# System Identification and Data Analysis

Master of Science in Engineering Management

Master of Science in Artificial Intelligence and Automation Engineering
(Curriculum Robotics and Automation)

9 cfu (approximately 72 hours)

# Data Analysis

Master of Science in Artificial Intelligence and Automation Engineering
(Curriculum Intelligent Systems)

Master of Science in Applied Mathematics
(part of the Data and Financial Analysis course)

6 cfu (approximately 48 hours)

A.Y. 2020-2021

Instructors: Andrea Garulli, Marco Casini

## Syllabus

### Part 1: Estimation theory

**1.1 Random variables**. Probability distributions. Mean and covariance. Conditional probability. Gaussian variables.

**1.2 Estimation theory**. Parametric estimation. Properties of estimators. Maximum likelihood estimators. Least squares and Gauss-Markov estimators. Bayesian estimation. Minimum mean square error estimators.

**1.3 Stochastic processes and time-series prediction**. Distributions, mean and covariance function. Stationary processes. Frequency domain representation. Stochastic dynamic systems. Time-series models: AR, MA, ARMA. Time-series prediction.

### Part 2: System Identification

**2.1 System identification theory**. Identification of linear systems: prediction error methods. Input-output models: ARX, ARMAX, OE, BJ. Least squares estimator for linear regression models. Model validation.

**2.2 Practical system identification**. Use of software tools for system identification.

### Part 3: State Estimation

**3.1 State estimation for linear systems**. Non stationary stochastic systems. The state estimation problem and the Kalman filter. Asymptotic properties of the Kalman filter. Recursive system identification.

**3.2 State estimation for nonlinear systems**. State estimation in nonlinear stochastic systems. The Extended Kalman Filter. Advanced nonlinear filtering techniques: unscented filter; sequential Monte Carlo methods.

# Bibliography

Main textbooks

- T. Soderstrom, *Discrete-time Stochastic Systems*, Springer London Ltd, 2nd ed., 2002.
  Library code: 269 - 269a.
- L. Ljung. *Identification: Theory for the user*, 2nd ed., Prentice-Hall, 1999.
  Library code: 71 - 71a - 71b.
- L. Ljung. *System Identification Toolbox: User's guide*, The Mathworks, 2001.
  Library code: A-1-197.

Other reference textbooks

- Probability theory:
  A. Papoulis. *Probability, random variables and stochastic processes*, 3rd edition, McGraw Hill, 1991.
- Systems modelling:
  L. Ljung and T. Glad. *Modelling of dynamical systems*. Prentice-Hall, 1994.
- State estimation, nonlinear filtering:
  F. L. Lewis, *Optimal Estimation*, John Wiley, 1986.
  E. W. Kamen and J. K. Su, *Introduction to Optimal Estimation*, Springer, 1999.

# What do I need to know before starting?

- Linear algebra and calculus
- Basics of probability theory and random variables (though there will be a very short review)
- Basics of dynamic systems theory: input-output and state space representations; transfer function; Z-transform (discrete-time systems)
- Basics of Matlab language

# How to pass the exam

In order to pass the **System Identification and Data Analysis** exam (9 cfu), you have to perform three steps:

(A) the system identification homework;

(B) the state estimation homework;

(C) the oral exam.

Steps (A) and (B) must be done before step (C). Details and schedule of steps (A) and (B) will be comunicated by the instructors during the course.

The date of the oral exams are published in `segreteriaonline`: to participate, you must sign up at least 7 days before the oral exam (a more detailed schedule may appear in the News section of the course website).

The oral exam involves the discussion of the homeworks and questions about all the topics treated during the course (yes, *all!*).

Students taking the **Data Analysis** course (6 cfu) have to do only steps (A) and (C).

# Golden rules

## Rule 1

The **right time** to do the exam is at the end of the course, during the summer exam session. You are strongly encouraged to do that!

In case you need to do the exam in subsequent sessions, you have to deliver the reports of homeworks (A) and (B) at least **7 days** before the date of the oral exam. Therefore, be sure to ask for the homeworks well in advance.

## Rule 2

Do the homeworks by yourself and report them carefully. During the oral exam you will have to explain and motivate the homework solutions.

Website of the course

http://control.dii.unisi.it/siada/

Every information will be there, including links to the recordings of the classes

Tentative schedule:

Part 1 - March (AG)
Part 2 - April (MC)
Part 3 - May (AG)
Exams - June / July

# Course Introduction

# Estimation problems

A large number of fundamental problems in engineering (and beyond) can be formulated as estimation problems

*Examples*

- Interpolation
- Signal filtering
- Time series prediction
- Estimation of mathematical models of dynamic systems (*system identification*)

*Estimation problem*
Find the values of one or more unknown quantities, by using available information on other quantities related to them

## Estimation problems: dealing with uncertainty

An estimation problem is characterized by three fundamental ingredients:

- some prior information on the quantities to be estimated
- a set of data
- a criterion to assess the quality of an estimate

A key feature of every estimation problem is the characterization of the **uncertainty** associated to the estimates.

