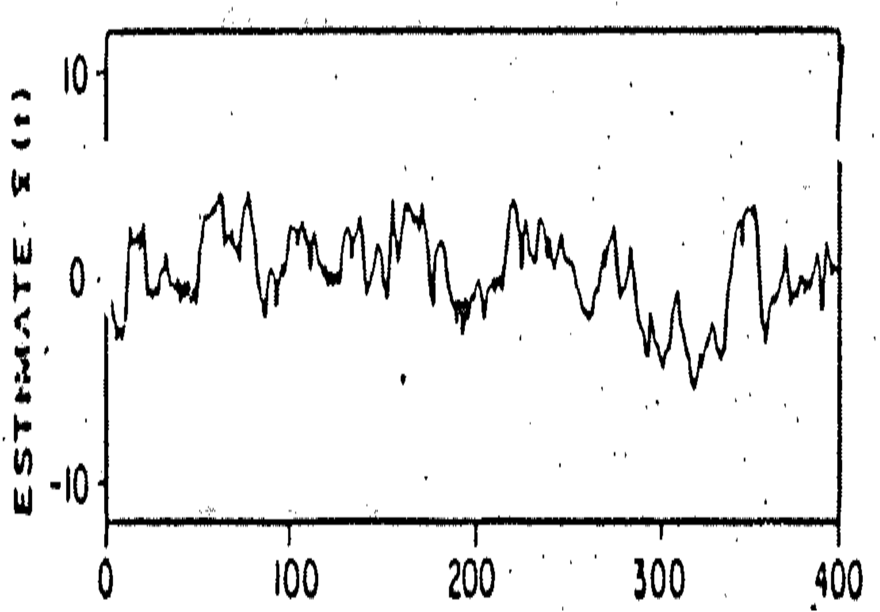


Dual control

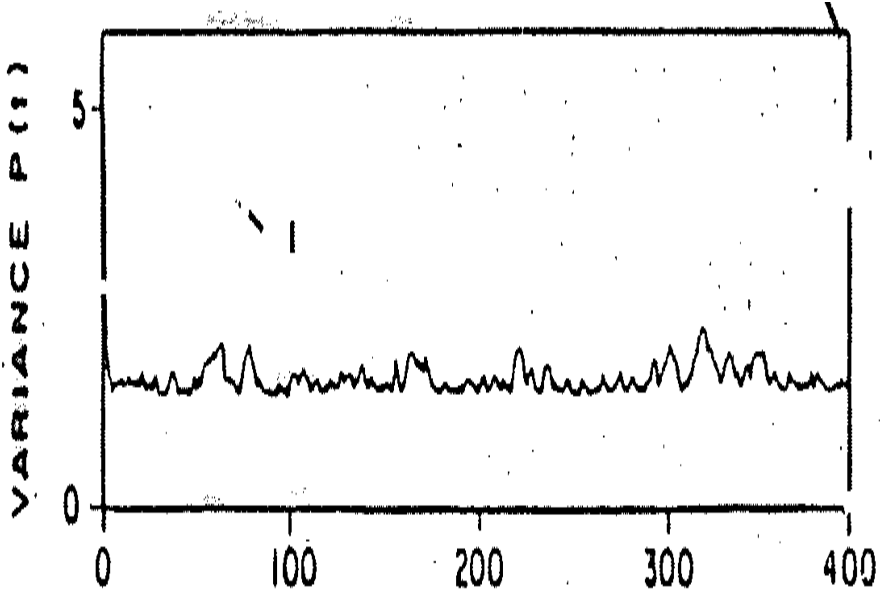



CONTROL u (t)
1
0
-1
0
100
200
300
400


PARAMETER x (t)
10
0
-10
0
100
200
300
400


ESTIMATE
10
0
-10
0
100
200
300
400

Dual control



Suboptimal Dual adaptive control

Constrained One-Step Minimizations

Limit the minimum value of the control action:

$$u_t = \begin{cases} u_{cautious} & \text{if } |u_{cautious}| > |u_{lim}| \\ u_{lim}\ sign(u_{cautious}) & \text{if } |u_{cautious}| < |u_{lim}| \end{cases}$$

Constrained uncertainty

Limit the variance of uncertainty.

The one-step cost is minimized subject to:

$$tr(P_{t+1}^{-1}) > M$$

or

$$P_b(t+1) < \begin{cases} \gamma \hat{b}_0^2 & \text{for } P_b(t) \leq \hat{b}_t^2 \\ \alpha P_b(t) & \text{otherwise} \end{cases}$$

Extension of the Loss function

Basicly want to "minimize" P_{t+1} as well

$$J = \mathbf{E}\Big\{ y_{t+1}^2 + \rho f(P_{t+1}) \Big\}$$

$$f(P_{t+1}) = P_b(t+1) \qquad f(P_{t+1}) = \frac{P_b(t+1)}{P_b(t)}$$

$$f(P_{t+1}) = -\varepsilon_{t+1}^2 \qquad f(P_{t+1}) = -\frac{det(P_t)}{det(P_{t+1})}$$

Active adaptive control

$$y_t = C_t^T \theta + e_t$$

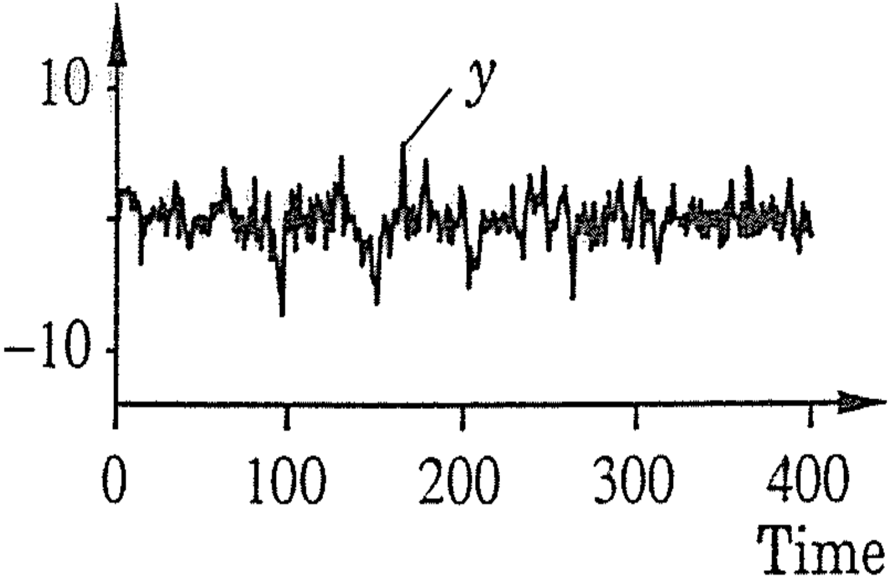
$$J_1 = \mathbf{E}\left\{ y_{t+1}^2 - \rho \frac{det(P_t)}{det(P_{t+1})} \right\} \qquad \rightarrow \qquad J_2 = \mathbf{E}\left\{ \, y_{t+1}^2 \middle| \, \theta = \hat{\theta} \right\} - \rho \frac{det(P_t)}{det(P_{t+1})}$$

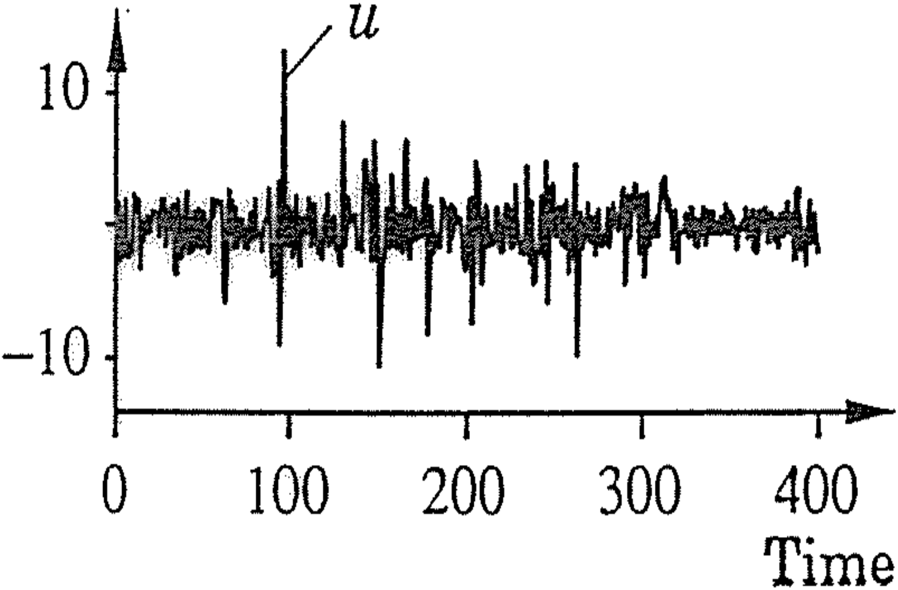
$$\mathbf{E}\left\{ y_{t+1}^2 | \theta, Y_t \right\} = (C_{t+1}^T \theta)^2 + \sigma^2 = (\phi_{t+1}^T \vartheta + b_0 u_t)^2 + \sigma^2$$

