



KAPITAŁ LUDZKI  
NARODOWA STRATEGIA SPÓJNOŚCI



Politechnika Wroclawska

UNIA EUROPEJSKA  
EUROPEJSKI  
FUNDUSZ SPOLECZNY



**ROZWÓJ POTENCJAŁU I OFERTY DYDAKTYCZNEJ POLITECHNIKI WROCŁAWSKIEJ**

Wrocław University of Technology

Computer Engineering

Krzysztof Brzostowski, Jarosław Drapała

# SYSTEMS MODELLING AND ANALYSIS

Wrocław 2011

Projekt współfinansowany ze środków Unii Europejskiej w ramach  
Europejskiego Funduszu Społecznego

Wrocław University of Technology

## Computer Engineering

Krzysztof Brzostowski, Jarosław Drapała

# SYSTEMS MODELLING AND ANALYSIS

Wrocław 2011

Reviewer

Jerzy Świątek

**Project Office**

ul. M. Smoluchowskiego 25, room 407

50-372 Wrocław, Poland

Phone: +48 71 320 43 77

Email: [studia@pwr.wroc.pl](mailto:studia@pwr.wroc.pl)

Website: [www.studia.pwr.wroc.pl](http://www.studia.pwr.wroc.pl)



## Contents

1. Dynamic Systems Modelling and Analysis .....	5
1.1. Differential Equation .....	5
1.2. State Vector Description .....	10
1.3. Transfer Function .....	16
1.4. Linear Systems Analysis .....	24
1.5. Nonlinear Systems Analysis .....	28
2. Identification .....	32
2.1. Determination of the System Parameters .....	33
2.2. The choice of the Best Model: Deterministic Case .....	40
2.3. Parameter Estimation of the System's Characteristic .....	51
2.4. The choice of the Best Model: Probabilistic Case .....	65
2.5. Example Application of the System Identification for Real Life Problem .....	71
Bibliography.....	75

# 1. Dynamic Systems Modelling and Analysis

## 1.1. Differential Equation

**Example 1.1 – Pendulum.** Let us consider a pendulum, composed of a ball attached to a solid wire (see Figure 1.1). The pendulum hangs down below the pivot point. Mass of the ball is  $m$ ,  $l$  is the wire's length,  $s$  stands for the distance made by the ball from its resting equilibrium position, and  $y$  denotes the angular displacement measured from the equilibrium position.

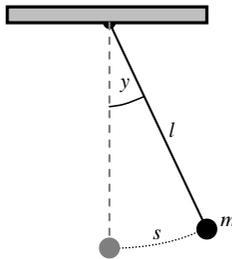


Figure 1.1. Pendulum.

After being displaced from its equilibrium position and released, pendulum begins to swing back and forth. Its motion may be described by changes of the angle  $y$  over time  $t$ . This description is given by a function  $y(t)$ . The question is: what does the function  $y(t)$  look like? For a given pendulum, specified by parameters  $l, m$  and for the initial displacement, function  $y(t)$  should describe the system behavior.

Later on we work out dynamic model of the pendulum in the form of differential equation and we introduce a method of solving linear differential equations.

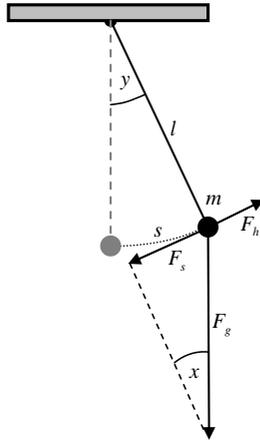
**Model.** We start analysis from the specification of forces that act upon a ball (see Figure 1.2).  $F_g$  is gravity,  $F_s$  denotes its component tangent to direction of movement and  $F_h$  stands for inertia. The force of interest is  $F_s$ , because it acts along the ball's trajectory. Note that

$$F_s = F_g \sin y(t). \quad (1.1)$$

For simplicity, let us assume that friction that acts in the pivot point is proportional to ball's velocity:

$$F_h = k \frac{ds(t)}{dt}, \quad (1.2)$$

where  $k$  is constant of proportionality.



**Figure 1.2. Forces acting upon a ball.**

Now, we can apply the second Newton's law of motion, which states that:

$$\sum_i F_i = m \frac{d^2 s(t)}{dt^2},$$

where  $\sum_i F_i$  is the net force acting upon the system. Thus, we have:

$$F_h - F_s = m \frac{d^2 s(t)}{dt^2}. \quad (1.3)$$

Evaluating formula (1.1):

$$F_s = F_g \sin y(t) = mg \sin y(t),$$

and substituting it, together with (1.2), to equation (1.3), yields:

$$k \frac{ds(t)}{dt} - mg \sin y(t) = m \frac{d^2 s(t)}{dt^2}.$$

Rearranging above formula in accordance with order of derivative gives:

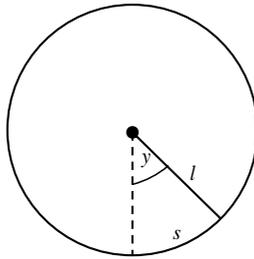
$$\frac{d^2 s(t)}{dt^2} - \frac{k}{m} \frac{ds(t)}{dt} + g \sin y(t) = 0. \quad (1.4)$$

Variable  $s$  may be expressed in terms of  $y$  and  $l$ . Figure 1.3 illustrates that proportion between the angle  $y$  and  $2\pi$  radians equals to proportion between the sector  $s$  and the circle circumference  $2\pi l$ :

$$\frac{y}{2\pi} = \frac{s}{2\pi l}.$$

The proportion given above results in:

$$s = ly.$$



**Figure 1.3. Relation between the angle  $y$  and the distance  $s$ .**

Equation (1.4) may be rewritten in the form:

$$\frac{d^2 y(t)}{dt^2} - \frac{k}{m} \frac{dy(t)}{dt} + \frac{g}{l} \sin y(t) = 0.$$

To simplify notation, let us use  $y'(t)$  to denote derivative of  $y$  with respect to  $t$ :

$$y''(t) + a_1 y'(t) + a_0 \sin y(t) = 0, \quad (1.5)$$

where

$$a_1 = -\frac{k}{m}, \quad a_0 = \frac{g}{l},$$

and initial conditions  $y(0)$  are known.

Equation (1.5) is the second order nonlinear differential equation describing motion of the pendulum presented on Figure 1.1.

Note, that for small angles  $\sin y \approx y$  holds, which allows to propose linear differential equation:

$$y''(t) + a_1 y'(t) + a_0 y(t) = 0. \quad (1.6)$$

This equations describes the system behavior only for small displacements from the resting position.

**Example 1.2 – Damper.** Let us consider a forced mass-spring-damper system, composed of a block attached to the spring and a dashpot filled up with liquid (see Figure 1.4). Mass of the block is  $m$ ,  $\mu$  is the damping factor,  $k$  stands for modulus of the spring elasticity,  $y$  is displacement from the equilibrium position and  $u$  is the force acting on the spring.

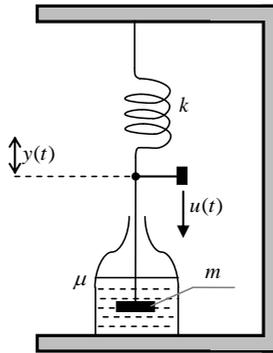


Figure 1.4. Forced mass-spring-damper system.

**Model.** The damping force generated by the dashpot is proportional to velocity of the block:

$$F_{\mu} = \mu \frac{dy(t)}{dt}. \quad (1.7)$$

The spring generates a force proportional to displacement:

$$F_k = k y(t). \quad (1.8)$$

The net force acting upon the block equals to  $u(t) - F_{\mu} - F_k$ , which gives the following form of the second Newton's law of motion:

$$u(t) - F_{\mu} - F_k = m \frac{d^2 y(t)}{dt^2}. \quad (1.9)$$

Substituting (1.7) and (1.8) to (1.9) yields:

$$u(t) - \mu \frac{dy(t)}{dt} - ky(t) = m \frac{d^2 y(t)}{dt^2},$$

and applying  $\frac{dy(t)}{dt} \equiv y'(t)$ :

$$my''(t) + \mu y'(t) + ky(t) = u(t).$$

Dividing both sides by  $m$  gives:

$$y''(t) + \frac{\mu}{m} y'(t) + \frac{k}{m} y(t) = \frac{1}{m} u(t).$$

Denoting:

$$a_1 = \frac{\mu}{m}, a_0 = \frac{k}{m}, b_0 = \frac{1}{m},$$

We obtain the second order differential equation:

$$y''(t) + a_1 y'(t) + a_0 y(t) = b_0 u(t). \quad (1.10)$$

Compare models (1.6) and (1.10). The first model describes linearized unforced system, but left sides of equations are identical. For  $u(t) = 0(t)$  both models are identical. Different physical objects have similar description, which means that variables of both systems are related in the same way. Therefore, analysis made on the model (1.10) applies for all possible systems that may be described by this equation. Another example of the system with description (1.10) is RLC circuit. Interpretation of variables is different, but model remains the same.

In general, **differential equation** is equation of the form (Bro06, Oga02):

$$f\left(\frac{d^m y(t)}{dt^m}, \frac{d^{m-1} y(t)}{dt^{m-1}}, \dots, \frac{dy(t)}{dt}, y(t); \frac{d^v u(t)}{dt^v}, \frac{d^{v-1} u(t)}{dt^{v-1}}, \dots, \frac{du(t)}{dt}, u(t)\right) = 0. \quad (1.11)$$

Linear differential equation is particular example of the above equation:

$$\begin{aligned} & \frac{d^m y(t)}{dt^m} + a_{m-1} \frac{d^{m-1} y(t)}{dt^{m-1}} + \dots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = \\ & = b_v \frac{d^v u(t)}{dt^v} + b_{v-1} \frac{d^{v-1} u(t)}{dt^{v-1}} + \dots + b_1 \frac{du(t)}{dt} + b_0 u(t), \end{aligned} \quad (1.12)$$

where  $m \geq v$  holds.

## Exercises

**Exercise 1.1.** Work out differential equation describing the process of hot cup of coffee cooling down. Assume, that temperature of environment is constant during the process. As a time varying variable take temperature of a cup of coffee  $y(t)$ . Initial value of this variable denote by  $y_0 \equiv y(0)$ .

**Exercise 1.2.** Work out the model of commonly known *RLC* circuit, where  $R$  stands for electrical resistance,  $L$  for inductance and  $C$  for capacitance. Potential drop across a capacitor equals to  $\frac{q}{C}$ , where  $q$  is the charge on the capacitor. Applied voltage  $e(t)$  is an input and resulting current  $i(t)$  is an output.

**Exercise 1.3.** Note similarities between models from exercise 1.1 and 1.2. Point out analogy between variables and parameters of both systems.

### 1.2. State Vector Description

In order to relate variables defined in the system, the following **state vector description** may be used (Oga02, Vuk03):

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = F(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(t) = G(\mathbf{x}(t), \mathbf{u}(t)) \end{cases}, \mathbf{x}(0), \mathbf{u}(t) \quad (1.13)$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x^{(1)}(t) \\ x^{(2)}(t) \\ \vdots \\ x^{(R)}(t) \end{bmatrix}, \quad \mathbf{u}(t) = \begin{bmatrix} u^{(1)}(t) \\ u^{(2)}(t) \\ \vdots \\ u^{(S)}(t) \end{bmatrix}, \quad \mathbf{y}(t) = \begin{bmatrix} y^{(1)}(t) \\ y^{(2)}(t) \\ \vdots \\ y^{(L)}(t) \end{bmatrix}.$$

The vector  $\mathbf{x}$  contains time-varying variables that uniquely describe system's state. The vector  $\mathbf{u}$  contains input variables associated to external action upon the system. The vector  $\mathbf{y}$  contains output variables associated to the system action upon its environment. Vector functions  $F$  and  $G$  define the system of state equations and output equations, respectively. Note that in contrary to differential equations, the state vector description uses intermediate state variables  $\mathbf{x}$  to express relation between the input  $u(t)$  and the output  $y(t)$ . There is differential equation relating  $\mathbf{u}(t)$  and  $\mathbf{x}(t)$  and static function  $G$  describing the output signal  $y(t)$ .

State equations have form:

$$\begin{cases} \frac{dx^{(1)}(t)}{dt} = f_1(x^{(1)}(t), x^{(2)}(t), \dots, x^{(R)}(t); u^{(1)}(t), u^{(2)}(t), \dots, u^{(S)}(t)) \\ \frac{dx^{(2)}(t)}{dt} = f_2(x^{(1)}(t), x^{(2)}(t), \dots, x^{(R)}(t); u^{(1)}(t), u^{(2)}(t), \dots, u^{(S)}(t)) \\ \dots \\ \frac{dx^{(R)}(t)}{dt} = f_R(x^{(1)}(t), x^{(2)}(t), \dots, x^{(R)}(t); u^{(1)}(t), u^{(2)}(t), \dots, u^{(S)}(t)) \end{cases} ,$$

and output equations have form:

$$\begin{cases} y^{(1)}(t) = g_1(x^{(1)}(t), x^{(2)}(t), \dots, x^{(R)}(t); u^{(1)}(t), u^{(2)}(t), \dots, u^{(S)}(t)) \\ y^{(2)}(t) = g_2(x^{(1)}(t), x^{(2)}(t), \dots, x^{(R)}(t); u^{(1)}(t), u^{(2)}(t), \dots, u^{(S)}(t)) \\ \dots \\ y^{(L)}(t) = g_L(x^{(1)}(t), x^{(2)}(t), \dots, x^{(R)}(t); u^{(1)}(t), u^{(2)}(t), \dots, u^{(S)}(t)) \end{cases} .$$

Linear system is described by a particular form of state vector description (1.13):

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases} , \quad (1.14)$$

where  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$  are matrices of the form:  $\mathbf{A} \in \mathcal{R}^{R \times R}$ ,  $\mathbf{B} \in \mathcal{R}^{R \times S}$ ,  $\mathbf{C} \in \mathcal{R}^{L \times R}$ ,  $\mathbf{D} \in \mathcal{R}^{L \times S}$ .

**Example 1.3 – Damper.** Let us take differential equation (1.10) describing behavior of the forced spring-mass-damper system. In order to describe it by state vector description, we introduce two state variables:

$$\begin{cases} x^{(1)}(t) \equiv x(t) \\ x^{(2)}(t) \equiv \frac{dx(t)}{dt} \end{cases} .$$

Variable  $x^{(1)}$  stands for displacement and  $x^{(2)}$  is velocity. These two time varying variables give full information about the current state of the system. Now, we work out state vector description on the basis of differential equation (1.10). It is required to find formulas for  $\frac{dx^{(1)}(t)}{dt}$  and  $\frac{dx^{(2)}(t)}{dt}$  as relations involving  $x^{(1)}$ ,  $x^{(2)}$  and system parameters. The first formula is a consequence of state variable definition:

$$\frac{dx^{(1)}(t)}{dt} \equiv \frac{dx(t)}{dt} = x^{(2)}(t).$$

The second formula is derived from (1.10):

$$\frac{dx^{(2)}(t)}{dt} + a_1 \frac{dx^{(1)}(t)}{dt} + a_0 x^{(1)}(t) = b_0 u(t),$$

which gives:

$$\frac{dx^{(2)}(t)}{dt} = -a_0 x^{(1)}(t) - a_1 x^{(2)}(t) + b_0 u(t).$$

As the output we decide to take  $y^{(1)}(t) = x^{(1)}(t)$ .

