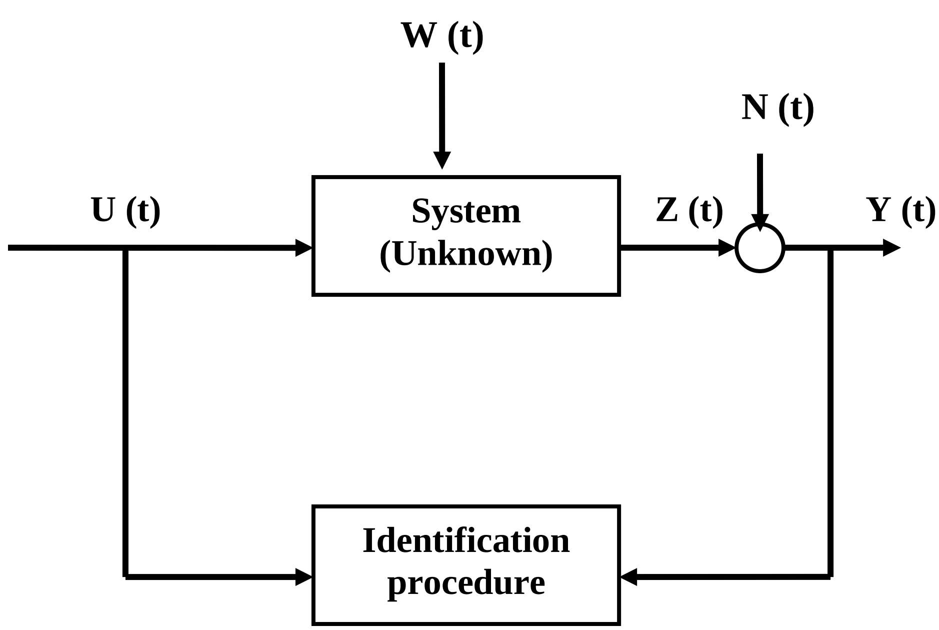
Introduction

The term identification was introduced by Zadeh (1956) as a generic expression for the problem of allowing you to build mathematical models of a dynamic system based on measured data. Essentially by adjusting parameters within a given model until its output coincides as well as possible with the measured output. A good test is to take a close look at the model's output compared to the measured one on a data set that wasn't used for the fit ("Validation Data"). If we denote input, output, and disturbances by **U**, Y, **W** and **N**, respectively, the relationship can be depicted in the following figure.



Where:

U(t) : is the input to the system

Z(t) : is the output

W(t) : is the input disturbance

N(t) : is the observation noise or disturbance

Y(t) : is the measured output

Basic elements of identification:

The procedure to determine a model of a dynamical system from observed Input - output data involves four basic ingredients:

- The nature of the input.
- Selection of model structure or determining the order of the linear model.
- Selection of a suitable criterion for determining the accuracy of the model.
- Model validation.

☒ ***The input***

1. The computations can be simplified if a special types of input signals are chosen such as step, impulse, PRBS, ---etc.
2. The input should excite all the modes of the system.
3. The choice of the input depends on the type of the input that the process might undergo under normal conditions (operations).
4. The level of the input is chosen such that the process will not drift to nonlinearity or to damage the product of the process.

☒ ***The model***

Models can be classified in different approaches

1. Linear and nonlinear:

- A linear system is easily identified because of the linear property of superposition.
- Linearization is very important as an approximation to a nonlinear system.

2. Stationary (time invariant) and non-stationary

- Non stationary systems have parameters that vary with time.

  Y = a (t) x (t)
- If the parameters of nonstationary systems vary slowly with time in comparison with time required for an adequate identification, it can be considered as stationary.

3. Continuous time- discrete time

- A mathematical model that describes the relationship between continuous signals is called time continuous. Differential equations are often used to describe such a relationship.
- A model that directly expresses the relationship between the values of the signals at the sampling instants is called discrete time or sampled model. Such model is typically described by difference equations.
- Transformation can be done straight forward.

4. Single input and multi inputs

Identification technique is simplified when the state of the system is affected by one input as compared with a state that is affected by a combination of several inputs.

5. Deterministic – stochastic

- Stochastic process: has mainly probabilistic knowledge of the exact state of the system. Relation between variables is given in terms of statical values.
- Deterministic (no probabilities): non-parametric models which can be described by (step, frequency,) response. But, parametric models which are expressed by differential equations, algebric equations, T.F's etc.

6. Lumped- distributed

- Distributed parameter model: many physical phenomena are described mathematically by partial differential equations. The events are dispersed over the space variables.
- Lumped models: the events are described by a finite number of changing variables; such models are usually expressed by ordinary differential equations.

