Introduction to Linear and Nonlinear System Identification

Dr.-Ing. Sudchai Boonto

Department of Control System and Instrumentation Engineering
King Mongkuts Unniversity of Technology Thonburi
Thailand

The Important of Models

Simulation : Study the system outputs for the given inputs

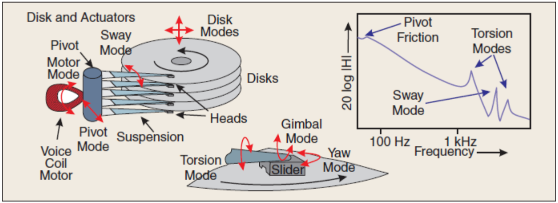
Chemical plant

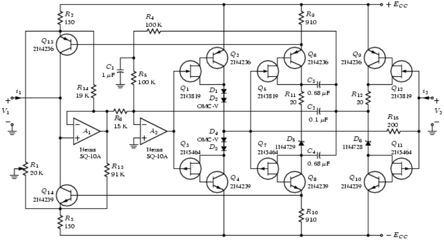
Thermal study

The Important of Models

Design : Compute the system parameters to have a desired output for a given input

Design a electrical, mechanical or chemical installations





The Important of Models

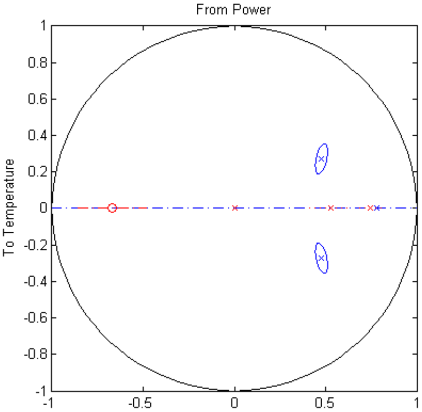
Prediction : Forecast the future values for the output

Weather forecasting, Flood forecasting (Too sad to use a picture in Thailand)

The Important of Models

Control : Model-based controller design

Pole placement controller design for tracking and disturbance rejection



Constructing Models

First principle method

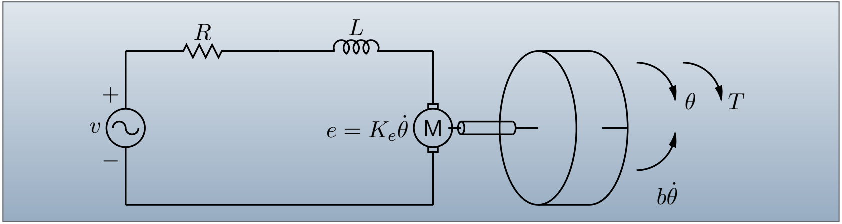
- The Newton's Law
- The law of conservation of energy
- KVL, KCL, etc.

System identification

Based on input/output measured data

- Parametric model
- Nonparametric model

Physical modeling



Assume

$$T = Ki,\ e = K\dot{\theta}$$

Based on Newton's law and KVL

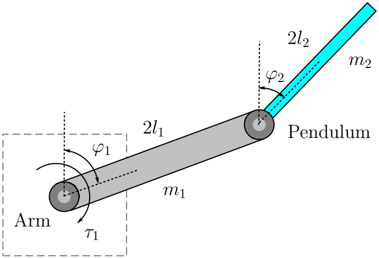
$$J\ddot{\theta} + b\dot{\theta} = Ki$$

$$L\frac{di}{dt} + Ri = v - K\dot{\theta}$$

$$\frac{\Theta}{V} = \frac{K}{s((Js+b)(Ls+R)+K^2)}$$

Physical modeling cont.