Eventually, we obtain state vector description in the form:

$$\begin{cases} \frac{dx^{(1)}(t)}{dt} = x^{(2)}(t) \\ \frac{dx^{(2)}(t)}{dt} = -\frac{k}{m} x^{(1)}(t) - \frac{\mu}{m} x^{(2)}(t) + \frac{1}{m} u(t) \\ y(t) = x^{(1)}(t) \end{cases} \quad (1.15)$$

Equivalent matrix representation has the form:

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = \begin{bmatrix} 0 & 1 \\ -k/m & -\mu/m \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 1/m \end{bmatrix} u(t) \\ \mathbf{y}(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}(t) \end{cases} \quad (1.16)$$

**Example 1.4 – State vector description and SIMULINK.** On the basis of state vector description it is easy to implement the system model in SIMULINK environment, which is a part of MATLAB software.

Implementation of state vector (1.15) is illustrated in Figure 1.5.

The model is composed of pre-defined building blocks representing mathematical operations. Blocks labeled by  $\frac{1}{s}$  are integrators and they integrate input value over time, beginning from the initial value, up to the total simulation time. Initial values are set up in the block's parameters window. Outputs of these two blocks are values of  $x^{(1)}(t)$  and  $x^{(2)}(t)$  at the time  $t$  solved numerically by Runge-Kutty method (Bro06). Another blocks allow to define constant parameters: square blocks with no inputs give constant value as output ('m' is  $m=2$ ), triangle blocks are constant gains ('k' is  $k=1.1$  and 'mu' is  $\mu=1.2$ ).

There are also blocks that do summation, division, and multiplication (the 'product' block). The 'clock' block returns current time  $t$ . The group of blocks on the bottom left implement input signal  $u(t) = e^{-0.1t} \sin t$ . Output signals  $x^{(1)}(t)$  and  $x^{(2)}(t)$  are illustrated in Figure 1.6.

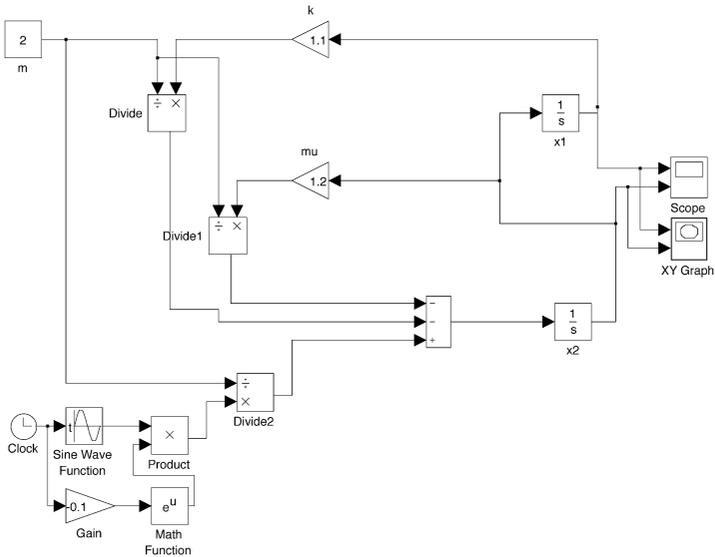


Figure 1.5. SIMULINK implementation of a forced mass-spring-damper system model.

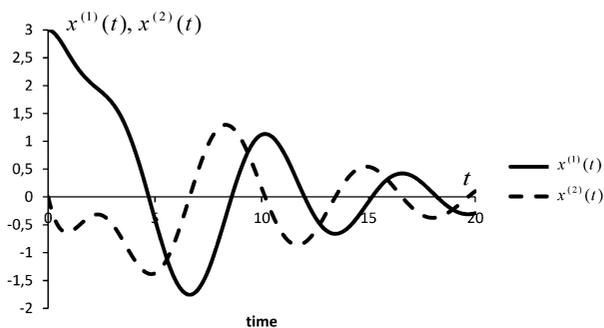


Figure 1.6. Solution of the model for initial conditions  $x^{(1)}(0) = 3, x^{(2)}(0) = 0$ .

Another view on the solution is given on the so called *phase plane* (see Figure 1.7) where axis stand for state variables.

Note that there is no time axis. Current state of the system is represented by a point and trajectory of this point represents systems behavior. Trajectory may be illustrated as a motion of a point or as a trace of this point (which is the case in Figure 1.7). We observe, that oscillation of a mass-spring-damper system driven by force  $u(t) = e^{-0.1t} \sin t$  tends to vanish.

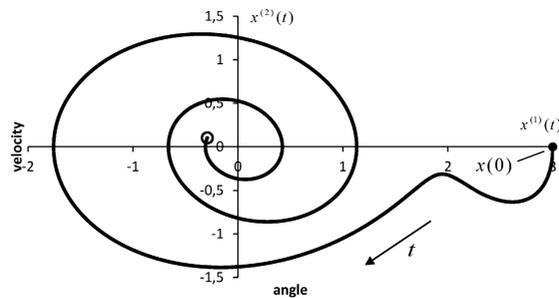


Figure 1.7. Solution of the model in the  $(x^{(1)}, x^{(2)})$  phase plane.

### Exercises

**Exercise 1.4.** Simulate forced mass-spring-damper system in SIMULINK for different initial conditions and for different input signals: step, impulse, constant, harmonic. Draw solution in the time domain and the phase plane. Could you decide when the system is stable or unstable?

**Exercise 1.5.** Simulate the so called Lotka-Volterra system describing relation between numbers of two populations individuals: preys ( $x^{(1)}$ ) and predators ( $x^{(2)}$ ). It is assumed that the number of preys grows exponentially and is reduced proportionally to the number of predators. Moreover, the predators growth factor is proportional to the number of preys. State vector description of the Lotka-Volterra system has the form:

$$\begin{aligned} \frac{dx^{(1)}(t)}{dt} &= (a_0 - a_1 x^{(2)}(t))x^{(1)}(t) \\ \frac{dx^{(2)}(t)}{dt} &= (b_1 x^{(1)}(t) - b_0)x^{(2)}(t) \end{aligned}$$

where  $a_0, a_1, b_0, b_1$  are constant parameters, characteristic for investigated pair of population. Try out different initial conditions. Work out SIMULINK model of the Lotka-Volterra system.

For different values of parameters and initial conditions analyze its behavior on the time domain and the phase plane. Does the system have any equilibrium states?

**Exercise 1.6.** Simulate the human heart dynamics using the Zeeman model of the form:

$$\begin{aligned}\frac{dx^{(1)}(t)}{dt} &= ax^{(1)}(t) - x^{(2)}(t) - (x^{(1)}(t))^3, \\ \frac{dx^{(2)}(t)}{dt} &= x^{(1)}(t) - b\end{aligned}$$

where  $x^{(1)}$  denotes length of heart's muscle fibers,  $x^{(2)}$  is electrochemical stimulus' strength,  $a$  stands for effect of blood pressure and parameter  $b$  determines position of the equilibrium point. Implement Zeeman model in SIMULINK. Analyze its behavior in the time domain and on the phase plane, trying different initial conditions. Can you observe the so called limit cycle? Simulate the system for different values of parameter  $a$ :  $a \leq 0$ , small value of  $a > 0$ , high value of  $a > 0$ , and different values of parameter  $b$ :  $b \leq \sqrt{a/3}$  and  $b > \sqrt{a/3}$ . On the  $(a, b)$  parameters plane draw areas where the system is stable and unstable.

**Exercise 1.7.** A simplified model of the weather behavior proposed by Lorenz has the form:

$$\begin{aligned}\frac{dx^{(1)}(t)}{dt} &= \sigma(x^{(2)}(t) - x^{(1)}(t)) \\ \frac{dx^{(2)}(t)}{dt} &= \rho x^{(1)}(t) - x^{(2)} - x^{(1)}x^{(3)}, \\ \frac{dx^{(3)}(t)}{dt} &= x^{(1)}x^{(2)} - \beta x^{(3)}(t)\end{aligned}$$

where  $x^{(1)}$  is velocity of the air ascendance,  $x^{(2)}$  is the difference between ascendant and descendant air columns temperatures,  $x^{(3)}$  is a measure of deviation of  $x^{(2)}(t)$  over time from linearity,  $\sigma, \rho, \beta$  are constant parameters. Work out SIMULINK model of the Lorenz system taking:  $\sigma = 10, \rho = 28, \beta = 8/3$ . Analyze its behavior in the time domain and on the phase plane. How to visualize three dimensional trajectory covered by the system? Compare two trajectories with slightly different initial positions. Do they stay close to each other as long as time goes by? Try also to analyze the system with  $\rho \in [145, 170]$ .

**Exercise 1.8.** Work out state vector description for inverted pendulum on a cart (see Figure 1.8), described by following set of differential equations:

$$(M + m)\frac{d^2x(t)}{dt^2} - ml\frac{d^2\theta(t)}{dt^2} + ml\left(\frac{d\theta(t)}{dt}\right)^2 \sin\theta(t) = F(t) , \quad (1.17)$$

$$\frac{d^2\theta(t)}{dt^2} = \frac{g}{l}\sin\theta(t) + \frac{1}{l}\frac{d^2x(t)}{dt^2}\cos\theta(t)$$

where  $x$  – position coordinate along the road,  $\theta$  – angular displacement,  $F$  – force driving the cart,  $m$  – mass of the block,  $M$  – mass of the cart,  $l$  – length of the wire,  $g = 9.81$  – gravity constant.

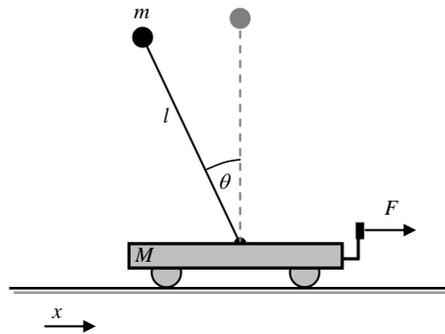


Figure 1.8. Inverted pendulum on a cart.

Work out SIMULINK model of the system and analyze its behavior in time domain and on the phase plane. Try out different values of parameters and different functions  $F(t)$  describing driving force.

### 1.3. Transfer Function

**The Laplace transform** relates functions  $f(t)$  of real variable  $t$  with functions  $F(s)$  of complex variable  $s$ . Laplace transform is defined as:

$$F(s) = \mathcal{L}[f(t)] \stackrel{\text{df}}{=} \int_0^{\infty} f(t) \cdot e^{-st} dt , \quad (1.18)$$

where  $f(t) = 0$  for  $t = 0$ .

Inverse Laplace transform is defined as:

$$f(t) = \mathcal{L}^{-1}[F(s)] \stackrel{\text{df}}{=} \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} F(s) \cdot e^{st} ds. \quad (1.19)$$

**Example 1.5 – Working out Laplace transform using definition.** Let us evaluate Laplace transform for function  $f(t) = e^{at}$ , where  $a \in \mathcal{R}$  is constant parameter. Following definition (1.19), we obtain:

$$\begin{aligned} \mathcal{L}[e^{at}] &= \int_0^{\infty} e^{at} \cdot e^{-st} dt = \int_0^{\infty} e^{at-st} dt = \int_0^{\infty} e^{-t(s-a)} dt = \int_0^{\infty} e^{-t(s-a)} dt = -\frac{1}{s-a} e^{-t(s-a)} \Big|_0^{\infty} = \\ &= -\frac{1}{s-a} \left( e^{-t(s-a)} \Big|_0^{\infty} \right) = -\frac{1}{s-a} \left( e^{-\infty(s-a)} - e^{0(s-a)} \right). \end{aligned}$$

Solution exists only if real part of  $(s-a)$  is greater than 0, which holds if  $\text{Re } s > a$ . Then we have:

$$-\frac{1}{s-a} \left( e^{-\infty(s-a)} - e^{0(s-a)} \right) = -\frac{1}{s-a} (0-1) = \frac{1}{s-a}.$$

Eventually we obtain:

$$\mathcal{L}[e^{at}] = \frac{1}{s-a}.$$

For functions commonly used in systems analysis mathematicians worked out Laplace transform and put results in tables, such as Table 1.1. Instead of performing integration, we may read out the transform directly from the table. Note, that we may also read the table backwards, i.e. for given function  $F(s)$  the table gives original function  $f(t)$ . The question arises: how to work out Laplace transform of more sophisticated functions, that are not found in the Laplace transforms table? In order to do this, properties of Laplace transform may be applied.

**Laplace transform properties.** Laplace transform has following properties:

**Property 1.** Linearity:  $\mathcal{L}[a_1 f_1(t) + a_2 f_2(t)] = a_1 \mathcal{L}[f_1(t)] + a_2 \mathcal{L}[f_2(t)] = a_1 F_1(s) + a_2 F_2(s),$

where  $a_1, a_2 \in \mathcal{R}$  are constant parameters.

Linearity property is a direct consequence of integral property.

**Table 1.1.** Table of commonly used Laplace transforms.

$f(t)$	$F(s)$	$f(t)$	$F(s)$
$\delta(t)$	$1(s)$	$t \sin at$	$\frac{2as}{(s^2 + a^2)^2}$
$1(t)$	$\frac{1}{s}$	$t \cos at$	$\frac{s^2 - a^2}{(s^2 + a^2)^2}$
$e^{at}$	$\frac{1}{s - a}$	$e^{at} \sin bt$	$\frac{b}{(s - a)^2 + b^2}$
$\sin at$	$\frac{a}{s^2 + a^2}$	$e^{at} \cos bt$	$\frac{s - a}{(s - a)^2 + b^2}$
$\cos at$	$\frac{s}{s^2 + a^2}$	$\frac{t^n}{n!}, n \in \mathcal{N}$	$\frac{1}{s^{n+1}}$
		$e^{at} \frac{t^n}{n!}, n \in \mathcal{N}$	$\frac{1}{(s - a)^{n+1}}$

**Property 2.** Differentiation:  $\mathcal{L}[f^{(n)}(t)] = s^n F(s) - s^{n-1} f(0) - s^{n-2} f'(0) - \dots - f^{(n-1)}(0)$ ,

where  $f^{(n)}(t)$  is the  $n$ -th order derivative of function  $f(t)$  and  $f(0), f'(0), \dots, f^{(n-1)}(0)$  are initial values.

Note, that when initial values are all equal to zero, we have simple rule to remember:

$$\mathcal{L}[f^{(n)}(t)] = s^n F(s).$$

Now, we prove the differentiation property in the case  $n = 1$ :  $\mathcal{L}[f'(t)] = sF(s) - f(0)$ .