7. Static – dynamic

- Static: if there are direct, instantaneous links between inputs and outputs, the system is termed static. The input and output are related by algebric equations.
- Dynamic: inputs and outputs are related by differential equations (which will make the current input affect future outputs also).

☒ ***The criterion***

The criterion is a minimization of a scalar loss function. The first attempt to implement identification is by Gauss who formulated the problem as an optimization problem using least squares.

The criterion is expressed as a function of the error between the process output and the model output.

$$v = \int e^2(t)dt$$

$$v = \sum_{k=1}^{N} e_k^2(k)$$

(i.e. minimization of the sum of the squares of errors)

☒ ***Model validation***

The identification model must undergo model validation to validate that the model represents the process or system. However, it must recognize that this objective of proving the model is correct can only be approached and not achieved.

Classification of identification methods

A. Identification according to I/P

ON-Line

Batch

Processing is done immediately after data collection

Real time

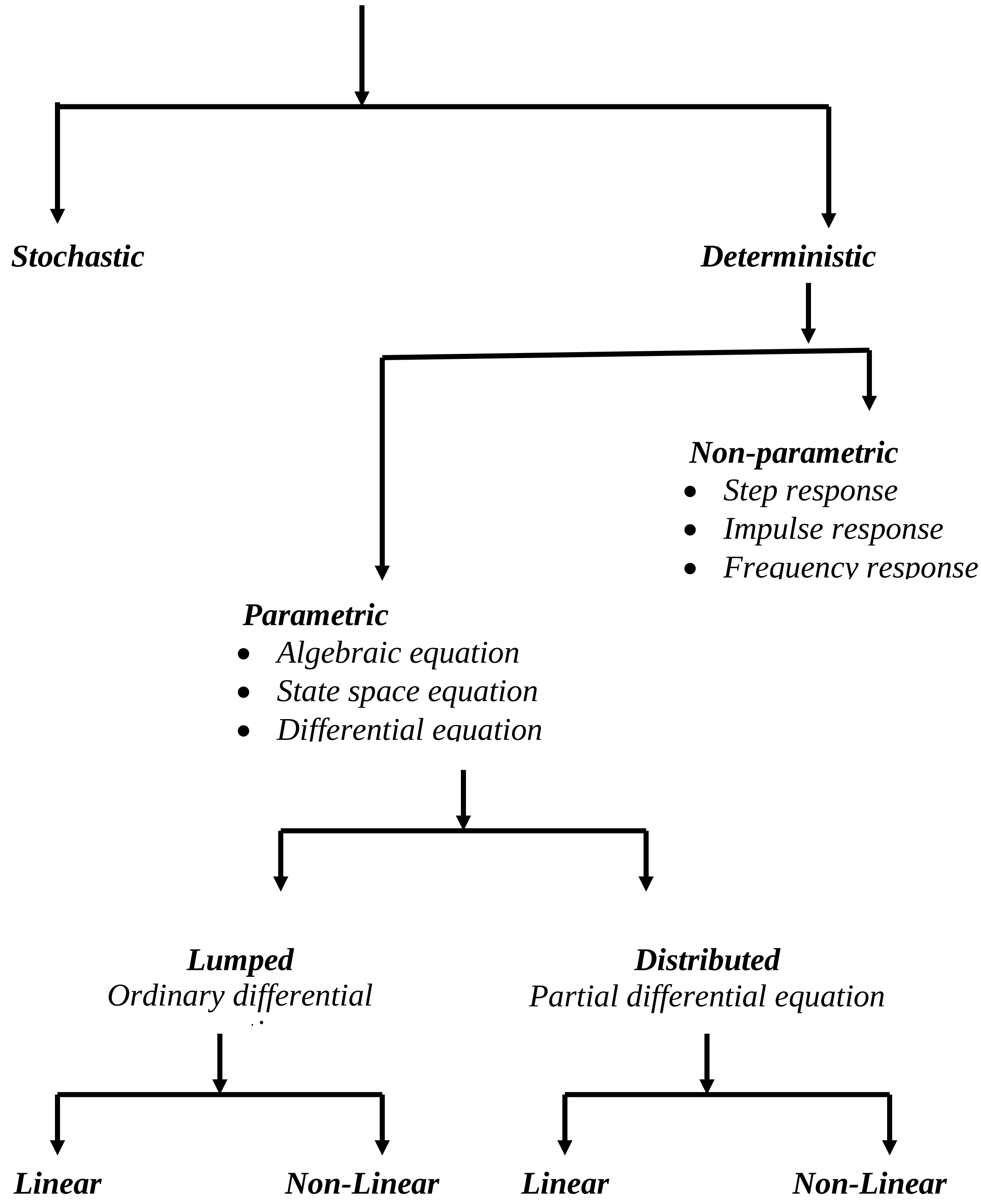
- *Processing is done during sampling interval*
- *Parameters are updated within the sampling period*

OFF-Line

Data are:

- *Collected*
- *Stored*
- *Processed after time*

B. Identification according to model



C. Identification according to criterion

Classical methods

- *Step response*
- *Impulse response*
- *Frequency response*
- *Correlation method*

Equation error methods.