Arm-driven inverted pendulum, Kajiwara, H. et al. 1999



The motion equation:

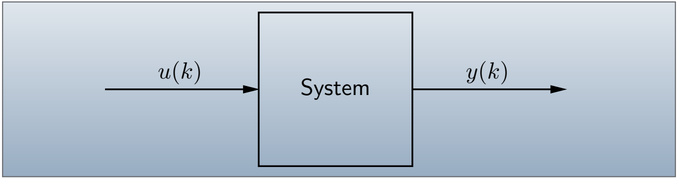
$$M(q)\ddot{q} + C(q,\dot{q}) + G(q) = \tau$$

$$(M_1 - M_2)\ddot{\varphi}_1 + R\cos(\varphi_1 - \varphi_2)\ddot{\varphi}_2$$
$$R\sin(\varphi_1 - \varphi_2)\dot{\varphi}_2^2 + (m_1 + 2m_2)l_1 g\sin(\varphi_1)$$
$$- m_2 l_2 g \sin(\varphi_2) = \tau_1$$

$$R\cos(\varphi_1 - \varphi_2)\ddot{\varphi}_1 + M_2\ddot{\varphi}_2 - R\sin(\varphi_1 - \varphi_2)\dot{\varphi}_1^2$$
$$- m_2 l_2 g \sin(\varphi_2) = 0$$

$$M_1 = \frac{4}{3}m_1 l_1^2 + \frac{4}{3}m_2 l_2^2 + 4m_2 l_1^2$$
$$M_2 = \frac{4}{3}m_2 l_2^2, \qquad R = 2m_2 l_1 l_2$$

System Identification



Apply a specific input $u(k)$ and measure the output $y(k)$ and represent it as a function of preceding values of input and output.

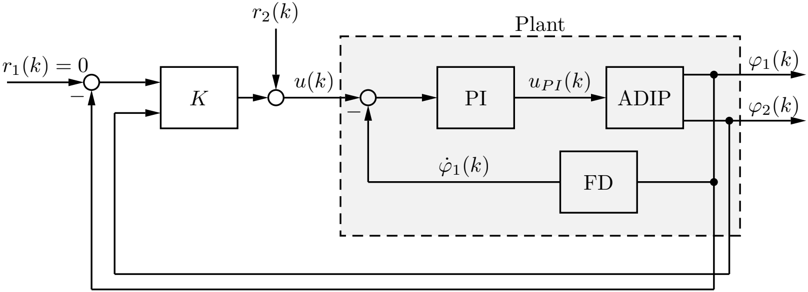
$$y(k) = G\left(u(k), u(k-1), u(k-2), \ldots, y(k-1), y(k-2), \ldots\right)$$

Transfer function representation

$$G(z) = \frac{bz}{z^2 + a_1 z + a_2}$$

System Identification

Arm-driven inverted pendulum, Kajiwara, H. et al. 1999



$$y(k) = G\big(u(k), u(k-1), \ldots, \varphi_1(k), \varphi_1(k-1), \ldots, \\ \varphi_2(k), \varphi_2(k-1), \ldots\big),$$

Physical or Identified Model?

Physical Models

$$G(s) = \frac{K}{s((Js+b)(Ls+R)+K^2)}$$

- ✔ Direct relation with physical parameters
- ✖ Need complete process knowledge
- ✖ Physical parameters should be known
- ✖ High order, approximative

Physical or Identified Model?

Identified Models

$$G(z) = \frac{bz}{z^2 + a_1 z + a_2}$$

- ✔ Appropriate for controller design
- ✔ Simple and efficient
- ✖ Limited validity (operating point, type of input), sensors, measurement noise
- ✖ Unknown model structure

Type of models

- Dynamic/Static
- SISO/MIMO
- Deterministic/Stochastic
- Linear/Nonlinear
- Time-invariant/Time-varying
- Causal/Noncausal
- Zero initial condition/Nonzero initial condition

Type of representations

- Input/Output representation
- State-space representation
- Time-domain representation
- Frequency-domain representation
- Continuous-time representation
- Discrete-time representation

Input/output time-domain continuous-time representation:

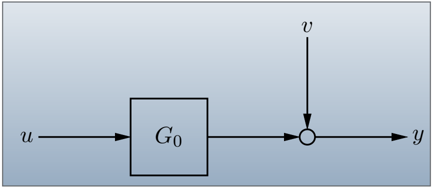
$$y(k) = H(u(k)), \ \ -\infty < \tau \leq t$$

State-space time-domain discrete-time representation:

$$\begin{aligned} x(k+1) &= f(x(k), u(k), k), \ x(k_0) = x_0 \\ y(k) &= g(x(k), u(k), k) \end{aligned}$$