Following the Laplace transform definition (1.18), we may write:  $\mathcal{L}[f'(t)] = \int_0^{\infty} f'(t)e^{-st} dt$ . By making

use of integration by parts, where  $h(t) = e^{-st}$ ,  $h'(t) = -se^{-st}$ , we obtain:

$$\int_0^{\infty} f'(t)e^{-st} dt = f(t)h(t)\Big|_0^{\infty} - \int_0^{\infty} f(t)h'(t) dt = f(t)e^{-st}\Big|_0^{\infty} - \int_0^{\infty} f(t)(-se^{-st}) dt.$$

Assuming  $\text{Re } s > 0$  gives:  $f(t)e^{-s\infty} - f(0)e^{-s \cdot 0} + s \int_0^{\infty} f(t)e^{-st} dt = sF(s) - f(0).$

**Property 3.** Integration:  $\mathcal{L}\left[\int_0^t f(\tau) d\tau\right] = \frac{1}{s} F(s).$

Recall, that in SIMULINK the integration block is labeled by  $\frac{1}{s}.$

**Property 4.** Multiplication of a function by  $t$ :  $\mathcal{L}[tf(t)] = -\frac{d}{ds} F(s).$

Now, we present short proof of the property above:

$$-\frac{d}{ds} \mathcal{L}[f(t)] = -\frac{d}{ds} \int_0^{\infty} f(t)e^{-st} dt = -\int_0^{\infty} \frac{d}{ds} e^{-st} f(t) dt = \int_0^{\infty} te^{-st} f(t) dt = \int_0^{\infty} tf(t)e^{-st} dt = \mathcal{L}[tf(t)].$$

**Property 5.** Division of a function by  $t$ :  $\mathcal{L}\left[\frac{f(t)}{t}\right] = \int_s^{\infty} F(s) ds.$

Now, we present short proof of the property above:

$$\begin{aligned} \int_s^{\infty} \mathcal{L}[f(t)] ds &= \int_s^{\infty} \left( \int_0^{\infty} f(t)e^{-st} dt \right) ds = \int_0^{\infty} f(t) \left( \int_s^{\infty} e^{-st} ds \right) dt = \int_0^{\infty} f(t) \left( -\frac{1}{t} \left( e^{-st} \Big|_s^{\infty} \right) \right) dt = \int_0^{\infty} f(t) \left( 0 + \frac{1}{t} e^{-st} \right) dt = \\ &= \int_0^{\infty} \frac{f(t)}{t} e^{-st} dt = \mathcal{L}\left[\frac{f(t)}{t}\right]. \end{aligned}$$

**Property 6.** Multiplication of a function by  $e^{at}$ :  $\mathcal{L}[e^{at} f(t)] = F(s-a).$

**Property 7.** Change of time scale:  $\mathcal{L}[f(at)] = \frac{1}{a} F\left(\frac{s}{a}\right),$  where  $a > 0.$

Now, we prove the property above. Following the Laplace transform definition (1.18), we may write:

$$\mathcal{L}[f(at)] = \int_0^{\infty} f(at)e^{-st} dt. \text{ Taking } x = at \text{ we obtain: } \int_0^{\infty} f(x)e^{-\frac{s}{a}x} \frac{dx}{a} = \frac{1}{a} \int_0^{\infty} f(x)e^{-\frac{s}{a}x} dx = \frac{1}{a} F\left(\frac{s}{a}\right).$$

**Property 8.** Delay:  $\mathcal{L}[f(t-t_0) \cdot 1(t-t_0)] = e^{-st_0} F(s)$ .

Now, we prove the delay property. Following the Laplace transform definition (1.18), we may write:

$$\mathcal{L}[f(t-t_0) \cdot 1(t-t_0)] = \int_0^{\infty} f(t-t_0) \cdot 1(t-t_0) e^{-st} dt = \int_{t_0}^{\infty} f(t-t_0) e^{-st} dt. \text{ Taking } \tau = t-t_0 \text{ we obtain:}$$

$$\int_0^{\infty} f(\tau) e^{-s(\tau+t_0)} d\tau = \int_0^{\infty} f(\tau) e^{-s\tau} e^{-st_0} d\tau = e^{-st_0} \int_0^{\infty} f(\tau) e^{-s\tau} d\tau = e^{-st_0} \mathcal{L}[f(t)] = e^{-st_0} F(s).$$

**Property 9.** Convolution:

$$\mathcal{L}[f_1(t) * f_2(t)] = \mathcal{L}\left[\int_0^t f_1(t-\tau) f_2(\tau) d\tau\right] = \mathcal{L}[f_1(t)] \mathcal{L}[f_2(t)] = F_1(s) F_2(s),$$

where  $f_1(t) * f_2(t)$  is convolution of functions  $f_1(t)$  and  $f_2(t)$ .

Now, we prove the convolution property. At the beginning note, that  $f_1(t-\tau)1(t-\tau) = 0$  holds for  $\tau > t$ , which was assumed while defining Laplace transform (1.18). Following the Laplace transform

definition (1.18), we may write:  $\int_0^t f_1(t-\tau) f_2(\tau) d\tau = \int_0^{\infty} f_1(t-\tau) 1(t-\tau) f_2(\tau) d\tau$ . Then:

$$\mathcal{L}\left[\int_0^t f_1(t-\tau) f_2(\tau) d\tau\right] = \mathcal{L}\left[\int_0^{\infty} f_1(t-\tau) 1(t-\tau) f_2(\tau) d\tau\right] = \int_0^{\infty} e^{-st} \left[\int_0^{\infty} f_1(t-\tau) 1(t-\tau) f_2(\tau) d\tau\right] dt.$$

Substituting  $t-\tau = \lambda$  and changing the order of integration gives:

$$\begin{aligned} \mathcal{L}\left[\int_0^t f_1(t-\tau) f_2(\tau) d\tau\right] &= \int_0^{\infty} \int_0^{\infty} f_1(t-\tau) 1(t-\tau) f_2(\tau) e^{-st} dt d\tau = \int_0^{\infty} f_1(\lambda) e^{-s(\lambda+\tau)} d\lambda \int_0^{\infty} f_2(\tau) d\tau = \\ &= \int_0^{\infty} f_1(\lambda) e^{-s\lambda} d\lambda \int_0^{\infty} f_2(\tau) e^{-s\tau} d\tau = \mathcal{L}[f_1(t)] \mathcal{L}[f_2(t)] = F_1(s) F_2(s). \end{aligned}$$

**Example 1.6 – Using table of Laplace transforms.** Work out Laplace transform of function

$$f(t) = 3e^{-2t} + 5\sin 2t - 4.$$

With use of linearity property no. 1 together with table 1, we write:

$$\mathcal{L}[3e^{-2t} + 5\sin 2t - 4] = 3\mathcal{L}[e^{-2t}] + 5\mathcal{L}[\sin 2t] - 4\mathcal{L}[1] = 3\frac{1}{s+2} + 5\frac{2}{s^2+4} - 4\frac{1}{s} = \frac{3}{s+2} + \frac{10}{s^2+4} - \frac{4}{s}.$$

Let us consider linear system with zero initial conditions, where dynamic relation between input  $u(t)$  and output  $y(t)$  is described by differential equation (1.12):

$$\begin{aligned} \frac{d^m y(t)}{dt^m} + a_{m-1} \frac{d^{m-1} y(t)}{dt^{m-1}} + \dots + a_1 \frac{dy(t)}{dt} + a_0 y(t) &= \\ = b_v \frac{d^v u(t)}{dt^v} + b_{v-1} \frac{d^{v-1} u(t)}{dt^{v-1}} + \dots + b_1 \frac{du(t)}{dt} + b_0 u(t). \end{aligned} \quad (1.20)$$

To transform both sides of this equation:

$$\mathcal{L}[y^{(m)}(t) + a_{m-1}y^{(m-1)}(t) + \dots + a_1 y'(t) + a_0 y(t)] = \mathcal{L}[b_v u^{(v)}(t) + b_{v-1} u^{(v-1)}(t) + \dots + b_1 u'(t) + b_0 u(t)].$$

Applying linearity property yields:

$$\begin{aligned} \mathcal{L}[y^{(m)}(t)] + a_{m-1}\mathcal{L}[y^{(m-1)}(t)] + \dots + a_1\mathcal{L}[y'(t)] + a_0\mathcal{L}[y(t)] &= \\ = b_v\mathcal{L}[u^{(v)}(t)] + b_{v-1}\mathcal{L}[u^{(v-1)}(t)] + \dots + b_1\mathcal{L}[u'(t)] + b_0\mathcal{L}[u(t)] \end{aligned}$$

Applying differentiation property no. 2, with zero initial values assumed, gives:

$$s^m Y(s) + a_{m-1} s^{m-1} Y(s) + \dots + a_1 s Y(s) + a_0 Y(s) = b_v s^v U(s) + b_{v-1} s^{v-1} U(s) + \dots + b_1 s U(s) + b_0 U(s).$$

Pulling  $Y(s)$  and  $U(s)$  before the bracket gives:

$$Y(s)(s^m + a_{m-1}s^{m-1} + \dots + a_1s + a_0) = U(s)(b_v s^v + b_{v-1}s^{v-1} + \dots + b_1s + b_0).$$

**The transfer function**  $K(s)$  of linear system with zero initial conditions is the ratio between Laplace transform of the system output and the system input:

$$K(s) \stackrel{df}{=} \frac{Y(s)}{U(s)} = \frac{b_v s^v + b_{v-1} s^{v-1} + \dots + b_1 s + b_0}{s^m + a_{m-1} s^{m-1} + \dots + a_1 s + a_0} = \frac{\sum_{i=0}^v b_i s^i}{s^m + \sum_{i=0}^{m-1} a_i s^i}. \quad (1.21)$$

**Example 1.7 – Evaluation of transfer function from differential equation.** A linear system is described by differential equation:

$$\frac{d^2 y(t)}{dt^2} + 3 \frac{dy(t)}{dt} + 2y(t) = \frac{du(t)}{dt} + 2u(t). \quad (1.22)$$

Passing both sides by Laplace transform gives:

$$s^2 Y(s) + 3s Y(s) + 2Y(s) = s U(s) + 2U(s).$$

Pulling  $Y(s)$  and  $U(s)$  before the bracket gives:

$$Y(s)(s^2 + 3s + 2) = U(s)(s + 2).$$

Rearranging above equation gives transfer function:

$$\frac{Y(s)}{U(s)} = \frac{s + 2}{s^2 + 3s + 2} = K(s). \quad (1.23)$$

Note, that description (1.23) makes sense for the system (1.22) only for zero initial conditions.

**Example 1.8 – Evaluation of differential equation from transfer function.** A linear system is described by following transfer function:

$$K(s) = \frac{1}{s + 1}. \quad (1.24)$$

We ask about differential equation describing the system:

$$\frac{Y(s)}{U(s)} = \frac{1}{s + 1}.$$

Cross multiplication gives:

$$Y(s)(s + 1) = U(s),$$

which is

$$sY(s) + Y(s) = U(s).$$

Now we may deduce original difference equation:

$$\frac{dy(t)}{dt} + y(t) = u(t). \quad (1.25)$$

Let us take a closer look at solution (1.23) of example 1.7:

$$K(s) = \frac{s + 2}{s^2 + 3s + 2} = \frac{s + 2}{(s + 2)(s + 1)} = \frac{1}{s + 1}.$$

We observe that two different differential equations (1.22) and (1.25) have the same transfer function. Transfer function loses some information about the system. Differential equation is more accurate description, however, transfer function is more convenient. When dealing with complex systems, that are composed of many input-output dynamic elements, it is easy to work out the complex system model on the basis of its elements transfer functions.

**Example 1.9 – SIMULINK model.** Consider a system described by transfer function (1.23):

$K(s) = \frac{s + 2}{s^2 + 3s + 2}$ . SIMULINK implementation of the system is illustrated in Figure 1.9. Two vector of

parameters were set up in the 'Transfer Fcn' block parameters window. The first one, [1 2], is for nominator and the second one is for denominator [1 3 2]. As an input signal, Dirac delta  $\delta(t)$  is applied. For such an input signal the output signal is called *impulse response* (see Figure 1.10).

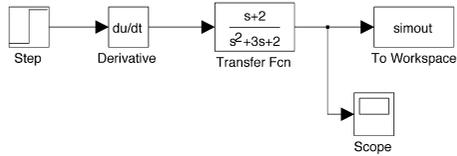


Figure 1.9. SIMULINK implementation of a system described by transfer function (1.23).

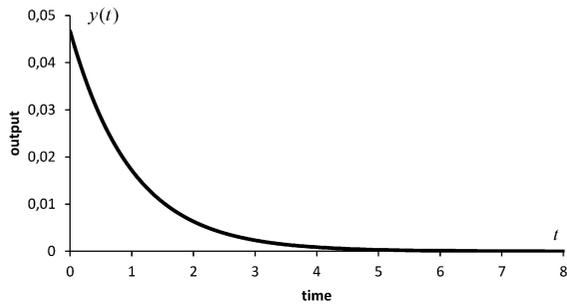


Figure 1.10. Impulse response of a system described by transfer function (1.23).

### Exercises

**Exercise 1.9.** Find Laplace transforms of the following functions:

a)  $f(t) = \cos 3t - 2$

d)  $f(t) = e^{-2t} \sin t + \delta(t)$

b)  $f(t) = \frac{1}{3}(\cos 2t + t \sin t)$

e)  $f(t) = 3t^2 e^{-3t}$

c)  $f(t) = \frac{e^{4t} - e^t}{2t}$

f)  $f(t) = 2t^2 - t^3 e^{-2t}$

**Exercise 1.10.** Proof: – linearity, – differential, – integration, – multiplication of a function by  $e^{at}$ , properties of Laplace transform.

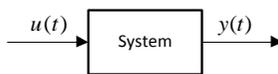
**Exercise 1.11.** Linear system fed up with step function  $u(t) = 1(t)$ , returned the output signal  $y(t) = e^{-t}$ . Work out the system's transfer function  $K(s)$ .

**Exercise 1.12.** Work out transfer function of forced mass-spring-damper system described by differential equation (1.10).

**Exercise 1.13.** With use of SIMULINK implement models of systems described by transfer functions (1.23), (1.24) and differential equations (1.22), (1.25). Compare their responses for different input signals. What do you observe?

## 1.4. Linear Systems Analysis

In general, we treat the system as a black box with input  $u(t)$  and output  $y(t)$  (see Figure 1.11). Dynamic relation between input and output signals may be expressed in terms of state vector description, differential equation or transfer function.



**Figure 1.11.** An input-output system.

The task of system analysis is to find output signal  $y(t)$  (system's behavior) for a given input signal  $u(t)$  (describing our activity upon the system). Having differential equation describing system's dynamics and initial conditions determining system's state at the beginning of the dynamic process, we ask about an unknown function  $y(t)$ , which is solution of differential equation. Thus, the task of system analysis reduces to solving differential equations describing the process.

The procedure of solving differential equations using Laplace transform involves the following steps (Bro06):

*Step 1.* Pass both sides of differential equation through Laplace transform. As a result, algebraic equation of complex variable is obtained.

*Step 2.* Rearrange algebraic equation in order to find formula for transform  $Y(s)$  of unknown function  $y(t)$ . As a result, rational function of complex variable  $s$  is obtained.

*Step 3.* If needed, apply partial fraction decomposition of rational function worked out in the previous step.

*Step 4.* Apply inverse Laplace transform in order to find solution  $y(t)$ .

The procedure described above applies for linear differential equations only.

**Example 1.10 – Solving linear differential equation.**

Let us consider an input-output system described by following differential equation:

$$\frac{d^2 y(t)}{dt^2} + y(t) = u(t).$$

Let us analyze the systems' behavior activated by input signal  $u(t) = \cos 2t$  from zero initial

conditions:  $y(0) = \frac{dy(0)}{dt} = 0$ . What does the output signal  $y(t)$  look like?