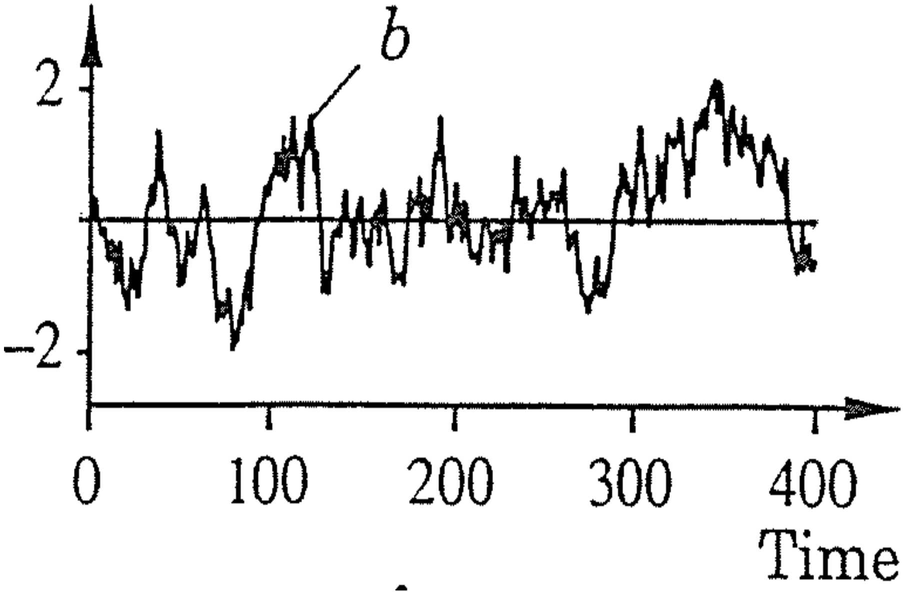
$$\frac{det(P_t)}{det(P_{t+1})} = 1 + C_{t+1}^T P_t C_{t+1} = 1 + \phi^T P_c \phi + 2\phi^T P_x u_t + P_b u_t^2$$

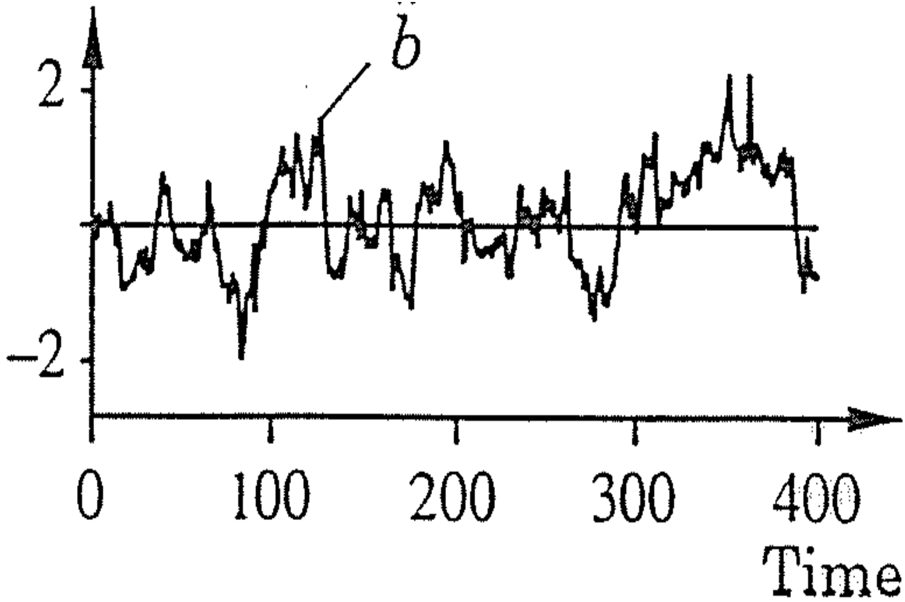
$$u_t = -\frac{\hat{b}_0 \phi_{t+1}^T \hat{\vartheta}_t - \rho P_x \phi_{t+1}}{\hat{b}_0^2 - \rho P_b}$$

- $\rho = 0$ gives a Ce controller
- $\rho > 1$ gives an active learning controller
- $\rho = -1$ gives a cautious controller

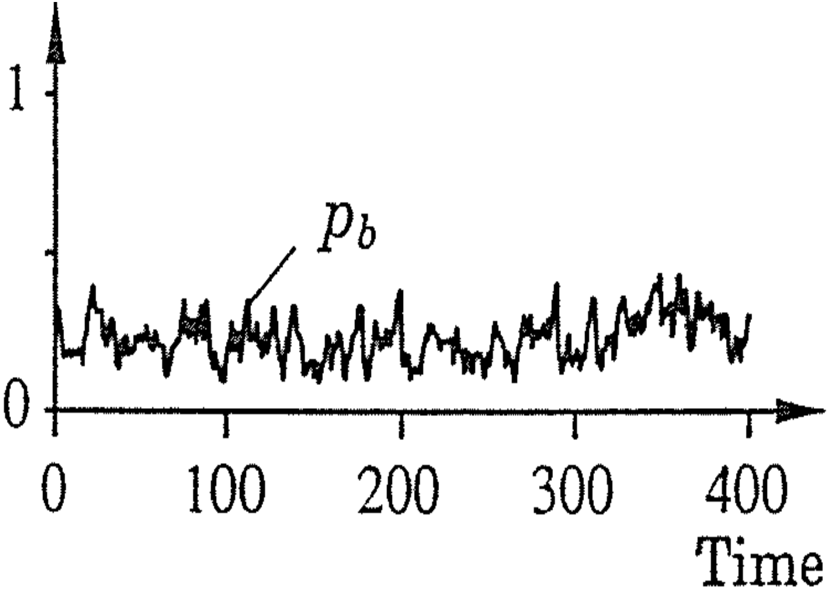

10
−10
y
0
100
200
300
400
Time


u
10
−10
0
100
200
300
400
Time


b
2
−2
0
100
200
300
400
Time



$\hat{b}$
2
−2
0
100
200
300
400
Time


1
0
p_b
0
100
200
300
400
Time

Probing

$$u_t = u_t^c + u_t^x$$

- PRBS
- Design of excitation signal

Persistenly exitation: i -> t roderi med m og n

Stochastic Adaptive Control (02421)

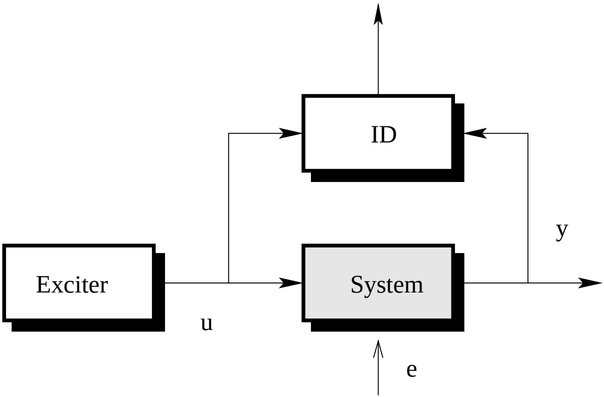
www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

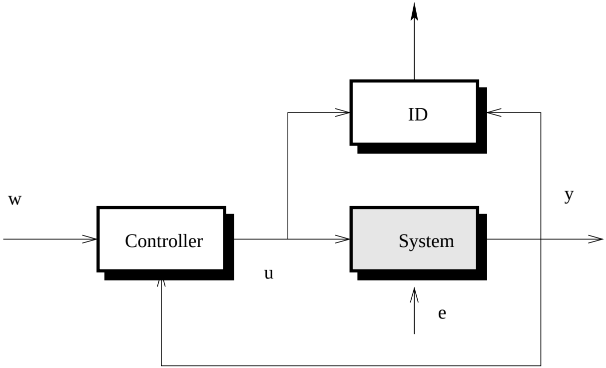
Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Closed Loop Identification (L25)



The input (u) do not depend on y (and e).