System Identification

To-be-identified system

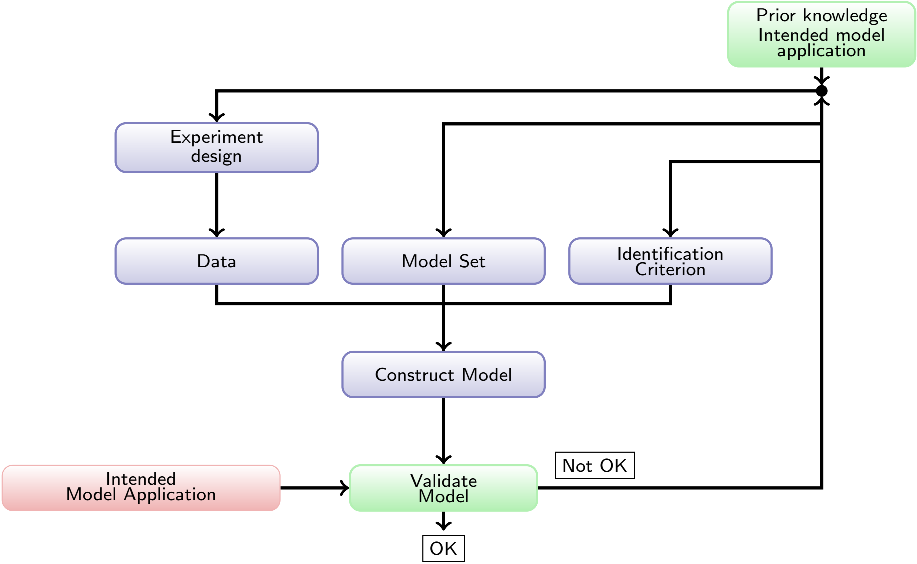


$u(k)$ is the discrete-time input which can be freely chosen

$y(k)$ is the discrete-tie output which can be measured and is made up of

- a contribution due to $u(k)$ i.e. $G_0 u(k)$
- a contribution independent of $u(k)$ i.e. the disturbance $v(k)$

System Identification procedure



Identification Criterion

Measure the distance between a data set $\{u, y\}_{t=1,\dots,N}$ and a particular model.

In this course, we will consider two criteria

- Prediction Error Method (PEM) delivering a discrete-time transfer function as model of G_0
- Empirical Transfer Function Estimate (ETFE) delivering an estimate of the frequency response of G_0

Identification Criterion

Why those

- PEM is the most used method in practice and the one delivering the most tools to validate a model
- ETFE is used to have the first idea of the system and facilitate the use of PEM

Other criteria: subspace identification, instrumental variable methods, Maximum likelihood method, ...

Model set

Complexity of models (order, number of parameters) to be determined. We will talk these topics later.

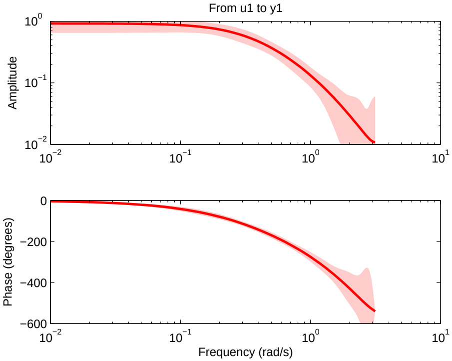
Experiment Design

- Choice of the type of excitation signal
  - sum of sinusoids (multisine)
  - realization of (filtered) white noise or alike
- which frequency content?
- which duration?

Experiment design is very important since it has a direct influence on the quality of the model.

Model Validation

- comparing the actual output of the system with the output predicted by the model
- determining the uncertainty of the system e.g. in the frequency domain



History

- basic principle least square (LS) from Gauss (1809)
- development based on theories of
  - stochastic processes
  - statistics
- strong growth in sixties and seventies Åström and Bohin (1965), Åström and Eykhoff (1971)
- brought to technological tools in nineties (Matlab Toolboxes for either time-domain of frequency domain), as well as to professional industrial control packages (Aspen, SMOC-PRO, IPCOS, Tai-Ji Control, AdaptX, ...)

Reference

1. Lecture note on *System Identification* Karimi, A., EPFL, Switzerland
2. Lecture note on *System Identification* Bombois, X. and Van den Hof, P.M.J., Netherlands