The problem reduces to solving unknown function  $y(t)$  from differential equation:

$$\frac{d^2 y(t)}{dt^2} + y(t) = \cos 2t. \quad (1.26)$$

At first, we pass both sides of the equation above by Laplace transform:

$$\mathcal{L}\left[\frac{d^2 y(t)}{dt^2} + y(t)\right] = \mathcal{L}[\cos 2t].$$

Applying linearity property yields:

$$\mathcal{L}\left[\frac{d^2 y(t)}{dt^2}\right] + \mathcal{L}[y(t)] = \mathcal{L}[\cos 2t].$$

Using differentiation property for the left side and Laplace transform table for the right side gives algebraic equation:

$$s^2\left(Y(s) - sy(0) - \frac{dy(0)}{dt}\right) + Y(s) = \frac{s}{s^2 + 2^2}.$$

Since all initial conditions are zero, we have:

$$s^2 Y(s) + Y(s) = \frac{s}{s^2 + 2^2}.$$

Taking  $Y(s)$  before the bracket gives:

$$Y(s)(s^2 + 1) = \frac{s}{s^2 + 4}.$$

We obtain rational function:

$$Y(s) = \frac{s}{(s^2 + 4)(s^2 + 1)},$$

which partial fraction decomposition has the following form:

$$Y(s) = \frac{s}{(s^2 + 4)(s^2 + 1)} = \frac{As + B}{s^2 + 4} + \frac{Cs + D}{s^2 + 1}, \quad (1.27)$$

where  $A, B, C, D$  are coefficients to be determined in such a way, that equation above holds true. To determine values of these coefficients, let us multiple both sides by denominator  $(s^2 + 4)(s^2 + 1)$ :

$$s = (As + B)(s^2 + 1) + (Cs + D)(s^2 + 4).$$

Brackets multiplication gives:

$$s = As^3 + As + Bs^2 + B + Cs^3 + 4Cs + Ds^2 + 4D.$$

After elements of the same order are joined:

$$s = (A + C)s^3 + (B + D)s^2 + (A + 4C)s + (B + 4D).$$

The equation above holds still if and only if the system of equations:

$$\begin{cases} A + C = 0 & \text{(a)} \\ B + D = 0 & \text{(b)} \\ A + 4C = 1 & \text{(c)} \\ B + 4D = 0 & \text{(d)} \end{cases}$$

holds true. By subtraction of equation (a) from (c) and equation (b) from (d):

$$\begin{cases} 3C = 1 & \text{(c) - (a)} \\ 3D = 0 & \text{(d) - (b)} \end{cases}$$

we find that  $D = 0$  and  $C = 1/3$ . Then, from (b) we find that  $B = 0$  and from (a) we have  $A = -1/3$ . Going back to expression (1.27), we may write its form after partial fraction decomposition:

$$Y(s) = \frac{s}{(s^2 + 4)(s^2 + 1)} = -\frac{1}{3} \frac{s}{s^2 + 4} + \frac{1}{3} \frac{s}{s^2 + 1}.$$

Now, it is easy to apply inverse Laplace transform:

$$y(t) = \mathcal{L}^{-1}[Y(s)] = \mathcal{L}^{-1}\left[-\frac{1}{3} \frac{s}{s^2 + 4} + \frac{1}{3} \frac{s}{s^2 + 1}\right] = -\frac{1}{3} \mathcal{L}^{-1}\left[\frac{s}{s^2 + 2^2}\right] + \frac{1}{3} \mathcal{L}^{-1}\left[\frac{s}{s^2 + 1}\right].$$

Eventually, using Laplace transforms table backwards, we obtain solution of differential equation (1.26):

$$y(t) = -\frac{1}{3} \cos 2t + \frac{1}{3} \cos t. \quad (1.28)$$

As a result of analysis we found that considered system, driven by the signal  $u(t) = \cos 2t$  from the zero state, responds with the signal  $y(t) = -1/3 \cos 2t + 1/3 \cos t$ .

## Exercises

**Exercise 1.14.** Work out numerical solution of differential equation (1.26) and compare it with the exact one (1.28).

**Exercise 1.15.** Solve following differential equations:

a)  $\frac{d^2 y(t)}{dt^2} + 4y(t) = \sin t,$

zero initial conditions

b)  $\frac{d^4 y(t)}{dt^4} + \frac{d^3 y(t)}{dt^3} = \cos t,$

zero initial conditions

c)  $2 \frac{dy(t)}{dt} + y(t) = t^2 + 2,$

$y(0) = 4$

d)  $\frac{d^2 y(t)}{dt^2} + 4y(t) = 3 \sin t + 10 \cos 3t,$

$y(0) = 2, \frac{dy(0)}{dt} = 3$

e)  $\frac{d^2 y(t)}{dt^2} + 2 \frac{dy(t)}{dt} + 5y(t) = 3e^{-2t},$

$y(0) = 1, \frac{dy(0)}{dt} = 1$

f)  $\frac{d^2 y(t)}{dt^2} - \frac{dy(t)}{dt} - 4y(t) + 3 = \frac{du(t)}{dt} - u(t)$

zero initial conditions

g)  $\frac{d^3 y(t)}{dt^3} + 3 \frac{d^2 y(t)}{dt^2} + 3 \frac{dy(t)}{dt} + y(t) = 6e^{-t}$

zero initial conditions

h)  $\frac{d^2 y(t)}{dt^2} + 9 \frac{dy(t)}{dt} + 14y(t) = \frac{1}{2} \sin t,$

$y(0) = 0, \frac{dy(0)}{dt} = 1$

i)  $\frac{d^2 y(t)}{dt^2} - 3 \frac{dy(t)}{dt} + 2y(t) = 4e^{3t},$

$y(0) = 2, \frac{dy(0)}{dt} = 6$

j)  $\frac{d^3 y(t)}{dt^3} - 3 \frac{d^2 y(t)}{dt^2} + 3 \frac{dy(t)}{dt} - y(t) = t^2 e^t,$

$y(0) = 1, \frac{dy(0)}{dt} = 2, \frac{d^2 y(0)}{dt^2} = 3$

k)  $\frac{d^2 y(t)}{dt^2} - y(t) = e^{-t}(t^2 - 3t + 3),$

$y(0) = 1, \frac{dy(0)}{dt} = -1$

**Exercise 1.16.** Try to solve differential equation  $\frac{d^2 y(t)}{dt^2} + y(t) = \frac{d^3 u(t)}{dt^3}$  for zero initial conditions. Use analytical method and numerical (in SIMULINK). Do you find anything wrong about this equation?

**Exercise 1.17.** Work out step response  $u(t) = 1(t)$  of linear system described by the transfer function

$$K(s) = \frac{1}{s+1}.$$

**Exercise 1.18.** Work out impulse and step responses of a system described by differential equation:

$$\frac{d^2y(t)}{dt^2} + 2\frac{dy(t)}{dt} + y(t) = u(t).$$

### 1.5. Nonlinear Systems Analysis

Analysis of nonlinear systems is more challenging problem, due to limited possibilities of solving nonlinear differential equations. There are some analytical methods, but for most nonlinear equations that arise in practice only numerical solutions may be obtained. However, it is possible to work out linear approximation of nonlinear system around a given state (usually called *set point*). Usually set point is chosen as an equilibrium state of interest. Results of linearized system analysis also apply for original nonlinear system in the neighborhood of a set point. For example, the stability of nonlinear system may be stated on the basis of its linear approximation around a set point of interest. Control algorithms may be designed for linearized systems and they should perform well as long as the systems' state stays close to the set point. If the systems goes far away from the chosen set point, it should be linearized once again for different set point.

One of the simplest linearization procedure (Vuk03), based on the Taylor's expansion, is given below.

Let us consider nonlinear system with the state vector description:

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = F(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(t) = G(\mathbf{x}(t), \mathbf{u}(t)) \end{cases} \quad (1.29)$$

Nonlinear functions  $F, G$  are vector functions of the form:

$$F(\mathbf{x}(t), \mathbf{u}(t)) = \begin{bmatrix} f_1(\mathbf{x}(t), \mathbf{u}(t)) \\ f_2(\mathbf{x}(t), \mathbf{u}(t)) \\ \vdots \\ f_R(\mathbf{x}(t), \mathbf{u}(t)) \end{bmatrix}, \quad G(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} g_1(\mathbf{x}(t), \mathbf{u}(t)) \\ g_2(\mathbf{x}(t), \mathbf{u}(t)) \\ \vdots \\ g_R(\mathbf{x}(t), \mathbf{u}(t)) \end{bmatrix}.$$

Let us denote a set point of interest by:

$$\begin{bmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \bar{x}^{(1)} \\ \vdots \\ \bar{x}^{(R)} \\ \bar{u}^{(1)} \\ \vdots \\ \bar{u}^{(S)} \end{bmatrix}.$$

Linear approximation of the system (1.29) has the general form:

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases}, \quad (1.30)$$

where  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$  are matrices containing constant number determined using the following definitions:

$$\mathbf{A} = \left. \frac{\partial F}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}} = \begin{bmatrix} \frac{\partial f_1}{\partial x^{(1)}} & \dots & \frac{\partial f_1}{\partial x^{(R)}} \\ \dots & \dots & \dots \\ \frac{\partial f_R}{\partial x^{(1)}} & \dots & \frac{\partial f_R}{\partial x^{(R)}} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}, \quad (1.31)$$

$$\mathbf{B} = \left. \frac{\partial F}{\partial \mathbf{u}} \right|_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}} = \begin{bmatrix} \frac{\partial f_1}{\partial u^{(1)}} & \dots & \frac{\partial f_1}{\partial u^{(S)}} \\ \dots & \dots & \dots \\ \frac{\partial f_R}{\partial u^{(1)}} & \dots & \frac{\partial f_R}{\partial u^{(S)}} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}, \quad (1.32)$$

$$\mathbf{C} = \left. \frac{\partial G}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}} = \begin{bmatrix} \frac{\partial g_1}{\partial x^{(1)}} & \dots & \frac{\partial g_1}{\partial x^{(R)}} \\ \dots & \dots & \dots \\ \frac{\partial g_L}{\partial x^{(1)}} & \dots & \frac{\partial g_L}{\partial x^{(R)}} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}, \quad (1.33)$$

$$\mathbf{D} = \left. \frac{\partial G}{\partial \mathbf{u}} \right|_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}} = \begin{bmatrix} \frac{\partial g_1}{\partial u^{(1)}} & \dots & \frac{\partial g_1}{\partial u^{(S)}} \\ \dots & \dots & \dots \\ \frac{\partial g_L}{\partial u^{(1)}} & \dots & \frac{\partial g_L}{\partial u^{(S)}} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}. \quad (1.34)$$

**Example 1.11 – Linearization of state vector description.**

Let us take a system described by the following state vector description:

$$\begin{cases} \frac{dx^{(1)}(t)}{dt} = 3x^{(2)}(t) + x^{(1)}(t)x^{(2)}(t) + (x^{(2)}(t))^2 - (u(t))^2 \\ \frac{dx^{(2)}(t)}{dt} = -2x^{(1)}(t) + e^{-x^{(2)}(t)} \sin x^{(1)}(t) + 2u(t) \\ y(t) = x^{(1)}(t) + u(t) \end{cases} \quad (1.35)$$

State vector for the system above has two dimensions:

$$\mathbf{x}(t) = \begin{bmatrix} x^{(1)}(t) \\ x^{(2)}(t) \end{bmatrix}, \text{ and } \mathbf{u}(t) = [u(t)].$$

Functions  $F, G$  also have two dimensions:

$$F(\mathbf{x}(t), \mathbf{u}(t)) = \begin{bmatrix} f_1(\mathbf{x}(t), \mathbf{u}(t)) \\ f_2(\mathbf{x}(t), \mathbf{u}(t)) \end{bmatrix}, \quad G(\mathbf{x}(t), \mathbf{u}(t)) = [g(\mathbf{x}(t), \mathbf{u}(t))],$$

where functions  $f_1, f_2, g$  are:

$$\begin{aligned} f_1(\mathbf{x}(t), \mathbf{u}(t)) &= 3x^{(2)}(t) + x^{(1)}(t)x^{(2)}(t) + (x^{(2)}(t))^2 - (u(t))^2 \\ f_2(\mathbf{x}(t), \mathbf{u}(t)) &= -2x^{(1)}(t) + e^{-x^{(2)}(t)} \sin x^{(1)}(t) + 2u(t) \\ g(\mathbf{x}(t), \mathbf{u}(t)) &= x^{(1)}(t) + u(t) \end{aligned}$$

Matrices introduced in (1.30) have the form given by (1.31)-(1.34):

$$\mathbf{A} = \begin{bmatrix} \frac{\partial f_1}{\partial x^{(1)}} & \frac{\partial f_1}{\partial x^{(2)}} \\ \frac{\partial f_2}{\partial x^{(1)}} & \frac{\partial f_2}{\partial x^{(2)}} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}, \quad \mathbf{B} = \begin{bmatrix} \frac{\partial f_1}{\partial u} \\ \frac{\partial f_2}{\partial u} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}, \quad \mathbf{C} = \begin{bmatrix} \frac{\partial g}{\partial x^{(1)}} & \frac{\partial g}{\partial x^{(2)}} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}, \quad \mathbf{D} = \begin{bmatrix} \frac{\partial g}{\partial u} \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}}. \quad (1.36)$$

Components of these matrices are evaluated as:

$$\begin{aligned} \frac{\partial f_1}{\partial x^{(1)}} &= x^{(2)}u(t), \quad \frac{\partial f_1}{\partial x^{(2)}} = 3 + x^{(1)}(t) + 2x^{(2)}(t), \\ \frac{\partial f_2}{\partial x^{(1)}} &= -2 + e^{-x^{(2)}(t)} \cos x^{(1)}(t), \quad \frac{\partial f_2}{\partial x^{(2)}} = -e^{-x^{(2)}(t)} \sin x^{(1)}(t), \\ \frac{\partial f_1}{\partial u} &= -2u(t), \quad \frac{\partial f_2}{\partial u} = 2, \\ \frac{\partial g}{\partial x^{(1)}} &= 1, \quad \frac{\partial g}{\partial x^{(2)}} = 0, \quad \frac{\partial g}{\partial u} = 1. \end{aligned}$$

Introducing the formulas above to (1.36), we obtain description of linearized systems:

$$\left\{ \begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \begin{bmatrix} x^{(2)}(t) & 3 + x^{(1)}(t) + 2x^{(2)}(t) \\ -2 + e^{-x^{(2)}(t)} \cos x^{(1)}(t) & -e^{-x^{(2)}(t)} \sin x^{(1)}(t) \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}} \mathbf{x}(t) + \begin{bmatrix} -2u(t) \\ 2 \end{bmatrix}_{\substack{\mathbf{x}=\bar{\mathbf{x}} \\ \mathbf{u}=\bar{\mathbf{u}}}} \mathbf{u}(t) \\ \mathbf{y}(t) &= [1 \quad 0] \mathbf{x}(t) + \mathbf{u}(t) \end{aligned} \right. \quad (1.37)$$

As an example, we chose the set point as the zero vector:  $\begin{bmatrix} \bar{x}^{(1)} \\ \bar{x}^{(2)} \\ \bar{u} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ .

For this particular point the system (1.37) has the form:

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = \begin{bmatrix} 0 & 3 \\ -1 & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \mathbf{u}(t) \\ \mathbf{y}(t) = [1 \quad 0] \mathbf{x}(t) + \mathbf{u}(t) \end{cases} .$$

### Exercises

**Exercise 1.19.** Linearize the system below:

$$\begin{cases} \frac{dx^{(1)}(t)}{dt} = \frac{x^{(2)}(t)}{1 + (x^{(1)}(t))^2 + (x^{(2)}(t))^2} - u(t) \sin u(t) \\ \frac{dx^{(2)}(t)}{dt} = ux^{(1)}(t)(1 - x^{(1)}(t)) \\ y(t) = x^{(1)}(t)x^{(2)}(t) + (u(t) - 1)^2 \end{cases}$$

around the zero set point:  $\begin{bmatrix} \bar{x}^{(1)} \\ \bar{x}^{(2)} \\ \bar{u} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$

**Exercise 1.20.**

Linearize the state vector description of inverted pendulum, obtained when solving exercise 1.8 around the unstable equilibrium position (zero angle of displacement and zero angular velocity). When the system is more sensitive to control actions and disturbances: for long or for short pendulum?

