System Identification, Lecture 1

Kristiaan Pelckmans (IT/UU, 2338),

Course code: 1RT880, Report code: 61800 - Spring 2014
F, FRI Uppsala University, Information Technology

20 January 2014
Lecture 1

• Course Overview.

• System Identification in a Nutshell.

• Applications.

• Course Outline.
Prerequisites

- Linear algebra and statistical techniques.

- 120 ECTS credits.

- Courses: Signals and systems, Automatic control I, Automatic control II.

- Ph.D. student.
Course Organisation

Part I.: Basics.

• 7 Lectures.
• 4 Exercise Sessions.
• 5 Computer Labs (Reports Mandatory, 0.5 ECTS).
• 1 Laboratory Session (Report Mandatory, 0.5 ECTS).

→ Written Exam (≈ 20 March 5 ECTS)

Part II. Advanced.

• 5 Lectures.
• Projects.

→ Presentation + Report project (4 ECTS).
Course Organisation

- Lectures (Kristiaan, kp@it.uu.se)
- Exercises (Liang, liang.dai@it.uu.se)
- Compute Labs (Margarida, margarida.silva@it.uu.se)
- Process lab (M. & L.)
Final Score

\[
\text{Score} = \left\lfloor \frac{5}{2} (\text{Score}_1 + \text{Score}_2) \right\rfloor
\]

- Score: 1,2 (U), 3,4,5

- Score\(_1\) = Score obtained on written exam (x/50)

- Score\(_2\) = Score obtained on projects (y/50)

- +Mandatory parts.
Part I.: Basics

SISO:

(i) Overview.

(ii) Least Squares Rulez.

(iii) Models & Representations.

(iv) Stochastic Setup.

(v) Prediction Error Methods.

(vi) Model Selection and Validation.

(vii) Recursive Identification.
Problem Solving Sessions:

1. Aspects of Least Squares.
2. Statistical Aspects: what can go wrong with OLS?
4. Recursive Identification.

5 Computer Labs:

1. Least Squares Estimation: do’s and dont’s.
2. Timeseries Modeling.
3. Recursive Identification.
4. The System Identification Toolbox.
5. MIMO: Kalman Filter and Subspace ID.
Part II.: Advanced

MIMO:

(i) State Space Models.

(ii) Realization Theory.

(iii) Subspace Identification.

(iv) Design of Experiments.

(v) Perspectives.
Projects:

• Identification of an industrial Petrochemical plant
• Identification of an Acoustic Impulse Response
• Identification of Financial Stock Markets
• Identification of a Multimedia stream
• *.*
Course Material

• Lecture Notes: Available from next week in lectures, or online.

• Slides: Available past lectures.

• Solutions exercises. Available past lectures.

• Book: “System Identification”, T. Söderström, P. Stoica, Prentice-Hall, 1989 \[1\]

---

\[1\]see http://www.it.uu.se/research/syscon/Ident, …
Desiderata

Students who pass the course should be able to

1. Describe the different phases that constitute the process of building models, from design of an identification experiments to model validation.
2. Explain why different system identification methods and model structures are necessary/useful in engineering practice.
3. Account for and apply the stochastic concepts used in analysis of system identification methods.
4. Describe and motivate basic properties of identification methods like the least-squares method, the prediction error method, the instrumental variable method, as well as to solve different problems that illustrate these properties.
5. Describe the principles behind recursive identification and its field of application.
6. Explain the usefulness of realization theory in the context of system identification, and how it is employed in subspace identification techniques.
7. Show hands-on experience with analyzing actual data, and have a working knowledge of the available tools. Reason about how to choose identification methods and model structures for real-life problems.
In order to pass the course, I need to have for each of the candidates:

1. Attendance of the lab. session, as well as a filled out copy of the lab report.

2. A filled out report of the computer sessions.

3. A successful written exam (∼ 20 March).

4. A project report.

5. A successful presentation of the project (possibly shared amongst partners in the group).
System

**System (S):** A defined part of the real world. Interaction with the environment are described by input signals, output signals and disturbances.

**Dynamical System:** A system with a memory, i.e. the input value at time $t$ will influence the output signal at the future, i.e. $t' > t$.

![Schematic picture of a system](image)

**Figure 1: Schematic picture of a system**
Figure 2: A System and A Model
Ex.: A Stirred Tank

Figure 3: A Stirred Tank
Ex.: Speech

Figure 1.7 Speech generation.

Figure 1.8 The speech system: $y$: output; $w$: unmeasured disturbance.

Figure 1.9 The speech signal (air pressure). Data sampled every 0.125 ms. (8 kHz sampling rate).
Ex. and...

- Stock (Shock) Market
• Acoustic Noise Cancellation Headset (Adaptive filtering)
● Evolution of the Temperature in the world
• Construction (Strength)
• Robots (Mechanical, Operational, Intellectual)
Social Behavior of Crowd (gossip)
• A human protein-protein interaction network
Models

**Model (M):** A description of a system. The model should capture the essential behavior of the system.

<table>
<thead>
<tr>
<th>Systems</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex</td>
<td>Approximative (Idealization)</td>
</tr>
<tr>
<td>Examine real</td>
<td>Models can answer</td>
</tr>
<tr>
<td>system is costly</td>
<td>many questions.</td>
</tr>
</tbody>
</table>
Applications

- Process Design. Ex. Designing new cars, planes, ... .