Why ?

- unstable plant
- in a production, economic and safety reasons
- inherent feedback mechanics
- adaptive control
- point of operation same as the intended use (⋆)
- input spectrum same as in intended use (⋆)

Closed loop identification

Avoid closed loop identification if possible (at all cost)

1987

Use closed loop identification whenever possible

Pitch fall - Example I

Plant:

$$y_t + a y_{t-1} = b u_{t-1} + e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

Control (we are only consider the regulation problem):

$$u_t = -f y_t$$

Spectral analysis

$$\hat{G}(e^{j\omega}) = \frac{\phi_{yu}(\omega)}{\phi_u(\omega)}$$

Let us analyse the convergence point.

$$\Psi_u(z) = H_u(z)H_u(z^{-1})\sigma^2$$

Here H_x is the transfer function from noise to x

$$\Psi_{yu}(z) = H_y(z)H_u(z^{-1})\sigma^2$$

$$\hat{G}(e^{j\omega}) = \frac{\phi_{yu}(\omega)}{\phi_u(\omega)} = \frac{H_y(z)}{H_u(z)}$$

$z = e^{j\omega}$

Example I

In closed loop we have:

$$y_t + (a + bf)y_{t-1} = e_t \qquad \text{and} \qquad u_t = -fy_t$$

or

$$H_y(z) = \frac{1}{1 + (a + bf)z^{-1}} \qquad H_u(z) = \frac{-f}{1 + (a + bf)z^{-1}}$$

Then:

$$\hat{G}(z) = \frac{H_y(z)}{H_u(z)} = \frac{1}{1 + (a + bf)z^{-1}} \times \frac{1 + (a + bf)z^{-1}}{-f} = -\frac{1}{f}$$

We might have to take care when doing ID in closed loop.

Here the problem is (lack of) ensuring causality.

Pitch fall - Example II

Prediction error method

The plant:

$$y_t + a y_{t-1} = b u_{t-1} + e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

have the prediction error

$$\varepsilon_t = y_t + \hat{a} y_{t-1} - \hat{b} u_{t-1} \qquad \hat{y}_{t|t-1} = -\hat{a} y_{t-1} + \hat{b} u_{t-1}$$

but in closed loop ($u_t = -f y_t$):

$$\varepsilon_t = y_t + (\hat{a} + \hat{b} f) y_{t-1} \qquad \hat{y}_{t|t-1} = -(\hat{a} + \hat{b} f) y_{t-1}$$

That means that any

$$\hat{a} = a_0 + \gamma f \qquad \hat{b} = b_0 - \gamma$$

(where γ is arbitrary scalar) gives equally good predictions.

The controller is too simple (ie. not complex enough).

Closed loop ID

If the plant

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

is controlled by

$$u_t = -\frac{S(q^{-1})}{R(q^{-1})}\ y_t$$

The closed loop is

$$\left[AR + q^{-k}BS\right]y_t = RCe_t$$

The number of equations outnumbers the number of parameters (in A and B) if

$$Max\left[n_r - n_b, n_s + k - n_a\right] \geq 1 + n_p$$

where n_p is the number of common factors in C and $AR + q^{-k}BS$.

The controller has to be adequately complex

DTU

Proof: Number of parameters:

$$n_a + n_b + 1$$

Order of denominator (equals number of equations):

$$Max(n_a + n_r, k + n_b + n_s) - n_p$$

Match:

$$Max\Big[n_r - n_b, n_s + k - n_a\Big] = 1 + n_p$$

Number of common factors (between denominator and numerator), n_p .

The Minimal Variance Controller

$$Ay_t = q^{-k}Bu_t + Ce_t \qquad J = \mathbf{E}\left\{y_{t+k}^2\right\}$$

$$C = AG + q^{-k}S \qquad y_{t+k} = \frac{1}{C}\left[BGu_t + Sy_t\right] + Ge_{t+k}$$

$$y_t = q^{-k}\frac{B}{A}u_t + \frac{C}{A}e_t = -q^{-k}\frac{B}{A}\frac{S}{BG}y_t + \frac{C}{A}e_t$$

$$\left(AG + q^{-k}S\right)y_t = CGe_t$$

$$n_r = n_b + k - 1 \qquad n_p = n_c \qquad n_s = n_a - 1$$

$$Max\left[n_r - n_b, n_s + k - n_a\right] \geq 1 + n_p$$

$$k \geq n_c + 2$$

Example II (again)

Plant:

$$y_t + ay_{t-1} = bu_{t-1} + e_t$$

Control:

$$u_t = -fy_t$$

Here:

$$Max\left[n_r - n_b, n_s + k - n_a\right] = 1 + n_p$$

$$= Max\left[0 - 0, 0 + 1 - 1\right] = 0 < 1 + 0 = 1$$

DTU

Informative data

Sufficient informative: A set of data, $\{z_t\}$, is *sufficient informative with respect to a model set* $\mathcal{M}$ if, for two models, $\mathcal{G}_1$ and $\mathcal{G}_2$, in the set

$$\bar{\mathbf{E}}\left\{\|(\mathcal{G}_1(q) - \mathcal{G}_2(q))z_t\|^2\right\} = 0$$

implies that $\mathcal{G}_1(e^{j\omega}) \equiv \mathcal{G}_2(e^{j\omega})$ for almost all ω.

Persistenly exciting

$$r_k = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbf{E}\left\{u_t u_{t-k}\right\}$$

$$R_m = \begin{bmatrix} r_0 & r_1 & \cdots & r_{m-1} \\ r_1 & r_0 & & \\ \vdots & \vdots & & \vdots \\ r_{m-1} & r_{m-2} & \cdots & r_0 \end{bmatrix}$$

$u_t \;\; pe(n)$ if $\;R_n > 0$

The bad news

- The closed loop experiment may be non-information even if the input in itself is persistently exciting. Reason: the controller might be too simple.

- Spectral analysis applied in the straightforward fashion will give erroneous results. The estimate of G will converge to

$$G_* = \frac{G_0\phi_r - F\phi_v}{\phi_r + |F|^2\phi_v} \qquad\qquad y = Gu + v$$

- Correlation methods will give biased estimate of the impulse response.

- OE gives unbiased estimate of G in open loop experiments, even if the additive noise (v) is not white. This is not true in close loop.

- The subspace methods will typically not give consistent estimate when applied to close loop data.

DTU

The good news

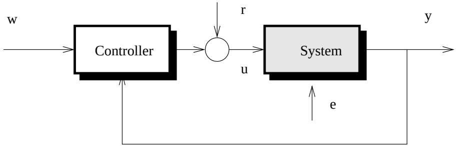
PE methods will give consistent estimate of the system if

- The data is informative.
- The model set contains the true system ($\mathcal{S} \in \mathcal{M}$).

The closed loop experiment is informative if the reference w_t (or another probe signal) is persistingly exciting.

Patches

- Persistingly exciting reference or probe signal



- Time-varying (adaptive) or nonlinear controller
- Shift between m different LTI controllers

$$u_t = -F_i(q)y_t \qquad\qquad i = 1, \; \ldots \; m$$

where

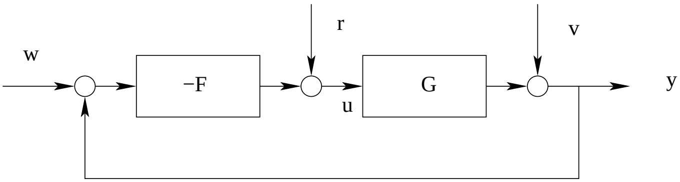
$$m \geq 1 + \frac{n_u}{n_y}$$

Closed loop identification methods

Approaches to closed loop identification

- Direct approach
- Indirect approach
- Joint Input- output Approach

We assume in the following that the reference (or a probe signal) is persistingly exciting.


w
−F
r
u
G
v
y

Direct Approach

The system is identified in exactly the same way as in open loop identification.