## 2. Identification

Identification is the term describing mathematical tools and algorithms to build models of real systems from measured data. These models are significant in many disciplines. They can be used to formulate both new research and engineering problems. The models are useful for system analysis, to get better understanding of the real systems (Bub80, Nel01).

In Figure 2.1., the process of model development and its applications is illustrated. In order to build a model of a real system, the process or phenomenon *investigator* collects measured data, information and observations. Then, applying identification tools and algorithms model of the real system is obtained. It must be stressed that the first results cannot be satisfying for *investigator*. In Figure 2.1, the process of model improvement is indicated in a block *comparison* and *adaptation*.

The real system model can be useful for formulating some new investigation problems, control and management. Other problems, in which models may be helpful, are related to diagnosis or fault detection.

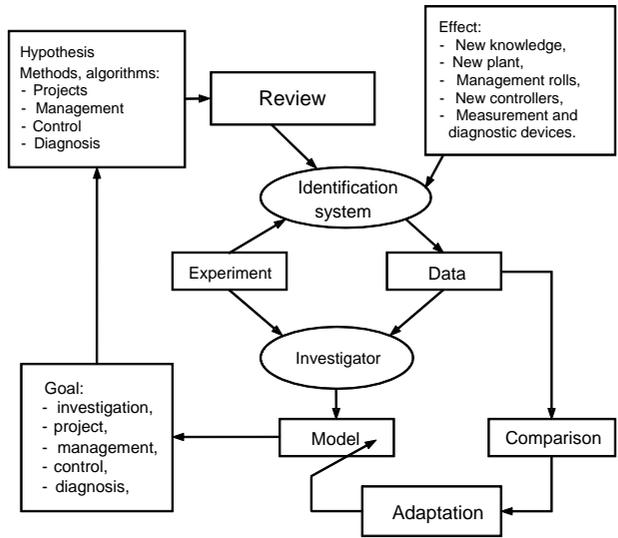


Figure 2.1. The Model in a System Analysis.

As the result of findings mentioned above problems, we have obtained some hypothesis and methods for new control, management and diagnosis algorithms. After verification of acquired outcomes, we have some new knowledge on investigated real systems. Developed control and management algorithms are applied in a design process of new controllers and management systems (Figure 2.1).

## 2.1 Determination of the System Parameters

Let us assume that we have a static system characteristic of the identification system (Figure 2.2) but we do not know its parameters (Bub80).

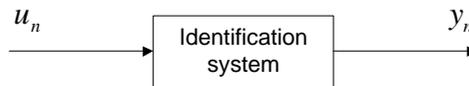


Figure 2.2. The identification system.

### The problem formulation:

The static system characteristic:  $y = F(u, \theta)$ ;

$F$  - known function;

$u$  - input vector  $u \in \mathcal{U} \subseteq \mathcal{R}^S$ ,  $\mathcal{U}$  - input domain;

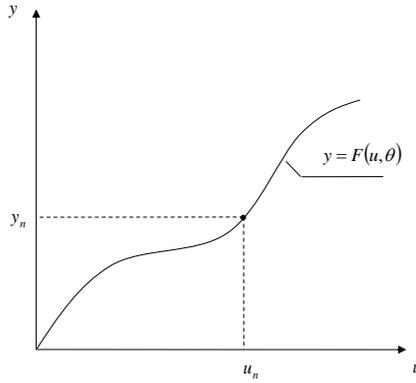
$y$  - output vector  $y \in \mathcal{Y} \subseteq \mathcal{R}^L$ ,  $\mathcal{Y}$  - output domain;

$\theta$  - unknown vector of the system characteristics parameters  $\theta \in \Theta \subseteq \mathcal{R}^R$ ;

$\Theta$  - parameters' domain.

As it was mentioned, in order to determine some parameters of the static system characteristic  $y = F(u, \theta)$ , collected data are needed.

In this case we assume, that we have noise free measurements, which means that these collected data are some selected points in the static system characteristic (Figure 2.4). In this case our problem is to determine parameters of the static system characteristic. Next few examples illustrates main issues connected with this problem.



**Figure 2.3. The static system characteristic.**

**Example 2.1 – Determination of the system parameters.** Let us assume, that we have the static system characteristic described by the following set of equations:

$$\begin{aligned} y^{(1)} &= \theta_1^{(1)} u^{(1)} + \theta_1^{(2)} u^{(2)} \\ y^{(2)} &= \theta_2^{(1)} u^{(1)} + \theta_2^{(2)} u^{(2)}. \end{aligned} \tag{2.1}$$

As a result of an experiment, the following data are collected:

**Table 2.1. Measurements for 1.1.**

n	1	2	3	4	5	6
$u^{(1)}$	1	0	-1	-1	1	2
$u^{(2)}$	1	1	1	-1	0	1
$y^{(1)}$	0	-1	-2	0	1	1
$y^{(2)}$	5	3	1	-5	2	7

**The problem:** To calculate some parameters of the static system characteristic (2.1).

**The solution:**

A given description of the system can be rewritten in the form:

$$y = \theta u, \tag{2.2}$$

where:

$$u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \end{bmatrix}, y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \end{bmatrix}, \theta = \begin{bmatrix} \theta_1^{(1)} & \theta_1^{(2)} \\ \theta_2^{(1)} & \theta_2^{(2)} \end{bmatrix}. \quad (2.3)$$

In the first step we have to decide, how many measurements are needed to solve the given problem. As we know, to calculate unknown parameters of the static system characteristic (2.1), following condition must be fulfilled:

$$N \times L \geq R, \quad (2.4)$$

where:

$N$  - number of measurements;

$L$  - dimensionality of the output vector;

$R$  - dimensionality of unknown vector of parameters for the static system characteristic.

Because, in our example,  $L = 2$ ,  $R = 4$ , it means that the minimal number of measurements, which are needed to solve the problem, is  $N = 2$ .

In the next step we have to select  $N = 2$  measurements from the table 2.1, which fulfills identification conditions. It means that for matrix of input measurements  $U_N$  we must obtain:

$$\det(U_N) \neq 0. \quad (2.5)$$

Let us choose from table 2.1 measurement  $n = 1$  and  $n = 4$ . Matrix  $U_N$  has form:

$$U_N = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}. \quad (2.6)$$

The identification condition can be verified in the following way:

$$\det(U_N) = \begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix} = 0. \quad (2.7)$$

The obtained result means that for selected measurements ( $n = 1$  and  $n = 4$ ) it is not possible to find the solution: a proposed sequence does not form identifiable sequence.

Let us try with a new set of measurements:  $n = 2$  and  $n = 5$ . Determinant of  $U_N$  is as follows:

$$\det(U_N) = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1. \quad (2.8)$$

Because  $\det(U_N) \neq 0$  means that we have found identifiable sequence, it can possible determine a set of parameters of static system's characteristic (2.1).

In the next step we have to determine identification algorithm, to calculate unknown parameters of a given characteristic (2.1).

Let us rewrite an equation (2.2), taking into account that  $N = 2$ :

$$Y_2 = \theta U_2, \quad (2.9)$$

where:

$$Y_2 = [y_1 \quad y_2], U_2 = [u_1 \quad u_2], \quad (2.10)$$

and

$$y_1 = \begin{bmatrix} y_1^{(1)} \\ y_1^{(1)} \end{bmatrix}, y_2 = \begin{bmatrix} y_2^{(1)} \\ y_2^{(1)} \end{bmatrix}, u_1 = \begin{bmatrix} u_1^{(1)} \\ u_1^{(1)} \end{bmatrix}, u_2 = \begin{bmatrix} u_2^{(1)} \\ u_2^{(1)} \end{bmatrix}. \quad (2.11)$$

In order to determine  $\theta$ , we have to solve the matrix equation (2.9):

$$Y_2 U_2^{-1} = \theta U_2 U_2^{-1}. \quad (2.12)$$

Finally, the algorithm to determine vector of parameters  $\theta$  has the form:

$$\theta = Y_2 U_2^{-1}. \quad (2.13)$$

Now, we can substitute in (2.13) measurements  $n = 2$  and  $n = 5$ . As the result, the following matrix equations are obtained:

$$\theta = \begin{bmatrix} -1 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1}, \quad (2.14)$$

and then we can get this solution:

$$\theta = \begin{bmatrix} -1 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}. \quad (2.15)$$

#### Verification of the solution:

Let us choose  $n = 1$  and  $n = 6$ . We can verify obtained results substituting in matrix  $U_2$  (2.9) data  $n = 1, n = 6$  and check that the determined values of the output are the same, as in the table 2.1:

$$Y_2 = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 5 & 7 \end{bmatrix}. \quad (2.16)$$

It is easy to notice that the same results will be obtained for other measurements set, providing that they form *identifiable* sequence.

**Example 2.2 – Determination of the system's parameter.** Let us assume that we have some static system characteristic described by following equation:

$$y = \theta^{(1)} + \theta^{(2)}u + \theta^{(3)}u^2. \quad (2.17)$$

**The problem:** To determine the parameters of the given characteristic (2.17) for the following set of measurements:

**Table 2.2. Measurements for example 2.2.**

n	1	2	3	4	5	6
$u$	-2	2	0	-1	3	1
$y$	9	25	5	4	44	12

Let us rewrite (2.17) in the form:

$$y = \begin{bmatrix} \theta^{(1)} & \theta^{(2)} & \theta^{(3)} \end{bmatrix} \begin{bmatrix} 1 \\ u \\ u^2 \end{bmatrix}. \quad (2.18)$$

Substituting:

$$\theta^T = \begin{bmatrix} \theta^{(1)} & \theta^{(2)} & \theta^{(3)} \end{bmatrix},$$

and

$$f(u) = \begin{bmatrix} 1 & u & u^2 \end{bmatrix}^T, \quad (2.19)$$

we can write connection (2.17) in the following form:

$$y = \theta^T f(u). \quad (2.20)$$

As it was considered in the first example, now we have to find some measurements, which form an *identifiable* sequence. Since, in this example  $L=1$  and  $R=3$ , it means that we need  $N=3$  measurements. For a given model (2.20) the system of equations has the form:

$$y_n = \theta^T f(u_n), \quad n = 1, 2, 3. \quad (2.21)$$

The system of equations (2.21) can be rewritten in the form (for  $N=3$ ):

$$Y_3 = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} = \begin{bmatrix} \theta^T f(u_1) & \theta^T f(u_2) & \theta^T f(u_3) \end{bmatrix}, \quad (2.22)$$

or:

$$Y_3^T = \bar{f}^T(U_3)\theta, \quad (2.23)$$

where:

$$\bar{f}(U_3) = [f(u_1) \quad f(u_2) \quad f(u_3)], \quad (2.24)$$

is square matrix.

In our example, (2.24) can be written as follows (for  $n = 1, 2, 3$ ):

$$\bar{f}(U_3) = \begin{bmatrix} 1 & 1 & 1 \\ -2 & 2 & 0 \\ 4 & 4 & 0 \end{bmatrix}, \quad (2.25)$$

and  $Y_3$ :

$$Y_3 = [9 \quad 25 \quad 5]. \quad (2.26)$$

Let us verify the *identification* condition:

$$\det(\bar{f}(U_3)) = -16. \quad (2.27)$$

It means that the proposed sequence is *identifiable*.

Now we may determine a vector of parameters  $\theta$ . Taking into account (2.23), the identification algorithm has the form:

$$\theta = \Psi_3(U_3, Y_3) = [\bar{f}^T(U_3)]^{-1} Y_3^T. \quad (2.28)$$

In the next step, substituting (2.25) and (2.26) to (2.28) results in obtaining some values of the parameter vector  $\theta$ :

$$\theta = \left( \begin{bmatrix} 1 & 1 & 1 \\ -2 & 2 & 0 \\ 4 & 4 & 0 \end{bmatrix}^T \right)^{-1} [9 \quad 25 \quad 5]^T = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}. \quad (2.29)$$

#### Verification of the solution:

Let us choose  $n = 4, n = 5$  and  $n = 6$ . We can verify obtained results substitute in (2.23) matrix  $\bar{f}^T(U_3)$  for  $n = 4, n = 5, n = 6$  and deduce that the determined values of the output are the same as in table 2.2:

$$Y_3^T = \bar{f}^T(U_3)\theta = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 3 & 1 \\ 1 & 9 & 1 \end{bmatrix}^T \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 44 \\ 12 \end{bmatrix}. \quad (2.30)$$

It is easy to notice that the same results will be obtained for other measurements set, provided that they form the identifiable sequence.

### Exercises

**Exercise 2.1.** Let us assume that we have some static system's characteristic, described by the following set of equations:

$$y^{(1)} = \theta_1^{(1)}u^{(1)} + \theta_1^{(2)}u^{(2)} + \theta_1^{(3)}u^{(3)}$$

$$y^{(2)} = \theta_2^{(1)}u^{(1)} + \theta_2^{(2)}u^{(2)} + \theta_2^{(3)}u^{(3)}$$

As a result of the experiment, the following data are collected:

**Table 2.3. Measurements for exercise 2.1.**

n	1	2	3	4	5	6
$u^{(1)}$	1	-1	-1	2	1	1
$u^{(2)}$	1	1	1	-1	-1	-1
$u^{(3)}$	1	2	1	-1	1	2
$y^{(1)}$	6	7	4	-3	2	5
$y^{(2)}$	4	-9	-6	11	0	-3

**The problem:** To calculate some parameters of the static system's characteristic.

**Exercise 2.2.** Let us assume that we have the static system's characteristic, described by the following set of equations:

$$y^{(1)} = \theta_1^{(1)} + \theta_1^{(2)}u^{(2)} + \theta_1^{(3)}(u^{(3)})^2$$

$$y^{(2)} = \theta_2^{(1)} + \theta_2^{(2)}u^{(2)} + \theta_2^{(3)}(u^{(3)})^2$$

As the result of the experiment, the following data are measured:

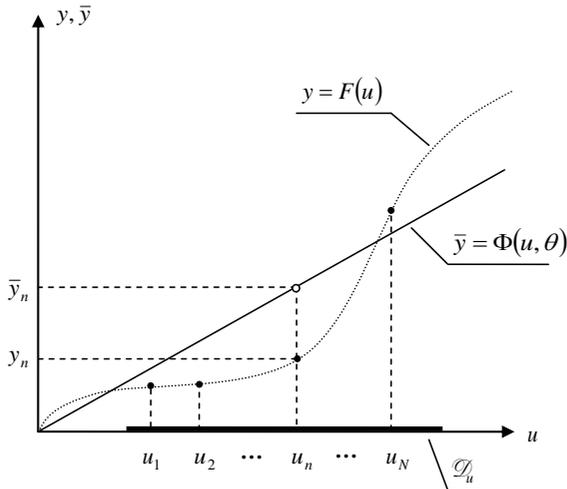
**Table 2.4. Measurements for exercise 2.2.**

n	1	2	3	4	5	6
$u^{(1)}$	2	-2	1	1	1	3
$u^{(2)}$	0	-2	-1	3	1	1
$y^{(1)}$	1	-3	0	8	0	4
$y^{(2)}$	2	-6	0	-8	0	2

**The problem:** To calculate some parameters of the static system's characteristic.

## 2.2. The choice of the Best Model: Deterministic Case

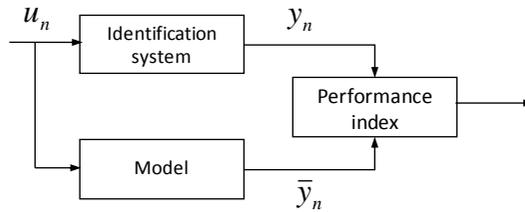
In this case we do not have any exact model of the static characteristic for investigated system, process or phenomenon. In order to solve given identification problem, an investigator determines arbitrary class of a model. Then, taking into account measured data, it is possible to find the best model for arbitrary given class of the model. This problem is called "the choice of the best model" (Bub80). In this section we assume, that the collected data are free noise measurements (Figure 2.4).