Least square

Model adjustment methods

Recursive least square

<u>The Basic Steps of System Identification</u>

The identification process amounts to repeatedly selecting a model structure, computing the best model in the structure, and evaluating this model's properties to see if they are satisfactory. The cycle can be itemized as follows:

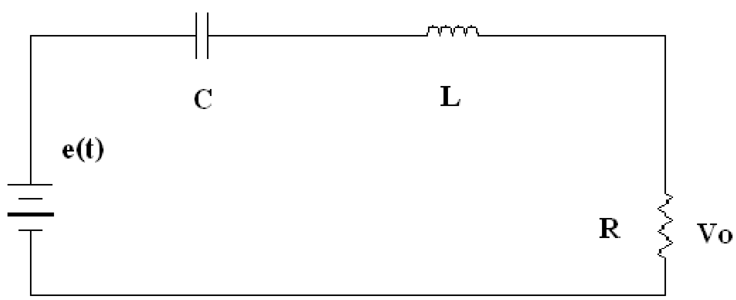
1. Design an experiment and collect input-output data from the process to be identified.
2. Examine the data. Polish it so as to remove trends and outliers, and select useful portions of the original data. Possibly apply filtering to enhance important frequency ranges.
3. Select and define a model structure according to the input-output data.
4. Compute the best model in the model structure according to the input-output data and a given criterion of fit.
5. Examine the obtained model's properties
6. If the model is good enough, then stop; otherwise go back to Step 3 to try another model set. Possibly also try other estimation methods (Step 4) or work further on the input-output data (Steps 1 and 2).

Types of models

☒ ***Analytical or white box***

1. Derived from first principles laws, physical, chemical, biological, etc.
2. They do not depend on data.
3. When they involve parameters that are unknown, these parameters have a physical meaning even if they are estimated from data.

Ex: Consider RLC circuit shown below, derive its state space equation.



Notes:

- State variables of a dynamic system are the smallest set of variables which determine the state of the dynamics system. If at least n variables $X_1(t)$, $X_2(t)$, Xn (t) are needed to describe the behavior of dynamic system, then such ***n*** variables are a set of state variables.
- The classical way of writing equations of electrical network is based on the loop method and the node method, which are formulated from the two laws of Kirchhoff.

  Let $X_1(t) = v_c(t)$

  $X_2(t) = i(t)$

- The reason for this choice is because the state variables are directly related to the energy storage elements of a system.
- The state equations for the network are written as:

<u>Current in C</u>

$$i_C = i(t) = C\frac{dv_C}{dt} = C * \dot{v}_C - 1 -$$

<u>By KVL</u>

$$e(t) - v_R - v_L - v_C = 0$$

$$e(t) - i(t) * R - L\frac{di_L}{dt} - v_C = 0 - - - - - - - - - - - - - - - - - - - 2$$

Re-arrange the state equations

$$\dot{v}_C = \dot{X}_1 = \frac{1}{C} * X_2$$

$$\dot{i}_L = \dot{X}_2 = \frac{1}{L} * e(t) - \frac{R}{L} * X_2 - \frac{1}{L} * X_1$$

The state equation in matrix form is written as:

$$\dot{X} = AX + BU - - - - - - - - \qquad \textbf{\textit{State equation}}$$

Where:

$X : n \times 1$ State vector

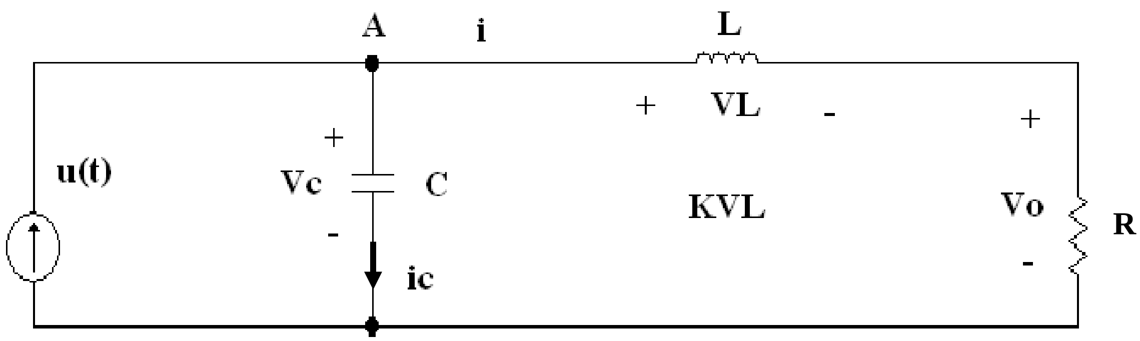
$U : p \times 1$ Input vector

$A : n \times n$ Coefficient matrix with constant elements

$B : n \times p$ Coefficient matrix with constant elements

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{C} \\ -\frac{1}{L} & -\frac{R}{L} \end{bmatrix} \times \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} \times e(t)$$