- It works regardless of the complexity of the controller and requires no knowledge about the character of the feedback.
- No special algorithms or software are required. (A word from our sponsers) .
- Consistency and optimal accuracy are obtained if model structure contains the true system.
- Unstable system can be handled without problems (as long as the closed loop and the predictor are stable).

Drawback: we need good noise models.

(Not a problem if true system (G,H) is contained in model structure).

If noise model is incorrect (fixed incorrectly or not contain the true noise model) bias of G will be introduced.

Indirect approach

Assume the plant is given by

$$y_t = Gu_t + v_t = Gu_t + He_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

and the control is (using partially the notation in LL):

$$u_t = w_t - Fy_t$$

Then the closed loop is given by

$$y_t = GSw_t + Sv_t$$

$$u_t = Sw_t - FSv_t$$

where the sensitivity functions is:

$$S = \frac{1}{1 + FG}$$

Indirect approach

In the indirect method the closed transfer functions are estimated and from these the plant parameters (or transfer functions) are determined.

$$y_t = G_{cl} w_t + H_{cl} e_t$$

where

$$G_{cl} = \frac{G}{1+FG} \qquad H_{cl} = \frac{H}{1+FG}$$

Consequently:

$$\hat{G} = \frac{\hat{G}_{cl}}{1-\hat{G}_{cl}F} \qquad \hat{H} = \hat{H}_{cl}(1+F\hat{G})$$

+ Any (open loop) method such as spectral analysis, instrumental variable, subspace and prediction error methods can be applied.
- Any error in F will transported directly to the estimate of the model. (Notice saturation a.o.)

Prediction error methods

The parameterization might be directly in the system parameters ie.

$$G_{cl} = \frac{G(\theta)}{1 + FG(\theta)} \qquad H_{cl} = \frac{H(\theta)}{1 + FG(\theta)}$$

and

$$\varepsilon_t = H_{cl}^{-1}\Big(y_t - G_{cl}u_t\Big)$$

We then have the mapping

$$\theta \to \Big[G_{cl}, H_{cl}\Big] \to \varepsilon_t \to J = \sum \frac{1}{2}\varepsilon_t^2$$

Other parameterizations exists. Methods based on the Youla-Kucera parameterization has been proposed by Hansen, Franklin and Kosut (1989) and Schrama (1991).

Let the system be

$$y_t = Gu_t + v_t \qquad\qquad v_t = He_t$$

If the controller is

$$u_t = w_t - Fy_t + z_t$$

where z_t ($z_t \perp w_t, v_t$) is an unknown signal, then the closed loop is characterized by

$$\begin{aligned} y_t &= GSw_t + Sv_t + GSz_t &= G_{cl}w_t + v_1 \\ u_t &= Sw_t - FSv_t + Sz_t &= G_{ur}w_t + v_2 \end{aligned}$$

where

$$S = \frac{1}{1 + FG}$$

Let the model

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = S \begin{bmatrix} G \\ 1 \end{bmatrix} w_t + S \begin{bmatrix} H & G \\ -FH & 1 \end{bmatrix} \begin{bmatrix} e_t \\ z_t \end{bmatrix}$$

or

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \mathcal{G} w_t + \mathcal{H} \begin{bmatrix} e_t \\ z_t \end{bmatrix} \qquad S = \frac{1}{1+FG}$$

be used.

Then a ML or a PEM method is basicly a minimization of

$$J = \sum_{i=1}^{t} \varepsilon_i^\top R^{-1} \varepsilon_i \qquad\qquad R = \mathbf{V}ar\left\{ \begin{bmatrix} e_t \\ z_t \end{bmatrix} \right\}$$

where

$$\varepsilon_t = \mathcal{H}^{-1} \left[\begin{bmatrix} y_t \\ u_t \end{bmatrix} - \mathcal{G} w_t \right]$$

where $G(\theta)$, $H(\theta)$ and $F(\theta)$.

This parameterization might be independent, i.e. $G(\theta)$, $H(\gamma)$ and $F(\eta)$.

Approach 1

Consider a special case,

$$R = \begin{bmatrix} \sigma_e^2 & 0 \\ 0 & \sigma_z^2 \end{bmatrix}$$

then

$$\begin{aligned} J &= \frac{1}{\sigma_e^2} \sum_{i=1}^{t} (H^{-1}(y_i - G u_i))^2 \\ &+ \frac{1}{\sigma_z^2} \sum_{i=1}^{t} (u_i - w_i + F y_i)^2 \end{aligned}$$

If the parameterization of (G, H) and F is independent, this is a direct method for estimation G, H and F.

Approach 2

Here the corelation is disregarded and the two equations are treated as separate and as

$$\begin{bmatrix} y_t \\ u_t \end{bmatrix} = \begin{bmatrix} G_{cl} \\ G_{uw} \end{bmatrix} w_t + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \qquad J = \frac{1}{\sigma_1^2}\sum_{i=1}^{t}(y_i - G_{cl}w_i)^2 + \frac{1}{\sigma_2^2}\sum_{i=1}^{t}(u_i - G_{uw}w_i)^2$$

This approach has many variants, which have in common that

$$\hat{G} = \frac{\hat{G}_{cl}}{\hat{G}_{uw}} \qquad G_{cl} = GS \qquad G_{uw} = S$$

If nothing is done to prevent it, $\hat{G}$ would be of unnecessarily high order. (Lack of perfect cancellation)

One way to prevents this is to use an independent parameterization of G and S which results in

$$G_{cl} = G(\vartheta)S(\eta) \qquad G_{ur} = S(\eta)$$

Approach 2 - A two stage approach

If additionally the two noise sources are (assumed to be) independent with a variance ratio β then the cost to be minimized is

$$J(\vartheta, \eta) = \beta \sum_{i=0}^{t} (y_i - G(\vartheta)S(\eta)w_t)^2 + \sum_{i=0}^{t} (u_i - S(\eta)w_i)^2$$

If $\beta \to 0$ then η is determined to fit the second part of the cost. A two step procedure could then consist of (Van den Hof+Schrama, Forssel+Ljung):

1. Estimate η for $\beta \to 0$.
2. Use

   $$\hat{u}_t = S(\hat{\eta})w_t$$

   and estimate ϑ by fitting :

   $$y_t = G(\vartheta)\hat{u}_t + v_1$$

One example on a parameterization of S is

$$S(\eta) = \sum_{k=-m}^{m} s_k q^{-k}$$

which is a non causal filter.

DTU

Stochastic Adaptive Control (02421)

www.imm.dtu.dk/courses/02421

Niels Kjølstad Poulsen

Build. 303B, room 016
Section for Dynamical Systems
Dept. of Applied Mathematics and Computer Science
The Technical University of Denmark

Email: nkpo@dtu.dk
phone: +45 4525 3356
mobile: +45 2890 3797

Design of Experiment (L26)

Design of Experiment

Plan

- Configuration
- Common signals
- Informative Experiments/Signals - Necessarily condition
- Optimal Experiments - Classical approach
- Optimal Experiments - Modern approach

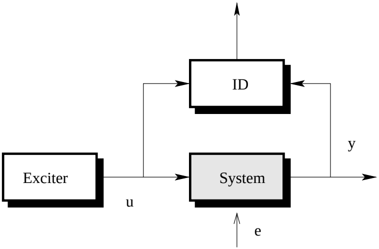
Literature

- Lennart Ljung (1999): System Identification (Ch. 13)
- Goodwin and Payne (1977): Dynamic System Identification: Experiment Design and Data Analysis.

DTU

Configurations

Configuration - Open loop



$$u_t = z_t$$

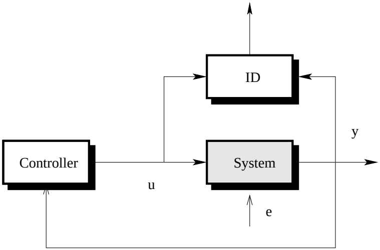
$$z_t = A\ sin(\omega\ t)$$

$$z_t = \xi_t \qquad \xi_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$$z_t = H(q)\xi_t \qquad \xi_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

$$u_t = z_t = f(t, \Omega)$$ which design parameters ?

Configuration - closed loop I

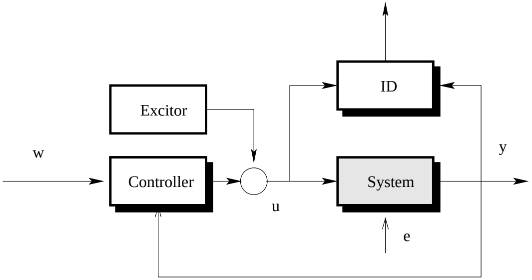


$$R(q^{-1})u_t = -S(q^{-1})y_t$$

$$u_t = -\frac{S(q^{-1})}{R(q^{-1})}y_t$$

which design parameters ?