**Figure 2.4. The choice of the best model – measurements and model's output.**

As it is shown in Figure 2.4, our problem is to find approximation of the static system's characteristic,  $F(u)$  by use an arbitrary given function  $\Phi(u, \theta)$  based on collected data of inputs and outputs.

In order to compare obtained results with collected data for each measurement point, it is necessary to define performance index, which is the measurement of difference between  $F(u)$  and  $\Phi(u, \theta)$  (Figure 2.5).



**Figure 2.5. The choice of the best model.**

**The problem formulation:**

Approximation function (model):  $\bar{y} = \Phi(u, \theta)$ ;

$\Phi$  - arbitrary given function;

$\bar{y}$  - model output vector  $\bar{y} \in \bar{\mathcal{Y}} \subseteq \mathbb{R}^L$ ,  $\bar{\mathcal{Y}}$  - output domain;

$u$  - input vector  $u \in \mathcal{U} \subseteq \mathbb{R}^S$ ,  $\mathcal{U}$  - input domain;

$\theta$  - vector of model parameters:  $\theta \in \Theta \subseteq \mathbb{R}^R$ ;

Experiment  $U_N = [u_1 \ u_2 \ \dots \ u_N]$ ,  $Y_N = [y_1 \ y_2 \ \dots \ y_N]$ ;

Measure of difference:  $\forall n = 1, 2, \dots, N \ q(y_n, \bar{y}_n) = q(y_n, \Phi(u_n, \theta))$ .

**Example 2.3 – The choice of the best model.** Let us assume that we have one dimensional system (SISO - Single Input Single Output) and measurements for three different values of the input and output:

**Table 2.5. Measurements for 2.1.**

$u$	-2	0	3
$y$	-2	1	3

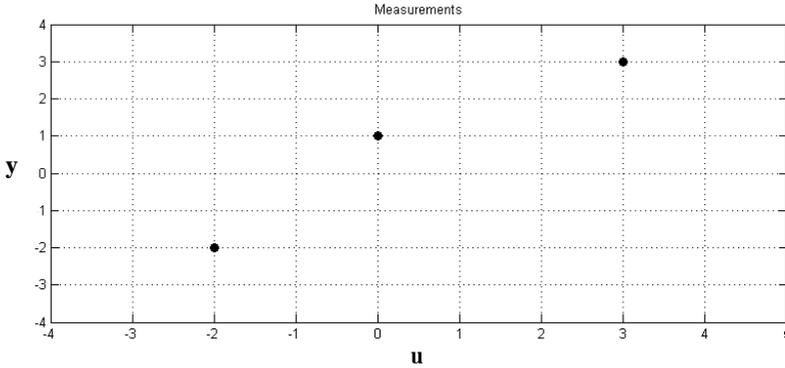


Figure 2.6. Measurements for 2.1.

**The problem:** To determine the optimal parameter value  $\theta$  for model  $\bar{y} = \theta u$  for given measurements and the following performance index:

$$Q_N(\theta) = \sum_{n=1}^N (y_n - \bar{y}_n)^2 . \quad (2.31)$$

In the first step we can substitute given measurements (table 2.5) in the performance index  $Q_N(\theta)$  (2.31):

$$Q_N(\theta) = (-2 - 2\theta)^2 + (1 - 0\theta)^2 + (3 - 3\theta)^2 . \quad (2.32)$$

After a few simple transformations, we determine:

$$Q_N(\theta) = 13\theta^2 - 26\theta + 14 . \quad (2.33)$$

In order to calculate some optimal value of  $\theta$  for model  $\bar{y} = \theta u$ , some data (from table 2.5) and the performance index (2.31), we have to solve the following equation:

$$\left. \frac{dQ_N(\theta)}{d\theta} \right|_{\theta=\theta_N^*} = 0 . \quad (2.34)$$

Taking into account (2.33) and (2.34), we obtain such equation as:

$$26\theta_N^* - 26 = 0 , \quad (2.35)$$

which means, that the optimal value of  $\theta$  for given conditions is:

$$\theta_N^* = 1. \quad (2.36)$$

Our model has form:  $\bar{y} = u$ . The results of the model determination are shown in Figure 2.7:

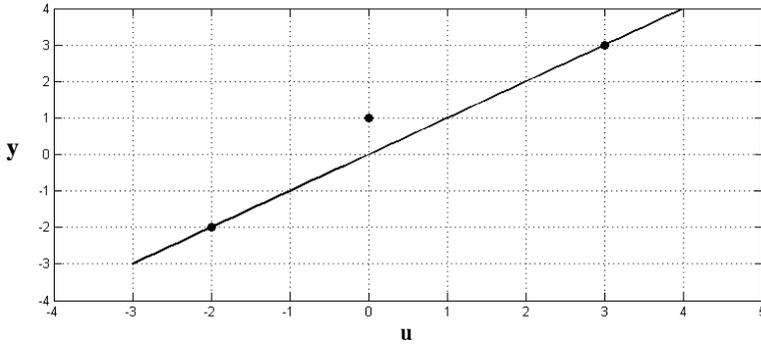


Figure 2.7. The results of parameters estimation.

**Example 2.4 – The choice of the best model.** For given measurements matrix  $\mathbf{U}_N$ ,  $\mathbf{Y}_N$  and quadratic performance index, the identification algorithm for following models must be found:

- a)  $\bar{y} = \theta u$ ;
- b)  $\bar{y} = \theta^{(1)} u + \theta^{(0)}$ .

We start with first example (a). Taking into account the performance index and the model  $\bar{y} = \theta u$ , we can write the following formula:

$$Q_N(\theta) = \sum_{n=1}^N (y_n - \theta u_n)^2. \quad (2.37)$$

In order to determine the optimal parameter  $\theta$ , we have to solve the following equation:

$$\left. \frac{dQ_N(\theta)}{d\theta} \right|_{\theta=\theta_N^*} = 0. \quad (2.38)$$

Next, we can make some transformations:

$$\frac{dQ_N(\theta)}{d\theta} = \sum_{n=1}^N 2(y_n - \theta u_n)(-u_n) = 2 \sum_{n=1}^N (-u_n y_n + \theta u_n^2), \quad (2.39)$$

and then

$$2 \sum_{n=1}^N (-u_n y_n + \theta_N^* u_n^2) = 0, \quad (2.40)$$

$$- \sum_{n=1}^N u_n y_n + \sum_{n=1}^N \theta_N^* u_n^2 = 0, \quad (2.41)$$

$$\sum_{n=1}^N \theta_N^* u_n^2 = \sum_{n=1}^N u_n y_n. \quad (2.42)$$

Finally, the algorithm to calculate the optimal value of parameter  $\theta$  has a form:

$$\theta_N^* = \frac{\sum_{n=1}^N u_n y_n}{\sum_{n=1}^N u_n^2}. \quad (2.43)$$

Let us now consider the same problem of calculating some optimal values of the parameters vector,  $\theta$  but for the second model (b)  $\bar{y} = \theta^{(1)} u + \theta^{(0)}$ . All conditions remain the same. The performance index for this problem has the form:

$$Q_N(\theta) = \sum_{n=1}^N (y_n - (\theta^{(1)} u_n + \theta^{(0)}))^2 = \sum_{n=1}^N (y_n - \theta^{(1)} u_n - \theta^{(0)})^2. \quad (2.44)$$

In this example the performance index depends on two variables i.e.:  $\theta^{(0)}$  and  $\theta^{(1)}$ . It means that we have to solve the set of equations which fulfills the following condition:

$$\nabla_{\theta} Q(\theta) \Big|_{\theta=\theta_N^*} = \mathbf{0}_2. \quad (2.45)$$

We can rewrite it as:

$$\frac{\partial Q(\theta)}{\partial \theta^{(0)}} \Big|_{\theta=\theta_N^{(0)}} = 0, \quad (2.46)$$

$$\frac{\partial Q(\theta)}{\partial \theta^{(1)}} \Big|_{\theta=\theta_N^{(1)}} = 0. \quad (2.47)$$

Taking into account equations (2.44) and (2.45), we can determine such derivatives for variable  $\theta^{(0)}$ :

$$\frac{\partial Q(\theta)}{\partial \theta^{(0)}} = \sum_{n=1}^N 2(y_n - \theta^{(1)} u_n - \theta^{(0)})(-1) = 2 \sum_{n=1}^N (-y_n + \theta^{(1)} u_n + \theta^{(0)}), \quad (2.48)$$

and  $\theta^{(1)}$ :

$$\frac{\partial Q(\theta)}{\partial \theta^{(1)}} = \sum_{n=1}^N 2(y_n - \theta^{(1)} u_n - \theta^{(0)})(-u_n) = 2 \sum_{n=1}^N (-u_n y_n + \theta^{(1)} u_n^2 + \theta^{(0)} u_n). \quad (2.49)$$

Now, we focus on the equation (2.48):

$$2 \sum_{n=1}^N (-y_n + \theta_N^{*(1)} u_n + \theta_N^{*(0)}) = 0, \quad (2.50)$$

$$-\sum_{n=1}^N y_n + \sum_{n=1}^N \theta_N^{*(0)} + \sum_{n=1}^N \theta_N^{*(1)} u_n = 0, \quad (2.51)$$

$$-\sum_{n=1}^N y_n + \theta_N^{*(0)} \sum_{n=1}^N 1 + \theta_N^{*(1)} \sum_{n=1}^N u_n = 0, \quad (2.52)$$

$$\theta_N^{*(0)} N + \theta_N^{*(1)} \sum_{n=1}^N u_n = \sum_{n=1}^N y_n, \quad (2.53)$$

$$\theta_N^{*(0)} + \theta_N^{*(1)} \frac{1}{N} \sum_{n=1}^N u_n = \frac{1}{N} \sum_{n=1}^N y_n, \quad (2.54)$$

and (2.49):

$$2 \sum_{n=1}^N (-u_n y_n + \theta_N^{*(1)} u_n^2 + \theta_N^{*(0)} u_n) = 0, \quad (2.55)$$

$$-\sum_{n=1}^N u_n y_n + \sum_{n=1}^N \theta_N^{*(0)} u_n + \sum_{n=1}^N \theta_N^{*(1)} u_n^2 = 0, \quad (2.56)$$

$$\theta_N^{*(0)} \sum_{n=1}^N u_n + \theta_N^{*(1)} \sum_{n=1}^N u_n^2 = \sum_{n=1}^N u_n y_n. \quad (2.57)$$

For sake of simplicity, let us substitute:  $\hat{u} = \frac{1}{N} \sum_{n=1}^N u_n$  and  $\hat{y} = \frac{1}{N} \sum_{n=1}^N y_n$ . Now, we can rewrite some connections between (2.54) and (2.57) in the form:

$$\theta_N^{*(0)} + \theta_N^{*(1)} \hat{u} = \hat{y}, \quad (2.58)$$

$$\theta_N^{*(0)} \hat{u} + \theta_N^{*(1)} \frac{1}{N} \sum_{n=1}^N u_n^2 = \frac{1}{N} \sum_{n=1}^N u_n y_n. \quad (2.59)$$

From (2.58) we can determine:

$$\theta_N^{*(0)} = \hat{y} - \theta_N^{*(1)} \hat{u}. \quad (2.60)$$

Substituting (2.60) in (2.59) we obtain:

$$(\hat{y} - \theta_N^{*(1)} \hat{u}) \hat{u} + \theta_N^{*(1)} \frac{1}{N} \sum_{n=1}^N u_n^2 = \frac{1}{N} \sum_{n=1}^N u_n y_n, \quad (2.61)$$

and then

$$\hat{y} \hat{u} - \theta_N^{*(1)} \hat{u}^2 + \theta_N^{*(1)} \frac{1}{N} \sum_{n=1}^N u_n^2 = \frac{1}{N} \sum_{n=1}^N u_n y_n. \quad (2.62)$$

We can rewrite (2.61) in the form:

$$\theta_N^{*(1)} \left( \frac{1}{N} \sum_{n=1}^N u_n^2 - \hat{u}^2 \right) + \hat{y} \hat{u} = \frac{1}{N} \sum_{n=1}^N u_n y_n, \quad (2.63)$$

and finally:

$$\theta_N^{*(1)} = \frac{\frac{1}{N} \sum_{n=1}^N u_n y_n - \hat{y} \hat{u}}{\frac{1}{N} \sum_{n=1}^N u_n^2 - \hat{u}^2}. \quad (2.64)$$

Let us summarize our calculations determining the algorithm to estimate two-dimensional variable  $\theta$ , i.e.:

$$\theta_N^{*(0)} = \hat{y} - \theta_N^{*(1)} \hat{u}, \quad (2.65)$$

$$\theta_N^{*(0)} = \frac{\frac{1}{N} \sum_{n=1}^N u_n y_n - \hat{y} \hat{u}}{\frac{1}{N} \sum_{n=1}^N u_n^2 - \hat{u}^2}. \quad (2.66)$$

**Example 2.5 – The static system’s approximation with the weight function.** For system described by the following static characteristic  $y = F(u) = \theta u^2$  find the best approximation  $\bar{y} = \Phi(u, \theta) = \theta u$  in  $\mathcal{U}_u = \{u \in \mathcal{R} : 0 \leq u \leq 1\}$  with weight function  $g_u(u)$ , presented in Figure 2.8:

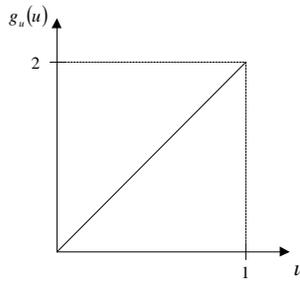


Figure 2.8. The weight function.

and the performance index:  $Q(\theta) = \int_{\mathcal{U}_u} (F(u) - \Phi(u, \theta))^2 g_u(u) du$ . The given problem of the static system’s approximation is illustrated in the Figure 2.9.

Let us start with the substituting system’s characteristic and model to the quadratic performance index:

$$Q(\theta) = \int_{\mathcal{U}_u} (u^2 - \theta u)^2 (2u) du. \quad (2.67)$$

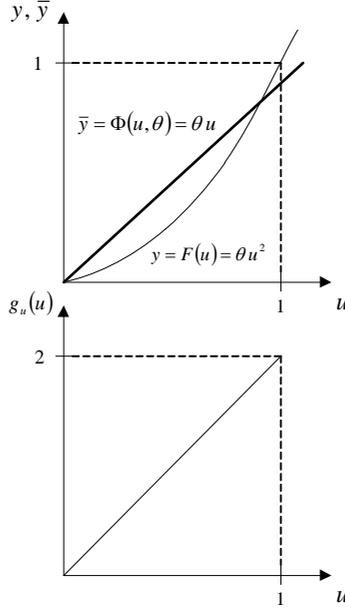


Figure 2.9. Approximation of the static system's characteristic with the weight function  $g_u(u)$ .