Once an uncertainty model has been defined, the estimation problems often boil down to the minimization of the uncertainty associated to the estimate.

## Part 1: Estimation Theory

- Mathematical models of non deterministic phenomena:
  - random variables
  - stochastic processes
  - stochastic systems
- Estimation approaches:
  - parametric
  - Bayesian
- Application: time-series prediction

# Part 2: System Identification

**System**: an object or a set of objects of which we want to study properties and behaviours

*Examples*

- An electrical circuit
- An industrial process
- An ecosystem
- The solar system

Possible approaches to system analysis:

1. Experimental tests collecting data
2. Modelling
   - Mental models
   - Verbal models
   - Structures and material models
   - **Mathematical models**

# Classification of mathematical models

| | | |
|---|---|---|
| Static | $\Leftrightarrow$ | Dynamic |
| Stationary | $\Leftrightarrow$ | Non stationary |
| Continuous-time | $\Leftrightarrow$ | Discrete-time |
| Linear | $\Leftrightarrow$ | Nonlinear |
| Deterministic | $\Leftrightarrow$ | Stochastic |
| Lumped parameters | $\Leftrightarrow$ | Distribuited parameters |
| Continuous variables | $\Leftrightarrow$ | Discrete events |

## Construction of Mathematical Models

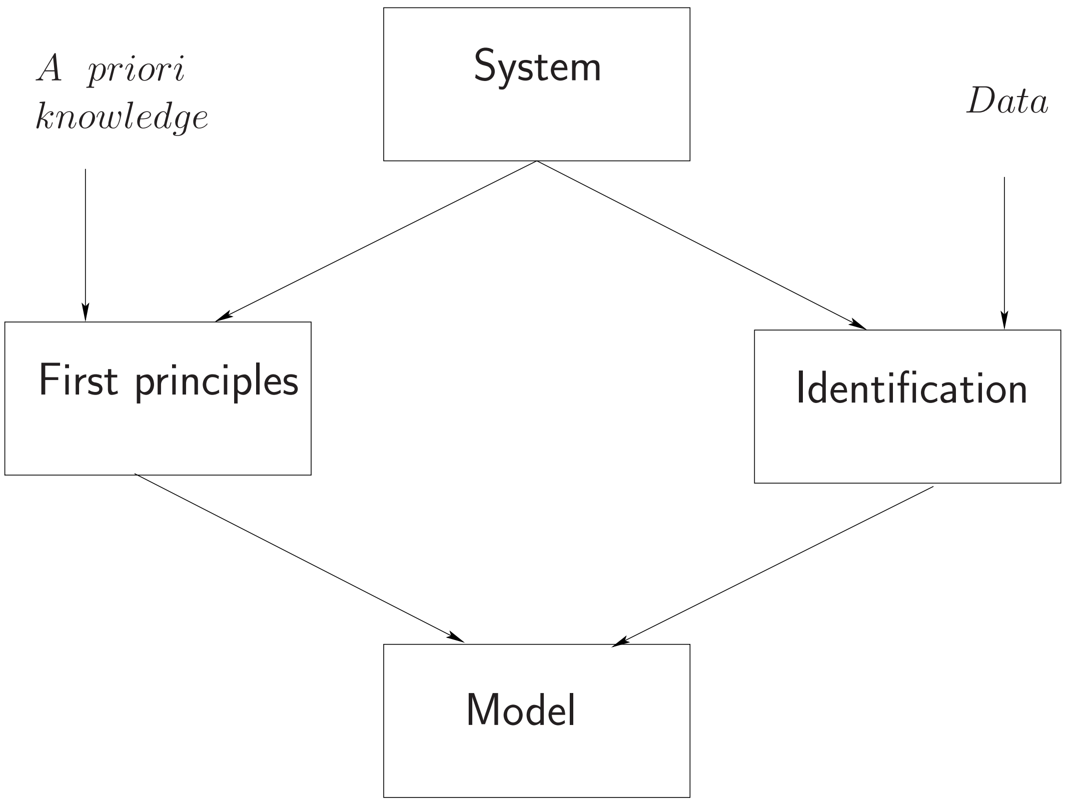
Two possible approaches

1. **Physical models**

   $\Rightarrow$ based on first principles and a priori knowledge

2. **System Identification**

   $\Rightarrow$ based on the observation of the system behaviour (the *data*)




*System Identification* $\longrightarrow$ estimation problem

# Part 3: State Estimation

State-space representation of dynamic systems

- Deterministic: state variables allow to propagate system evolution in future times
- Stochastic: Markov processes, a posteriori distribution

**State estimation problem**: compute an estimate of the state variables, based on the observation of input-output signals

Applications (largely incomplete list):

- mobile robotics
- aerospace
- population dynamics
- ecosystems
- financial analysis
- ...

**Hope you enjoy the course!**