Configuration - closed loop II



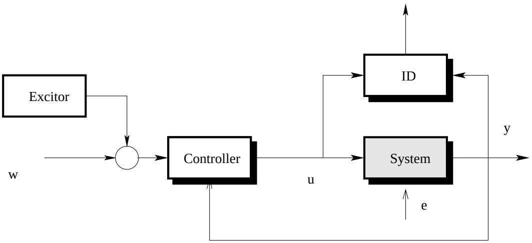
$$R(q^{-1})u_t = Q(q^{-1})w_t - S(q^{-1})y_t$$

$$u_t = \frac{Q(q^{-1})}{R(q^{-1})}w_t - \frac{S(q^{-1})}{R(q^{-1})}y_t + \underline{v_u}$$

$$u_t = f(t, \Omega, y_t)$$

which design parameters ?

Configuration - closed loop III



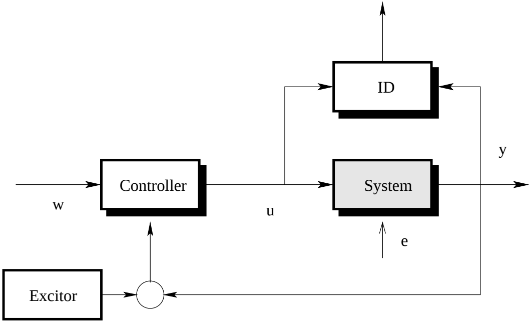
$$R(q^{-1})u_t = Q(q^{-1})w_t - S(q^{-1})y_t$$

$$u_t = \frac{Q(q^{-1})}{R(q^{-1})}\left(w_t + v_w\right) - \frac{S(q^{-1})}{R(q^{-1})}y_t$$

$$u_t = f(t, \Omega, y_t)$$

which design parameters ?

Configuration - closed loop IV



$$R(q^{-1})u_t = Q(q^{-1})w_t - S(q^{-1})y_t$$

$$u_t = \frac{Q(q^{-1})}{R(q^{-1})}w_t - \frac{S(q^{-1})}{R(q^{-1})}\left(y_t + v_y\right)$$

$u_t = f(t, \Omega, y_t)$ which design parameters ?

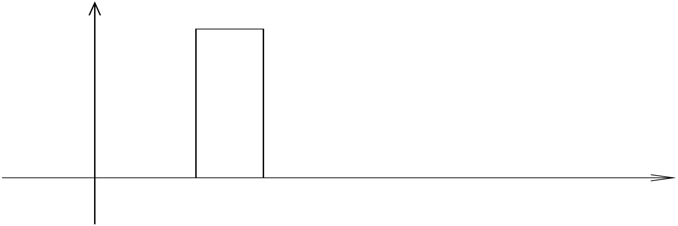
Common Excitation Signals

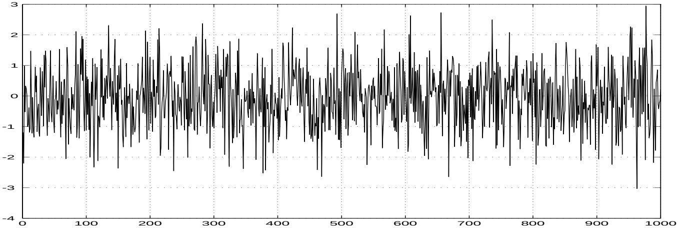
Common Excitation Signals

Considerations:

- Degree of excitation (separate models, improve precision)
- Generation (realization)
- Portability
- Reproduceability

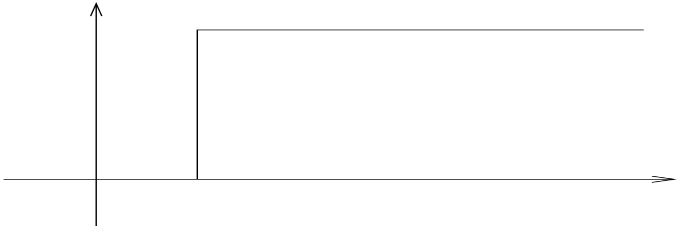
Impulse and white noise (D and C)

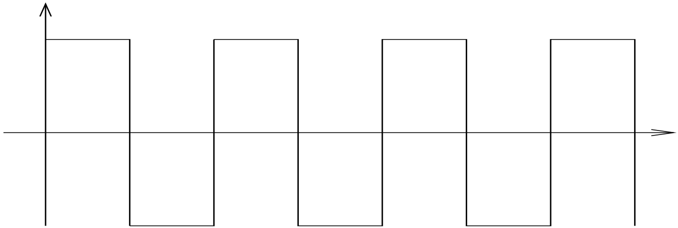




Realize the signals

Step and Square wave





For a constant signal:

$$\frac{d}{dt}u_t = 0 \qquad\qquad (1 - q^{-1})u_i = 0$$

For a ramp function

$$\frac{d^2}{dt^2}u_t = 0 \qquad\qquad (1 - q^{-1})^2 u_i = 0$$

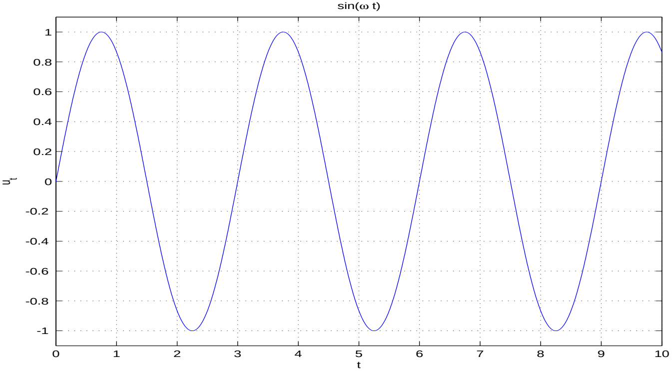
Harmonic function

$$u_t = A\ sin(\omega\ t)$$

$$u_i = A\ sin(\omega T_s i)$$

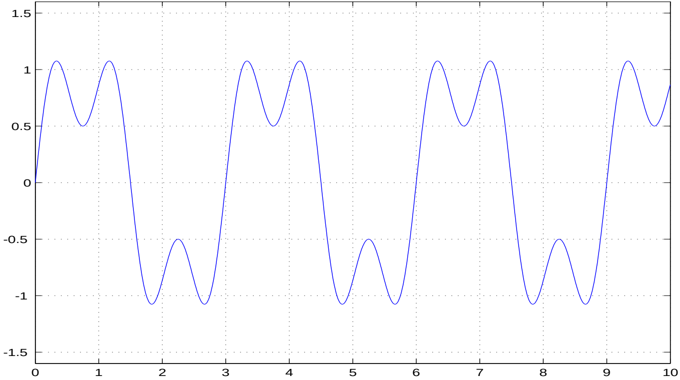
$$(s^2 - \omega^2)u_t = 0$$

$$(q^2 + 2\cos(\omega T_s)q + 1)u_i = 0$$



Sum of harmonic functions

$$u_t = \sum_{k=1}^{n} A_k \; sin(\omega_k \; t + \phi_k)$$

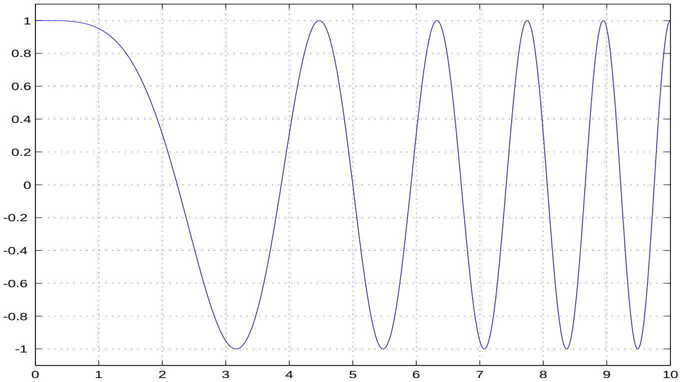


DTU

Chirp

$$u_t = A \sin(\omega_t \; t) \qquad \omega_t = \omega_0 + \alpha \; t$$

for example



JRJ's 7 tone signal

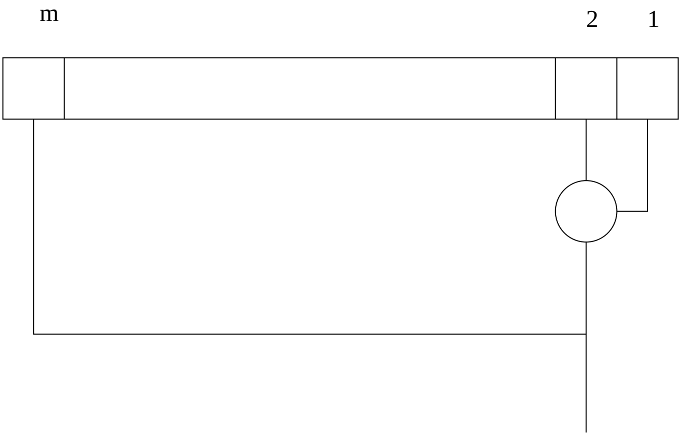
```
function [u,t]=jrj7sign(T,h,per)
% Usage: [u,t]=jrj7sign(T,h,per)
%
% T: length of signal
% h: sampling period
% per: period of signal

z=[1 per 0;
   1 per/2 0
   1 per/4 0
   1 per/8 0
   1 per/16 0
   1 per/32 0
   1 per/64 0];
[u,t]=harmsin(T,h,z);
```

Jens R. Jensen: Founder of Servolaboratoriet (now: Automation and Control, DTU-EE), Cofounder of IFAC.

JRJ's 7 tone signal

```
function [u,t]=harmsin(T,h,z)
% Usage: u=harmsin(T,h,z)
%
% T: length of signal
% h: sample period
% z: matrix (ns*3) each row [ Amplitude period phase], and one row per
%    component

[ns,n2]=size(z);
t=0:h:T; t=t(:);
n=length(t);
u=zeros(n,1);
for i=1:ns,
 u=u+z(i,1)*sinwave(T,h,z(i,2),z(i,3));
end
```

PRBS (Pseudo Random Binary Signal)

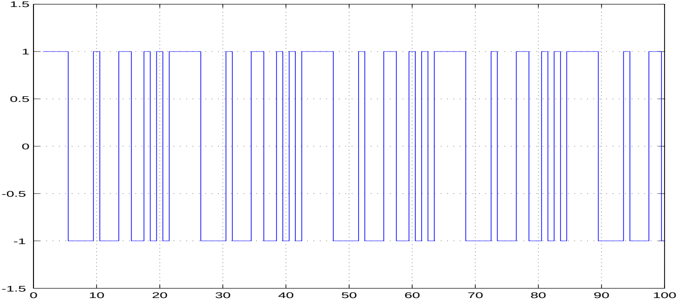


Here only 1 and 0.

PRBS

```
function u=prbs(n,m)
%function u=prbs(n,m)
% n: length of sequence
% m: length of shift register

% Programmed 1994 by Niels K. Poulsen
% Department of Mathematical Modelling,
% Technical University of Denmark

u=zeros(n,1);
x=(osc(m)+1)/2;
x=fix(x);

xor=[0 1; 1 0];

for i=1:n,
 y=xor(x(1)+1,x(2)+1);
 x(1:m-1)=x(2:m);
 x(m)=y;
 u(i)=y;
 end
 u=2*(u-0.5); % u in [-1,1]
```

PRBS



$m = 5$ (short shift registre), notice the short period.

A more modern (LL chp. 13) implementation:

$$z_t = rem(B(q)z_{t-1}, 2) \qquad \in \{0, 1\} \qquad\qquad B(q^{-1}) = b_1 + b_2 q^{-1} + \dots b_m q^{1-m}$$

Maximum length is $M = 2^m - 1$, which is obtained for special cases of B.

The coefficients of B are 1 or 0.

m	$M = 2^m - 1$	$b_k = 1$ for k
2	3	1,2
3	7	2,3
8	255	1,2,7,8
11	2047	9,11
18	262143	11,18

For a maximum length PRBS shifting between $\pm\bar{z}$;

$$|\mu| = \frac{\bar{z}}{M} \qquad\qquad R_k = \begin{cases} \bar{z}^2 & \text{for } k = 0, \pm M, \pm 2M, \dots \\ -\frac{\bar{z}^2}{M} & \text{otherwise} \end{cases}$$

is persistently exiting of order $M - 1$.

DTU

```
IDINPUT Generates input signals for identification.
   U = IDINPUT(N,TYPE,BAND,LEVELS)

   U: The generated input signal. A column vector or a N-by-nu matrix.
   N: The length of the input.
   N = [N Nu] gives a N-by-Nu input (Nu input channels).
   N = [P Nu M] gives a M*P-by-Nu input, periodic with period P
       and with M periods.
   Default values are Nu = 1 and M =1 ;
   TYPE: One of the following:
         'RGS': Generates a Random, Gaussian Signal.
         'RBS': Generates a Random, Binary Signal.
         'PRBS': Generates a Pseudo-random, Binary Signal.
         'SINE': Generates a sum-of-sinusoid signal.
         Default: TYPE = 'RBS'.
   BAND: A 1 by 2 row vector that defines the frequency band for the
         input's frequency contents.
         For the 'RS', 'RBS' and 'SINE' cases BAND = [LFR,HFR], where
         LFR and HFR are the lower and upper limits of the passband,
         expressed in fractions of the Nyquist frequency (thus always
         numbers between 0 and 1).
         For the 'PRBS' case BAND = [0,B], where B is such that the
         signal is constant over intervals of length 1/B (the Clock Period).
         Default: BAND =[0 1].
   LEVELS = [MI, MA]: A 2 by 1 row vector, defining the input levels.
         For 'RBS', 'PRBS', and 'SINE', the levels are adjusted so
         that the input signal always is between MI and MA.
         For the 'RGS' case, MI is the signal's mean value minus one
         standard deviation and MA is the signal's mean plus one standard
         deviation.
         Default LEVELS = [-1 1].
```

Random signals

- white (Gaussian (randn), uniform (rand) or other distributions)
- filtered noise or colored noise

  $$u_i = H(q)e_i \qquad e_i \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

  Use e.g. *filter* in Matlab.

Design of experiments

Some Basics

Ergodicity

$$J = \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} \varepsilon_t^2 \quad = \quad \mathbf{E}\left\{\varepsilon_t^2\right\}$$

Parseval

$$\varepsilon_t = H(q)e_t \qquad e_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right) \qquad\qquad \phi_\varepsilon(\omega) = H(z)H(z^{-1})\sigma^2 \geq 0 \qquad z = e^{j\omega}$$

$$\mathbf{E}\left\{\varepsilon_t^2\right\} = \int_{-\pi}^{\pi} \phi_\varepsilon(\omega) d\omega$$

DTU

Objectives

1. Distinguish between models
2. improve precision of estimate
3. improve performance on the application of the model

Informative Experiments

Loose definition: Informative experiments means we can see the difference between two models.

Difference between models means difference in predictions or prediction errors.

More precise: When can we from

$$\mathbf{E}\left\{D\varepsilon^2\right\} = 0$$

conclude

$$DG = 0, \quad DH = 0$$

That depend on

$$\mathcal{X}, \mathcal{M}$$

Informative Experiments - the stage settings

Let

$$\mathcal{S}: \quad y_i = G_0 u_i + H_0 e_i$$

$$\mathcal{M}_1: \quad y = G_1 u + H_1 \varepsilon_1$$

(time index and ˆ omitted in the following)

$$\mathcal{M}_2: \quad y = G_2 u + H_2 \varepsilon_2$$

$$D\varepsilon = \varepsilon_1 - \varepsilon_2$$

$$DG = G_1 - G_2 \qquad DH = H_1 - H_2$$

We have initially:

$$\begin{aligned}\varepsilon_1 &= \frac{1}{H_1}\Big[y - G_1 u\Big] \\ &= \frac{1}{H_1}\Big[G_0 u + H_0 e - G_1 u\Big] \\ &= \frac{1}{H_1}\Big[(G_0 - G_1)u + H_0 e\Big]\end{aligned}$$

and in a similar way:

$$\varepsilon_2 = \frac{1}{H_2}\Big[(G_0 - G_2)u + H_0 e\Big]$$

This results in

$$H_1\varepsilon_1 - H_2\varepsilon_2 = -(G_1 - G_2)u \qquad \rightarrow \qquad \varepsilon_1 = \frac{1}{H_1}\Big[-DGu + H_2\varepsilon_2\Big]$$