- Control Design.
  1. Simple regulators
  2. Simple models, optimal regulators,
  3. sophisticated models.

- Prediction. Ex. Forecast the weather, Predict the Stock market.


- Simulation. Ex. Train new nuclear plant operators, try new operating strategies.

- Fault Detection. Ex. VISA.
Type of Models

• Mental, intuitive or verbal. Ex. Driving a car.


Mathematical Models

- **Analytical Models** *(White-Box models)* Basic laws from physics (...) are used to describe the behavior of a phenomenon (system).
  - Know the physics.
  - Yields physical Interpretation
  - Quite general models. Often Nonlinear

- **System Identification**
  - **Black-Box models** *(Konfektionsmodeller)* ”Choose a standard model (class) and tune up the parameters (...) to the data”.
    * Easy to construct and use.
    * Less general. Linear (-ized)
  - **Grey-Box models** *(Skräddarsydda Modellerer)* ”Derive the model from laws and tune ’some’ parameters to data”.
    * Combines Analytical models and black-box identification.
Figure 4: White-, Black- and Grey-Box Models
Examples of Models

• Nonlinear vs. Linear (superposition principle):

"The net response at a given place and time caused by two or more stimuli is the sum of the responses which would have been caused by each stimulus individually." (Wiki)

• Time-continuous versus Time-discrete

• Deterministic versus Stochastic
System Identification (SI)

**Def.** System Identification is the study of *Modeling* dynamic *Systems* from *experimental data*.

- System Identification is art as much as science.
- Software available (MATLAB)
  - Estimation (Gauss (1809)),
    - Modern System Identification (Åström and Bohlin (1965), Ho and Kalman (1966)),
    - Recent System Identification (L. Ljung, 1977-1978)
    - Textbooks (Ljung 1987, Söderström and Stoica, 1989).
The System Identification Procedure

1. **Collect Data.** If possible choose the input signal such that the data is maximally informative. Display data, and try to get some intuition about the problem at hand.

2. **Choose Model Structure.** Use application knowledge and engineering intuition. Most important and most difficult step (don’t estimate what you know already)

3. **Choose Identification Approach.** How would a good model look like?

4. **Do.** Choose best model in model structure (Optimization or estimation)

5. **Model Validation.** Is the model good enough for our purpose?
FIGURE 1.3 Schematic flowchart of system identification.
Typical Problems to Answer

• How to design the experiment. How much data samples to collect?

• How to choose the model structure?

• How to deal with noise?

• How to measure the quality of a model?

• What is the purpose of the model?

• How do we handle nonlinear and time-varying effects?
System Identification Methods

- **Non-parametric Methods.** The results are (only) curves, tables, etc. These methods are simple to apply. They give basic information about e.g. time delay, and time constants of the system.

- **Parametric Methods (SI)** The results are values of the parameters in the model. These may provide better accuracy (more information), but are often computationally more demanding.
Course Outline

SISO:

(i) Overview.

(ii) Least Squares Rulez.

(iii) Models & Representations.

(iv) Stochastic Setup.

(v) Prediction Error Methods.

(vi) Model Selection and Validation.

(vii) Recursive Identification.
Conclusion

• System identification is the art of building mathematical models of dynamical systems using experimental data. It is an iterative procedure.
  – A real system is often very complex. A model is merely a good approximation.
  – Data contain often noise, individual measurements are unreliable.

• Analytical methods versus system identification (white-, black-, grey box)

• Non-parametric versus Parametric Methods

• Procedure: (a) Collect data, (b) Choose Model Structure, (c) Determine the best model within a structure, (d) Model validation.
An example

Identify a hairdryer: air is fanned through a tube and heated at the inlet. Input $u(t)$: power of the heating device. Output $y(t)$: air temperature.

```matlab
>> load dryer2
>> z2 = [y2(1:300) u2(1:300)];
>> idplot(z2, 200:300, 0.08)
```
Nonparametric Modeling

```matlab
>> z2 = dtrend(z2);
>> ir = cra(z2);
>> stepr = cumsum(ir);
>> plot(stepr)
```
Parametric modeling:

\[ y(t) + a_1 y(t - 1) + a_2 y(t - 2) = b_1 u(t - 3) + b_2 u(t - 4) \]

```matlab
>> model = arx(z2, [2 2 3]);
>> model = sett(model,0.08);
>> u = dtrend(u2(800:900));
>> y = dtrend(y2(800:900));
>> yh = idsim(u,model);
>> plot([yh y]);
```
Pole-zero plot of the model:

```matlab
>> zpth = th2zp(model);
>> zpplot(zpth);
```
Compare the transfer functions obtained from non- and parametric methods:

```matlab
>> gth = th2ff(model);
>> gs = spa(z2); gs = sett(gs,0.08);
>> bodeplot([gs gth]);
```