In order to find the optimal value of the parameter  $\theta$ , we have to solve the following equation:

$$\frac{dQ(\theta)}{d\theta} = 4 \int_{\mathcal{U}_u} (u^2 - \theta u)(u)(-u) du \Big|_{\theta=\theta^*} = 0. \quad (2.68)$$

Since  $\mathcal{U}_u = \{u \in \mathcal{R} : 0 \leq u \leq 1\}$ :

$$\int_0^1 (-u^4 + \theta^* u^3) du = 0, \quad (2.69)$$

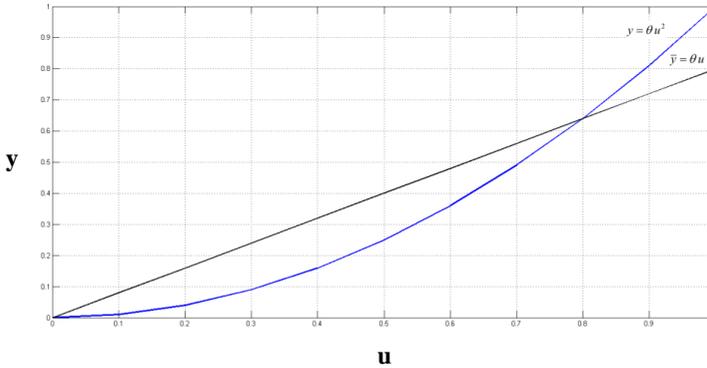
and then:

$$-\frac{u^5}{5} \Big|_0^1 + \theta^* \frac{u^4}{4} \Big|_0^1 = 0, \quad (2.70)$$

$$-\frac{1}{5} + \theta^* \frac{1}{4} = 0. \quad (2.71)$$

Finally, the optimal value of  $\theta$  is:  $\theta^* = \frac{4}{5}$ . (2.72)

**Verification of the solution:**



**Figure 2.10.** Characteristic of the system and its approximation.

### Exercises

**Exercise 2.3.** Let us assume that we have one dimensional system (SISO - Single Input Single Output) and some measurements for three different values of the input and output:

**Table 2.6.** Measurements for exercise 2.3.

$u$	-1	1	3
$y$	-2	1	2

**The problem:** To determine the optimal parameter value  $\theta$  for the model  $\bar{y} = \theta u$  for given measurements and the following performance index:  $Q_N(\theta) = \sum_{n=1}^N (y_n - \bar{y}_n)^2$ .

**Exercise 2.4.** For the system described by following characteristic  $y = F(u) = \theta u^2$ , find the best approximation  $\bar{y} = \Phi(u, \theta) = \theta u$  in  $\mathcal{U}_u = \{u \in \mathcal{R} : 0 \leq u \leq 1\}$  with the weight function  $g_u(u)$ , as in the Figure 2.11, and the performance index:  $Q(\theta) = \int_{\mathcal{U}_u} (F(u) - \Phi(u, \theta))^2 g_u(u) du$ .

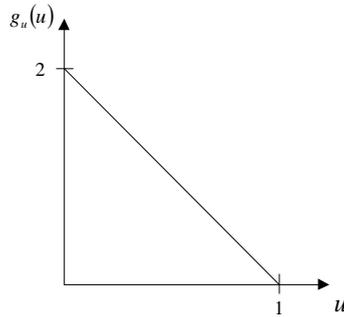


Figure 2.11 The weight function in Exercise 2.4.

**Exercise 2.5.** For the model described by the following characteristic  $y = F(u) = \theta u^2$ , find the best approximation  $\bar{y} = \Phi(u, \theta) = \theta^{(1)}u + \theta^{(0)}$  in  $\mathcal{G}_u = \{u \in \mathcal{R} : 0 \leq u \leq 1\}$  with the weight functions  $g_u(u)$ , as below:

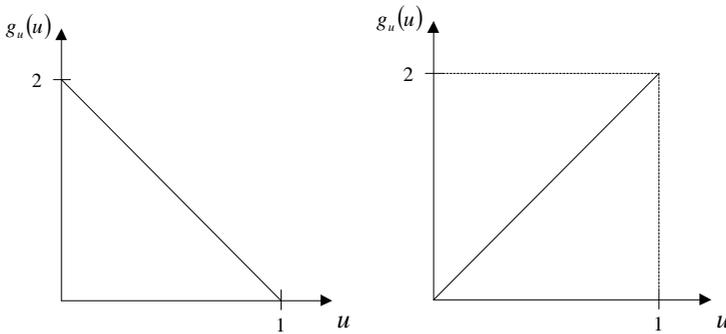


Figure 2.12. The weight functions for Exercise 2.5.

and the performance index:  $Q(\theta) = \int_{\mathcal{G}_u} (F(u) - \Phi(u, \theta))^2 g_u(u) du$ .

### 2.3 Parameter Estimation of the System's Characteristic

In section 2.1 we have considered the problem of the system parameters' determination, based on the noise free measurements. Here, we take into account that collected data are not free noise measurements, i.e. for the given input collected output is not free noise (Bub80). This problem is illustrated in Figure 2.13.

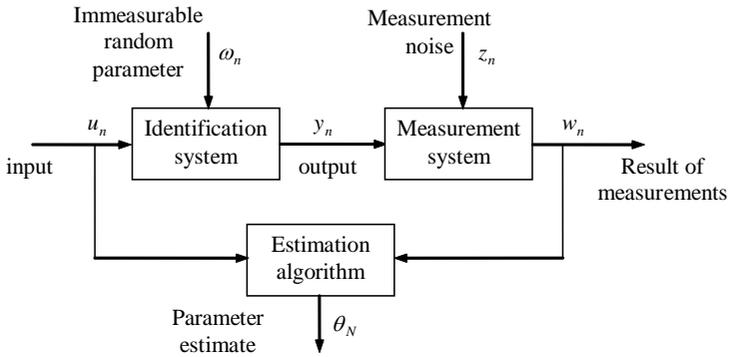


Figure 2.13. System parameter estimation problem.

This problem is called “parameter estimation of the system’s characteristic”. It is possible to distinguish such following cases as:

- a) Noised measurements of the physical values;
- b) Estimation of the system parameter with noisy measurements;
- c) Estimation of the system parameter with random value.

#### Noised measurements of the physical values

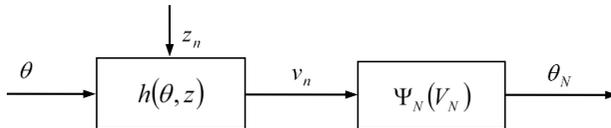


Figure 2.14. Noised measurements of the physical values.

**The problem formulation:**

Measurement noise:

$z_n$  – value of random variable  $\underline{z}$  from space  $\mathcal{Z}$  ;

$f_z(z)$ – probability density function;

$\theta$  – observed vector of parameters, value of random variable  $\underline{\theta}, \theta \in \Theta \subseteq \mathcal{R}^R$  ;

$f_\theta(\theta)$  – probability density function;

Measurements:  $V_N = [v_1 \quad v_2 \quad \dots \quad v_N]$ .

**Estimation of the system parameter with noisy measurements**

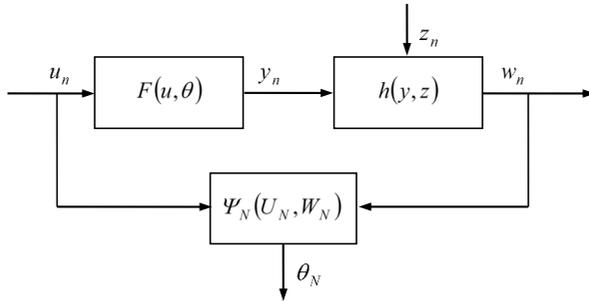


Figure 2.15. Estimation of the system parameter with noisy measurements.

**The problem formulation:**

Measurement system's description:  $w = h(y, z)$ ;

Where:  $w \in \mathcal{W}$ ,  $h$  is known one-to-one function such as  $h: \mathcal{Y} \times \mathcal{Z} \rightarrow \mathcal{W}$  ;

$\mathcal{W}$  – measurements domain ( $\dim y = \dim z = L$ );

Measurement noise:

$z_n$  – the observed vector of parameters, value of random variable  $\underline{\theta}, \theta \in \Theta \subseteq \mathcal{R}^R$  ;

$f_\theta(\theta)$  – value of random variable  $f_\theta(\theta)$  from space  $\mathcal{X}$  ;

$f_z(z)$ – probability density function;

Measurements:  $U_N = [u_1 \ u_2 \ \dots \ u_N]$ ,  $W_N = [w_1 \ w_2 \ \dots \ w_N]$ .

### Estimation of the system parameter with random value

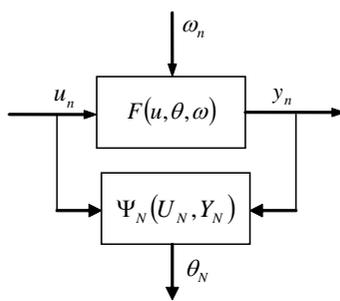


Figure 2.16. Estimation of the system parameter with random value.

### The problem formulation:

Static system characteristic:  $y = F(u, \theta, \omega)$ ;

Random system parameter:  $\omega \in \Omega \subseteq \mathcal{R}^L$ , ( $\dim y = \dim \omega = L$ );

$F$  – one-to-one mapping  $\omega = F_\omega^{-1}(u, \theta, y)$ ;

$\omega_n$  – value of random variable  $\omega$  from  $\Omega \subseteq \mathcal{R}^L$ ;

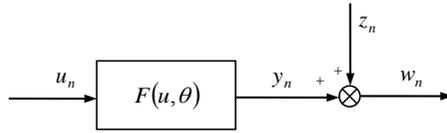
Probability density function  $f_\omega(\omega)$  is given;

Measurements:  $U_N = [u_1 \ u_2 \ \dots \ u_N]$ ,  $Y_N = [y_1 \ y_2 \ \dots \ y_N]$ .

**Example 2.6 – Estimation of the system parameter with noisy measurements: Least Squares**

**Method.** For linear system with static characteristic  $y = \theta^T \varphi(u)$ , we have  $N$  measurements. Input and output sequences are collected in matrix  $U_N = [u_1 \ u_2 \ \dots \ u_N]$  and  $W_N = [w_1 \ w_2 \ \dots \ w_N]$  respectively. We also know that the noise is additive, i.e.  $w_n = y_n + z_n$ , with  $m_z = 0$  (which means that expected value of the noise signal is zero) and  $\sigma_z < \infty$  (variance of the noise is not infinite). The **problem:** To determine the estimator of  $\theta$  parameter.

**The solution:** The given problem of the parameter estimation is depicted in the figure below:



**Figure 2.17. Noised measurements of the identification system known as static's characteristics.**

Since the noise is additive, the expected value of the noise signal is zero ( $m_z = 0$ ) and the variance of the noise is not infinite ( $\sigma_z < \infty$ ). It means that it is possible to apply Least Squares Method. Our problem is to find  $\theta_N$ , which minimize an empirical variance of a noise signal, i.e.:

$$\text{Var}_{z_N}(\theta; \mathbf{U}_N, \mathbf{Y}_N) = \frac{1}{N} \sum_{n=1}^N (w_n - y_n)^2 = \frac{1}{N} \sum_{n=1}^N (w_n - \theta^T \varphi(u_n))^2. \quad (2.73)$$

Let us formulate the given problem as an optimization task:

$$\theta_N \rightarrow \text{Var}_{z_N}(\theta_N; U_N, W_N) = \min_{\theta \in \Theta} \text{Var}_{z_N}(\theta; U_N, W_N). \quad (2.74)$$

To determine  $\theta_N$  (2.74) we have to solve the following set of equations:

$$\nabla_{\theta} \text{Var}_{z_N}(\theta; U_N, W_N) \Big|_{\theta=\theta_N} = 0_R. \quad (2.75)$$

Taking into account (2.73) and (2.75), we can write the following connection:

$$\nabla_{\theta} \sum_{n=1}^N (w_n - \theta^T \varphi(u_n))^2 = 0_R. \quad (2.76)$$

In order to solve the following equation, we have to make some transformations:

$$2 \sum_{n=1}^N (w_n - \theta_N^T \varphi(u_n)) (-\varphi(u_n)) = 0_R, \quad (2.77)$$

we can simplify our connection:

$$- \sum_{n=1}^N w_n \varphi(u_n) - \sum_{n=1}^N (-\varphi(u_n)) (\theta_N^T \varphi(u_n))^T = 0_R, \quad (2.78)$$

$$- \sum_{n=1}^N w_n \varphi(u_n) + \sum_{n=1}^N \varphi(u_n) (\varphi(u_n)^T \theta_N) = 0_R, \quad (2.79)$$

$$- \sum_{n=1}^N w_n \varphi(u_n) + \sum_{n=1}^N \varphi(u_n) \varphi(u_n)^T \theta_N = 0_R. \quad (2.80)$$

Finally, an algorithm to determine  $\theta_N$  has form:

$$\theta_N = \sum_{n=1}^N (\varphi(u_n) \varphi(u_n)^T)^{-1} \sum_{n=1}^N w_n \varphi(u_n). \quad (2.81)$$

**Example 2.7 – Estimation of the system parameter with noisy measurements: Maximum Likelihood**

**Method.** For fixed sequence of inputs  $u_1, u_2, \dots, u_N$  and for linear system with characteristic  $y = \theta u$ , it was collected following sequence of noised outputs  $w_1, w_2, \dots, w_N$ . **The problem:** to determine the estimator of  $\theta$  parameter for the additive noise ( $w_n = y_n + z_n$ ) and probability density function

$$f_z(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z - m_z)^2}{2\sigma_z^2}\right).$$

**The solution:** A given problem of the parameter estimation is illustrated in Figure 2.17.

Since these following assumptions are fulfilled:

- a) The measurement system is described by  $w_n = y_n + z_n$ , which is one-to-one invertible function,
- b) Probability density function for  $f_z(z)$  is given,

it is possible to apply Maximum Likelihood Method.

At the beginning, let us denote a measurement system  $w_n = y_n + z_n$  as:

$$w_n = h(y_n, z_n). \quad (2.82)$$

In order to build the likelihood function, we need to determine  $f_w(w_n, \theta)$ , which is probability density function of  $w$  variable. Because the function describes the influence of noise, ( $w_n = y_n + z_n$ ) is one-to-one invertible function, we can find unknown connection from the following equations:

$$f_w(w_n, \theta) = f_z(h_z^{-1}(y_n, w_n)) \cdot |J_h|, \quad (2.83)$$

where  $J_h$  is Jacobi matrix of inverse transformation and  $h_z^{-1}$  is inverse function with respect to  $z$ .