Furthermore:

$$D\varepsilon = \varepsilon_1 - \varepsilon_2 = \frac{1}{H_1}\Big[-DGu + H_2\varepsilon_2\Big] - \varepsilon_2 = -\frac{1}{H_1}\Big[DGu + DH\varepsilon_2\Big]$$

If we are using

$$\varepsilon_2 = \frac{1}{H_2}\Big[(G_0 - G_2)u + H_0 e\Big]$$

we get:

$$D\varepsilon = -\frac{1}{H_1}\left[DGu + \frac{DH}{H_2}\Big((G_0 - G_2)u + H_0 e\Big)\right]$$

Informative Experiments

$$D\varepsilon = -\frac{1}{H_1}\left[DGu + \frac{DH}{H_2}\Big((G_0 - G_2)u + H_0 e\Big)\right]$$

(just a copy - with some colors)

$$D\varepsilon = -\frac{1}{H_1}\left[DG + \frac{G_0 - G_2}{H_2}DH\right]u \; - \; DH\frac{H_0}{H_1 H_2}e$$

Suppose the experiment is in open loop, then:

$$\mathbf{E}\Big\{(D\varepsilon)^2\Big\} = \int_{-\pi}^{\pi} \phi_1(\omega) + \phi_2(\omega) d\omega$$

where

$$\phi_1(\omega) = \Big|\frac{1}{H_1}\Big|^2 \Big|DG + \frac{G_0 - G_2}{H_2}DH\Big|^2 A_u(\omega)$$

$$\phi_2(\omega) = \Big|\frac{DH H_0}{H_1 H_2}\Big|^2 \sigma^2$$

Informative Experiments - open loop

Now, suppose

$$\mathbf{E}\Big\{(D\varepsilon)^2\Big\} = \int_{-\pi}^{\pi} \phi_1(\omega) + \phi_2(\omega) d\omega = 0$$

Since $\phi_1 \geq 0$ and $\phi_2 \geq 0$, that means both ϕ_1 and ϕ_2 has to be zero for all frequencies.

Consider:

$$\phi_2(\omega) = \left|\frac{DHH_0}{H_1 H_2}\right|^2 \sigma^2$$

If $|H_0(e^{j\omega})| > 0, \ \forall \omega$ then

$$DH(e^{j\omega}) = 0$$

The first term:

$$\phi_1(\omega) = \left|\frac{1}{H_1}\right|^2 \left|DG + \frac{G_0 - G_2}{H_2} DH\right|^2 A_u(\omega)$$

takes then the form

$$\left|DG(e^{j\omega})\right|^2 A_u(\omega) = 0$$

The data is sufficiently informative if we from that can conclude $DG(e^{j\omega}) = 0$.

Informative Experiments - open loop (half way summary)

The experiment is informative if we from

$$\mathbf{E}\left\{(D\varepsilon)^2\right\} = 0 \qquad (1)$$

can conclude

$$DG \equiv 0 \qquad DH \equiv 0$$

The data is sufficiently informative iff we from

$$\left|DG(e^{j\omega})\right|^2 A_u(\omega) = 0 \qquad (2)$$

can conclude

$$DG(e^{j\omega}) = 0$$

The data is not sufficiently informative if (2) is satisfied even $DG(e^{j\omega}) \neq 0$.

The problem is $DG(e^{j\omega})$ can be non-zero at frequencies where $A_u(\omega) = 0$.

Informative Experiments - Persistently excitation

Definition: A quasi stationary signal, u_t with spectrum $A_u(\omega)$ is said to be persistently exciting of order n if for all filters of the form

$$M(q) = m_0 + m_1 q^{-1} + \ \dots \ + m_{n-1} q^{1-n}$$

the relation

$$\left| M(e^{j\omega}) \right|^2 A_u(\omega) = 0 \tag{3}$$

implies that

$$M(e^{j\omega}) = 0$$

for all ω - or equivalently that

$$m_i = 0 \qquad\qquad i = 0,\ 1,\ \dots,\ n-1$$

Notice, u_t is $pe(n)$, while $ord(M) = n-1$ and M has n coefficients and $n-1$ zeros. That implies, $M(z)M(z^{-1})$ can have at most $n-1$ different zeros on the unit circle.

In order to uniquely to determine the n coefficients in M, $A_u(\omega)$ has to be non-zero at (least at) n different frequencies.

If the signal u_t is persistently exciting of order n, then $A_u(\omega)$ is different from zero at least n different points in the interval $[-\pi, \pi]$.

Consider a constant signal

$$u_t = c$$

and is pe(1).

Consider a harmonic signal

$$u_t = A\ sin(\omega t)$$

which have a frequency component at $\pm\omega$ and consequently is pe(2).

Consider the signal

$$u_i = c(-1)^i \qquad \omega T_s = \pi \qquad \text{(Nyquist frequency)}$$

which is obtained for $\omega T_s = \pi$ and is pe(1).

Consider the signal

$$u_t = \sum_{k=1}^{n} A_k sin(\omega_k t)$$

Each sinusoid give rise to a spectral line at $-\omega_k$ and ω_k. This signal is $pe(2n)$ (if $\omega_k \neq 0$ and $\omega_k T_s \neq \pi$).

If one $\omega_k = 0$ then it is $pe(2n-1)$.

If one $\omega_k T_s = \pi$ then it is $pe(2n-1)$.

If one $\omega_k T_s = \pi$ and one $\omega_k = 0$ then it is $pe(2n-2)$.

Consider the signal

$$v_t = M(q^{-1})u_t$$

which has the spectral density:

$$\phi_v(\omega) = \left|M(e^{j\omega})\right|^2 \phi_u(\omega)$$

A signal that is $pe(n)$ can not be filtered to zero by a (non-zero) MA filter of order $n-1$ (or less). But a (non-zero) MA filter of order n (or higher) might.

Consider $u_t = const$. This signal is $pe(1)$. The filter of order 1 $M_1(q^{-1}) = 1 - q^{-1}$ can filter the signal to zero, whereas the filter $M_0 = 1$ can't.

Informative Experiments

Consider a model set given by

$$G = q^{-k} \frac{b_0 + b_1 q^{-1} + \ \dots \ + b_{n_b} q^{-n_b}}{1 + f_1 q^{-1} + \ \dots \ f_{n_f} q^{-n_f}}$$

An open loop experiment with u_t being $pe(n_f + n_b + 1)$ is sufficiently informative with respect to this set.

Proof: For two different models we have

$$DG = \frac{B_1 F_2 - B_2 F_1}{F_1 F_2}$$

and

$$\left| B_1(e^{j\omega}) F_2(e^{j\omega}) - B_2(e^{j\omega}) F_1(e^{j\omega}) \right|^2 A_u(\omega) = 0$$

The order of DG is $n_f + n_b$

Notice, the order of persistently excitation equals the number of parameters to be estimated.

It is sufficient to use $n + 1$ sinusoids to identify an n order system (assuming numerator and denominator to be estimated have the same number of parameters).

Let u_t be a quasi stationary signal and

$$R = \begin{bmatrix} r_0 & r_1 & \cdots & r_{n-1} \\ r_1 & r_0 & \cdots & r_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n-1} & r_{n-2} & \cdots & r_0 \end{bmatrix}$$

then u_t is $pe(n)$ if (and only if) R is non singular.

Proof: It is (said that it is) easy to verify that (look at the variance of $v_t = M(q^{-1})u_t = m^T\vec{u}$):

$$\mathbf{E}\Big\{(M(q^{-1})u_t)^2\Big\} = \mathbf{E}\Big\{(m^T \begin{bmatrix} u_t \\ \vdots \\ u_{t-n+1} \end{bmatrix})^2\Big\} = m^T R m$$

and then

$$m^T R m = \int_{-\pi}^{\pi} \left|M(e^{j\omega})\right|^2 A_u(\omega)\, d\omega$$

R is nonsingular iif

$$m^T R m = 0 \Rightarrow m = 0$$

or iff

$$\left|M(e^{j\omega})\right|^2 A_u(\omega) = 0 \Rightarrow M(e^{j\omega}) = 0$$

Informative Experiments

A quasi stationary signal, u_t with a spectrum $A(\omega)$ is in open loop persistently exciting if

$$A(\omega) > 0$$

for almost all ω.