Ex: Consider RLC circuit shown below; derive its state space equation and its transfer function



There are two states:

Let $X_1(t) = v_c$

$X_2(t) = i_L$

KCL at node A

$$u(t) = i_C + i_L$$

$$\because i_C = C\frac{dv_C}{dt}$$

$$\therefore u(t) = C\dot{v}_C + i_L$$

$$\dot{v}_C = \frac{1}{C}u(t) - \frac{1}{C}i_L$$

$$\dot{X}_1 = \frac{1}{C}u(t) - \frac{1}{C}X_2 \text{ -------------------------------- } 1$$

By KVL

$$-v_R - v_L + v_C = 0$$

$$v_L = v_C - v_R$$

$$L\frac{di_L}{dt} = v_C - i_L * R$$

$$\dot{i}_L = \dot{X}_2 = \frac{1}{L}X_1 - \frac{R}{L}X_2 \text{ ------------------------------ } 2$$

The output: $V_O = i_L R = RX_2$

The state equation in matrix form is written as:

$$\dot{X} = AX + BU \text{ -------- State equation}$$

$$\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{C} \\ \frac{1}{L} & -\frac{R}{L} \end{bmatrix} \times \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} \frac{1}{C} \\ 0 \end{bmatrix} \times u(t)$$

The output equation in matrix form:

$$Y = V_O = CX$$

$$\therefore V_O = \begin{bmatrix} 0 & R \end{bmatrix} * \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

The transfer function is defined as:

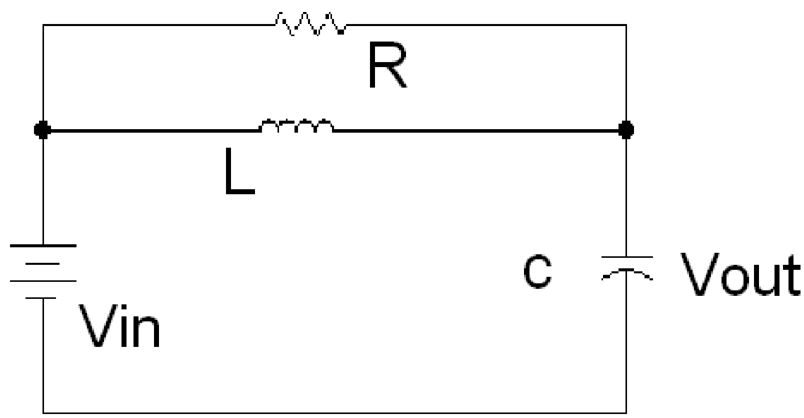
$$G(S) = \frac{output}{input} = C*(SI - A)^{-1} * B$$

Note: $A^{-1} = \frac{adjoint\, A}{determinant}$

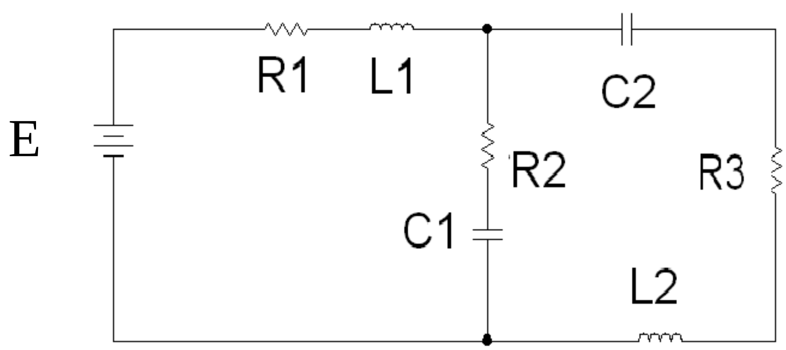
$$G(S) = \frac{\frac{R}{LC}}{S^2 + \frac{R}{L}*S + \frac{1}{LC}}$$

Home work

A) Consider the RLC circuit shown below, write state equations, then derive the transfer function



B) Consider the RLC circuit shown below, write the state equations ONLY



☒ ***Experimental or black box***

1. Both model structure and parameters are derived from measured data.
2. No direct relationship to first principles.