Almost all means A may be zero on a set of measure zero (countable number of points).

DTU

Corollar: An **open loop** experiment is informative if the input is persistently exciting.

▸ jump to definition

Optimal Design

Consider the ARX structure:

$$A(q^{-1})y_t = B(q^{-1})u_t + e_t$$

which can reformulated as a linear regression model

$$y_t = C_t^T \theta + e_t$$

For correct model structure ($\mathcal{S} \in \mathcal{M}$) the estimate is unbiased and asymptotic

$$\tilde{\theta} \in \mathbf{N}\left(0, P\right) \qquad P = \frac{1}{N}\mathbf{E}\Big\{C_t C_t^T\Big\}^{-1}$$

For the general (external) model

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t$$

we can rearrange the model to

$$y_t = C_t^T\theta + e_t$$

where C_t contains signals which are dependent on the system parameters. In this case

$$\tilde{\theta} \in \mathbf{N}(0, P) \qquad P = \frac{1}{N}\mathbf{E}\left\{\psi\psi^T\right\}^{-1} \qquad \psi = \frac{\partial}{\partial\theta}\hat{y} = -\frac{\partial}{\partial\theta}\varepsilon$$

In the general (LTI) case we have for correct model structure that

$$\tilde{\theta} \in \mathbf{N}(0, P)$$

where

$$P = \frac{1}{N} M^{-1} \qquad M = \mathbf{E}\left\{\psi\psi^T\right\} \qquad \psi = \frac{\partial}{\partial\theta} log(f(y_t|Y_{t-1}, \theta))$$

Optimal design I

Many method is based on a minimization of

$$J = \alpha(M) = \alpha(\mathbf{E}\left\{\psi\psi^T\right\})$$

Examples are:

$$\alpha(M) = -log[tr(M)] \qquad\qquad \alpha(M) = -log[det(M)]$$

The minimization has to be performed with respect to some constraints

$$\mathbf{E}\left\{u^2\right\} < c \qquad \mathbf{E}\left\{y^2\right\} < c$$

or some combination hereof.

Determines A_i, ω_i, ϕ_i in

$$u_t = \sum_{i=1}^{n} A_i sin(\omega_i t + \phi_i)$$

such that the performance index

$$J = -\det \bar{E}\{\psi\psi^T\} \qquad \psi = \frac{\partial \hat{y}}{\partial \theta}$$

is minimized subject to

$$\bar{E}\{u^2\} < c$$

- Trace
- Constraint on y instead - or a combination hereof.

Optimal open loop excitation

Determines the parameters in:

$$u_t = H(q)\zeta_t \qquad \zeta_t \in \mathbf{N}_{iid}\left(0, \sigma^2\right)$$

such that the performance index

$$J = -\det \bar{E}\{\psi\psi^T\} \qquad \psi = \frac{\partial \hat{y}}{\partial \theta}$$

is minimized subject to

$$\bar{E}\{u^2\} < c$$

- Trace
- Constraint on y instead - or a combination hereof.

Closed loop excitation

Determine the parameters in

$$u_t = -\frac{S(q^{-1})}{R(q^{-1})} y_t$$

such the performance index

$$J = \det \bar{E}\{\psi\psi^T\} \qquad \psi = \frac{\partial \hat{y}}{\partial \theta}$$

is maximized subject to

$$\bar{E}\{u^2\} < c$$

- Trace
- Constraint on y instead - or a combination hereof.

Optimal Design II

True system

$$y_t = G_0 u_t + H_0 e_t = T_0 \mathcal{X}$$

$$T_0 = \begin{bmatrix} G_0 & H_0 \end{bmatrix} \qquad \mathcal{X} = \begin{bmatrix} u_t \\ e_t \end{bmatrix}$$

Objective Assume the objective can be formulated as a minimization of

$$J = \int_{-\pi}^{\pi} \mathbf{E}\Big\{\|\tilde{T}\|_C^2\Big\} d\omega = \int_{-\pi}^{\pi} \mathbf{E}\Big\{\tilde{T} C(\omega) \tilde{T}^T\Big\}\, d\omega$$

where

$$\tilde{T} = \hat{T} - T_0 = \begin{bmatrix} \hat{G} - G_0 & \hat{H} - H_0 \end{bmatrix}$$

More specific

$$J = \int_{-\pi}^{\pi} \mathbf{E}\Big\{\tilde{T} C(\omega) \tilde{T}^T\Big\}\, d\omega = \int_{-\pi}^{\pi} tr\Big\{\Pi(\omega) C(\omega)\Big\} d\omega$$

where

$$C(\omega) = \begin{bmatrix} C_{11}(\omega) & C_{12}(\omega) \\ C_{21}(\omega) & C_{22}(\omega) \end{bmatrix} \qquad \Pi(\omega) = \mathbf{E}\Big\{\tilde{T}^T(\omega) \tilde{T}(\omega)\Big\}$$

jump

Optimal Design II

Basicly, minimize the performance degradation due to modeling errors.

Simulation

$$C(\omega) = \begin{bmatrix} A_u(\omega) & 0 \\ 0 & 0 \end{bmatrix}$$

Prediction

$$C(\omega) = \frac{1}{|H_0(\omega)|^2} \begin{bmatrix} A_u(\omega) & A_{ue}(\omega) \\ A_{ue}^T(\omega) & \sigma^2 \end{bmatrix}$$

DTU

Optimal Design II

$$\Pi(\omega) = \mathbf{E}\left\{\tilde{T}^T(\omega)\tilde{T}(\omega)\right\} = B(\omega)^T B(\omega) + \frac{1}{N}\Sigma(\omega)$$

where

$$B(\omega) = T^* - T_0 \qquad \Sigma(\omega) = T'(\omega)[NP]T'(\omega)$$

$$J = J_p + J_b$$

$$J_p = \frac{1}{N}\int_{-\pi}^{\pi} tr\left[\Sigma(\omega)C(\omega)\right] d\omega$$

$$J_b = \int_{-\pi}^{\pi} B(\omega)C(\omega)B^T(\omega)d\omega d\omega$$

$$\frac{1}{N}\Sigma(\omega) \simeq \frac{n}{N}A(\omega)\begin{bmatrix} A_u & A_{ue} \\ A_{ue}^T & \sigma^2 \end{bmatrix}$$

In this case

$$J_p = \int_{-\pi}^{\pi} tr\left[\Sigma(\omega)C(\omega)\right]d\omega$$

becomes

$$tr\left[\Sigma(\omega)C(\omega)\right] = \frac{\sigma^2 C_{11} - 2Re(C_{21}A_{ue}) + C_{22}A_u}{\sigma^2 A_u - |A_{ue}|^2}A_v$$

Optimal Design II

Theorem:

For the special case $C_{12} = C_{22} = 0$ the optimization of J_p subject to the constraints

$$\beta \mathbf{E}\left\{y^2\right\} + \alpha \mathbf{E}\left\{u^2\right\} < 1$$

is given by

$$u_t = w_t - F y_t$$

where F is the solution to the standard LQG control problem

$$J_c = \beta \mathbf{E}\left\{y^2\right\} + \alpha \mathbf{E}\left\{u^2\right\}$$

and the reference is given by the spectrum

$$A_w = \mu \sqrt{A_v C_{11}} \frac{|1 + G_0 F|^2}{\sqrt{\alpha + \beta |G_0|^2}}$$

Basic observations

Optimal design depend on system and noise characteristics

For pure input constraints ($\beta = 0$) an open loop solution is optimal and the input has a spectrum

$$A(\omega) = \mu\sqrt{A_v C_{11}}$$

If $\beta > 0$ it is always optimal to use a feedback solution.

Special courses - Master projects

- Control of dynamical/stochastic systems, dynamics optimization
- Modelling and System identification
- Kalman filtering (state estimation and monitoring) and fault diagnosis

- Navigation (mobile robots)[Sonardyne, Rovsing, DTU-elektro (or,mb)]
- Wind energy (wind turbine and farms) [hm, Vestas, Siemens, Risø]
- Artificial pancreas [jbj, hm, Novo]
- Fault diagnosis [DTU-elektro (hhn)]

- Adv. system identification (02904)
- Time series analysis (ord. and adv.) [hm]
- Robust and fault tolerant [mb, hhn]
- MPC course [jbj]
- Static and dynamic optimization (42111/02711)