In our example, function  $h_z^{-1}(y_n, w_n)$  is equal:

$$h_z^{-1}(y_n, w_n) = w_n - y_n. \quad (2.84)$$

Now, we can determine Jacobi  $J_h$  matrix by using the following formula:

$$J_h = \frac{\partial h_z^{-1}(y_n, w_n)}{\partial w} = \frac{d}{dw}(w_n - y_n) = 1. \quad (2.85)$$

The likelihood function has the form:

$$L_N(\theta; W_N, U_N) = \prod_{n=1}^N f_w(w_n, \theta) = \prod_{n=1}^N f_z(h_z^{-1}(y_n, w_n)) |J_h|. \quad (2.86)$$

Because  $J_h = 1$  and  $y = \theta u$ , we can rewrite (2.86) as below:

$$L_N(\theta; W_N, U_N) = \prod_{n=1}^N f_z(h_z^{-1}(\theta u_n, w_n)). \quad (2.87)$$

In order to find the optimal value of the parameter  $\theta$  for the connection (2.87), we have to formulate and then solve an optimization problem. In our example, the optimization task can be formulated as follows:

$$\theta_N \rightarrow L_N(\theta; W_N, U_N) = \max_{\theta \in \Theta} L_N(\theta; W_N, U_N). \quad (2.88)$$

Taking into account given probability density function for  $f_z(z)$ , we can rewrite (2.88) as:

$$L_N(\theta; W_N, U_N) = \prod_{n=1}^N \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{(w_n - \theta u_n - m_z)^2}{2\sigma_z^2}\right). \quad (2.89)$$

Let us now rewrite (2.89) in the form:

$$L_N(\theta; W_N, U_N) = \prod_{n=1}^N \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\frac{(w_n - \theta u_n - m_z)^2}{2\sigma_z^2}\right). \quad (2.90)$$

It is possible to write (2.90) in the form:

$$L_N(\theta; W_N, U_N) = \frac{1}{\sqrt{2\pi}\sigma_z} \exp\left(-\sum_{n=1}^N \frac{(w_n - \theta u_n - m_z)^2}{2\sigma_z^2}\right). \quad (2.90)$$

To find the estimated value of  $\theta$ , we need to consider only a part of connection(2.90). Let us denote:

$$\Gamma(\theta) = -\sum_{n=1}^N \frac{(w_n - \theta u_n - m_z)^2}{2\sigma_z^2}. \quad (2.91)$$

Let us formulate the following optimization problem:

$$\left. \frac{d\Gamma(\theta)}{d\theta} \right|_{\theta=\theta_N} = 0. \quad (2.92)$$

Substituting (2.91) in (2.92), we obtain:

$$\frac{d\Gamma(\theta)}{d\theta} = \frac{d}{d\theta} \left( -\sum_{n=1}^N \frac{(w_n - \theta u_n - m_z)^2}{2\sigma_z^2} \right) = 0. \quad (2.93)$$

In the next few steps, we will determine an algorithm to find the optimal value of  $\theta$ :

$$\frac{d}{d\theta} \left( -\frac{1}{2\sigma_z^2} \sum_{n=1}^N (w_n - \theta u_n - m_z)^2 \right) = 0, \quad (2.94)$$

$$-\frac{2}{2\sigma_z^2} \sum_{n=1}^N (w_n - \theta u_n - m_z)(-u_n) = 0, \quad (2.95)$$

$$\frac{1}{\sigma_z^2} \left( \sum_{n=1}^N w_n u_n - \theta_N \sum_{n=1}^N u_n^2 - m_z \sum_{n=1}^N u_n \right) = 0. \quad (2.96)$$

Since  $\sigma_z^2$  is constant:

$$\theta_N \sum_{n=1}^N u_n^2 = \sum_{n=1}^N w_n u_n - m_z \sum_{n=1}^N u_n. \quad (2.97)$$

Finally, the formula to determine the optimal value of  $\theta_N$  has the form:

$$\theta_N = \frac{\sum_{n=1}^N w_n u_n - m_z \sum_{n=1}^N u_n}{\sum_{n=1}^N u_n^2} = \frac{\sum_{n=1}^N (w_n - m_z) u_n}{\sum_{n=1}^N u_n^2}. \quad (2.98)$$

**Example 2.8 – Estimation of the system’s parameter with noisy measurements: Bayesian Method.**

For linear system  $y = \theta u$  it was obtained N measurements, i.e. for input sequence  $u_1, u_2, \dots, u_N$  output sequence  $w_1, w_2, \dots, w_N$  was collected. It is assumed that the noise is additive ( $w_n = y_n + z_n$ ) and probability density function is Gaussian, which means  $m_z = 0$  and variance  $\sigma_z < 0$ . **The**

**problem:** to determine the estimator of  $\theta$  parameter for following loss function  $L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$  and probability density function for  $\theta$  parameter population described by

$$f_\theta(\theta) = \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp\left(-\frac{(\theta - m_\theta)^2}{2\sigma_\theta^2}\right).$$

**The solution:** The given problem of the parameter estimation is illustrated in Figure 2.17.

To solve the problem by use of Bayesian method, we have to start with the verification that following conditions for this method are fulfilled:

- a) Measurement system is described by  $w_n = y_n + z_n$  which is one-to-one invertible function;
- b) Probability density functions for  $f_z(z)$  and  $f_\theta(\theta)$  are given;
- c) Loss function is defined i.e.:  $L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$ , where  $\delta$  is Dirac delta function.

Since  $L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$ , the condition risk has form:

$$r(\bar{\theta}, W_N) = -\int_{\Theta} \delta(\theta - \bar{\theta}) f'(\theta | W_N) d\theta = -f'(\bar{\theta} | W_N). \quad (2.99)$$

For (2.99) we can formulate an optimization problem:

$$\theta_N \rightarrow f'(\theta_N | W_N) = \max_{\bar{\theta} \in \Theta} f'(\bar{\theta} | W_N), \quad (2.100)$$

such as

$$f_\theta(\theta_N) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \theta_N), w_n)) | J_h| = \max_{\bar{\theta} \in \Theta} f_\theta(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \bar{\theta}), w_n)) | J_h|. \quad (2.101)$$

In this example the noise is additive i.e.:  $w_n = y_n + z_n$ . Let us denote it as follows:

$$w_n = h(y_n, z_n) = y_n + z_n. \quad (2.102)$$

Because  $w_n = h(y_n, z_n)$  is one-to-one invertible function, we can find:

$$z_n = h_z^{-1}(y_n, w_n). \quad (2.103)$$

which  $h_z^{-1}$  is an inverse function with respect to  $z$ .

Substituting  $y_n = \theta u_n$  in (2.102), we have obtained:

$$z_n = h_z^{-1}(\theta u_n, w_n) = w_n - \theta u_n. \quad (2.104)$$

Let us now determine Jacobi matrix:

$$J_h = \frac{\partial h_z^{-1}(\theta u_n, w_n)}{\partial w} = \frac{d}{dw} (w_n - \theta u_n) = 1. \quad (2.105)$$

Taking into account above calculations (2.102 – 2.105), we can determine probability density function of the observed value  $f_w(w_n|\theta)$  i.e.:

$$f_w(w_n|\theta) = f_z(h_z^{-1}(\theta u_n, w_n)) |J_h| = \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[ -\frac{(w_n - \theta u_n)^2}{2\sigma_z^2} \right] \cdot |1|, \quad (2.106)$$

and rewrite the connection(2.101):

$$\begin{aligned} f_\theta(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \bar{\theta}), w_n)) |J_h| &= \\ &= \frac{1}{\sigma_\theta \sqrt{2\pi}} \exp \left[ -\frac{(\bar{\theta} - m_\theta)^2}{2\sigma_\theta^2} \right] \prod_{n=1}^N \frac{1}{\sigma_z \sqrt{2\pi}} \exp \left[ -\frac{(w_n - \bar{\theta} u_n)^2}{2\sigma_z^2} \right], \end{aligned} \quad (2.107)$$

which is *a posteriori* of probability density function.

Let us rearrange the formula (2.107):

$$\begin{aligned} f_\theta(\bar{\theta}) \prod_{n=1}^N f_z(h_z^{-1}(F(u_n, \bar{\theta}), w_n)) |J_h| &= \\ &= \frac{1}{\sigma_\theta \sqrt{2\pi}} \left( \frac{1}{\sigma_z \sqrt{2\pi}} \right)^N \exp \left[ -\frac{(\bar{\theta} - m_\theta)^2}{2\sigma_\theta^2} - \sum_{n=1}^N \frac{(w_n - \bar{\theta} u_n)^2}{2\sigma_z^2} \right]. \end{aligned} \quad (2.108)$$

In order to find value of  $\theta_N$ , we have to maximize the following part of (2.108) the equation:

$$\Gamma(\bar{\theta}) = -\frac{(\bar{\theta} - m_\theta)^2}{2\sigma_\theta^2} - \sum_{n=1}^N \frac{(w_n - \bar{\theta} u_n)^2}{2\sigma_z^2}, \quad (2.109)$$

which means that we have to solve the following equation:

$$\left. \frac{d\Gamma(\bar{\theta})}{d\bar{\theta}} \right|_{\bar{\theta}=\theta_N} = 0. \quad (2.110)$$

Substituting (2.109) in (2.110):

$$\left. \frac{d\Gamma(\bar{\theta})}{d\bar{\theta}} = \frac{d}{d\bar{\theta}} \left( -\frac{(\bar{\theta} - m_\theta)^2}{2\sigma_\theta^2} - \sum_{n=1}^N \frac{(w_n - \bar{\theta} u_n)^2}{2\sigma_z^2} \right) \right|_{\bar{\theta}=\theta_N} = 0. \quad (2.111)$$

Then, we obtain an algorithm to determine the value of  $\theta_N$ :

$$-\frac{2(\theta_N - m_\theta)}{2\sigma_\theta^2} - 2 \sum_{n=1}^N \frac{(w_n - \theta_N u_n)(-u_n)}{2\sigma_z^2} = 0. \quad (2.112)$$

We can simplify this equation (2.112):

$$\frac{-\theta_N + m_\theta}{\sigma_\theta^2} - \sum_{n=1}^N \frac{(-w_n u_n + \theta_N u_n^2)}{\sigma_z^2} = 0, \quad (2.113)$$

$$\theta_N \left( -\frac{1}{\sigma_\theta^2} - \frac{1}{\sigma_z^2} \sum_{n=1}^N u_n^2 \right) = -\frac{1}{\sigma_z^2} \sum_{n=1}^N w_n u_n - \frac{1}{\sigma_\theta^2} m_\theta, \quad (2.114)$$

$$\theta_N = \frac{\frac{1}{\sigma_z^2} \sum_{n=1}^N w_n u_n + \frac{1}{\sigma_\theta^2} m_\theta}{\frac{1}{\sigma_\theta^2} + \frac{1}{\sigma_z^2} \sum_{n=1}^N u_n^2}. \quad (2.115)$$

Finally:

$$\theta_N = \frac{m_\theta + \left( \frac{\sigma_\theta}{\sigma_z} \right)^2 \sum_{n=1}^N w_n u_n}{1 + \left( \frac{\sigma_\theta}{\sigma_z} \right)^2 \sum_{n=1}^N u_n^2}. \quad (2.116)$$

Discussion of the results:

1.  $N$  is a small number,  $(\sigma_z \gg \sigma_\theta)$  - some poor quality measurements and  $\theta_N \approx m_\theta$ .

2.  $N \rightarrow \infty$ ,  $(\sigma_z \ll \sigma_\theta)$  - some good quality measurements and  $\theta_N \approx \frac{\sum_{n=1}^N w_n u_n}{\sum_{n=1}^N u_n^2}$ .

**Example 2.9 – Estimation of the system parameter with noisy measurements: Maximum Likelihood Method.** Parameter  $\theta$  of the system was measured with some noise. Probability density function of

the noise is described by  $f_z(z) = \begin{cases} 1 + \frac{z}{2} & z \in [-2, 0] \\ 0 & z \notin [-2, 0] \end{cases}$ . It is known that the noise is additive, which

means that  $v_n = \theta + z_n$  for  $n=1,2$ . The measurements of noised value of the parameter  $\theta$  are  $v_1 = 1$  and  $v_2 = 2$ . **The problem:** determination of the parameter  $\theta$  estimator for given measurements and probability density function.

**The solution:**

In the first step we have to decide which method can be applied in this example. Let us calculate mean value of variable  $z$  :

$$\begin{aligned} m_z &= \int_{-2}^0 f_z(z)z dz = \int_{-2}^0 \left(1 + \frac{z}{2}\right)z dz = \int_{-2}^0 \left(z + \frac{z^2}{2}\right) dz = \\ &= \left. \frac{z^2}{2} + \frac{z^3}{6} \right|_{-2}^0 = -\frac{2}{3} \end{aligned} \quad (2.117)$$

Mean value of  $z$  is not equal 0. We are limited to methods such as Maximum Likelihood Method (MLM). Let us check some conditions for MLM:

d) Measurement system is described by  $v_n = \theta + z_n$  which is one-to-one invertible function;

e) Probability density function for  $f_z(z)$  is given.

Let us denote  $v_n = \theta + z_n$  in general form:

$$v_n = h(\theta, z_n). \quad (2.118)$$

Next, we have to calculate the probability density function  $f_v(v, \theta)$  of observed variable  $v$ . To end it, we can apply the following formula:

$$f_v(v, \theta) = f_z(h_z^{-1}(\theta, v)) \cdot |J_h|, \quad (2.119)$$

where  $J_h$  is Jacobi matrix of the inverse transformation.

In our example  $f_v(v, \theta)$  has form:

$$f_v(v, \theta) = f_z(v_n - \theta) \cdot |J_h|. \quad (2.120)$$

Jacobi matrix can be determined using the following formula:

$$J_h = \frac{\partial h_z^{-1}(\theta, v)}{\partial v} = \frac{d}{dv}(v_n - \theta) = 1. \quad (2.121)$$

Finally, we can write the probability density function of  $f_v(v, \theta)$  as:

$$f_v(v, \theta) = f_z(v_n - \theta). \quad (2.122)$$

Taking into account the probability density function of  $f_z(z)$ , which is given, we can rewrite the formula (2.122) in the form:

$$f_v(v, \theta) = \begin{cases} 1 + \frac{v_n - \theta}{2} & \text{for } \forall n \quad (v_n - \theta) \in [-2 \quad 0] \\ 0 & \text{otherwise} \end{cases} \quad (2.123)$$

Now we rewrite (2.123) in the form:

$$f_v(v, \theta) = \begin{cases} 1 + \frac{v_n - \theta}{2} & \text{for } \forall n \quad \theta \in [v_n \quad v_n + 2] \\ 0 & \text{otherwise} \end{cases} \quad (2.124)$$

The maximum likelihood function has form:

$$L_N(V_N, \theta) = \prod_{n=1}^N f_v(v_n, \theta) = \prod_{n=1}^N f_z(h_z^{-1}(\theta, v_n)) \cdot |J_h|, \quad (2.125)$$

where  $V_N = [v_1 \quad v_2 \quad \dots \quad v_N]$ . For our example formula (2.125) can be written as:

$$L_N(V_N, \theta) = \begin{cases} \left(1 + \frac{v_1 - \theta}{2}\right) \left(1 + \frac{v_2 - \theta}{2}\right) & \text{for } \theta \in [v_1 \quad v_1 + 2] \wedge [v_2 \quad v_2 + 2] \\ 0 & \text{otherwise} \end{cases} \quad (2.126)$$

Taking into account  $v_1 = 1$  and  $v_2 = 2$ , we can rewrite (2.126):

$$L_N(V_N, \theta) = \begin{cases} \left(1 + \frac{1-\theta}{2}\right) \left(1 + \frac{2-\theta}{2}\right) & \text{for } \theta \in [1 \ 1+2] \wedge [2 \ 2+2]. \\ 0 & \text{otherwise} \end{cases} \quad (2.127)$$

Let us rewrite (2.127) in the form:

$$L_N(V_N, \theta) = \begin{cases} \left(\frac{3-\theta}{2}\right) \left(2 - \frac{\theta}{2}\right) & \text{for } \theta \in [2 \ 3]. \\ 0 & \text{otherwise} \end{cases} \quad (2.128)$$

After a few simple transformations we obtain:

$$L_N(V_N, \theta) = \begin{cases} \frac{\theta^2}{4} - \frac{7}{4}\theta + 3 & \text{for } \theta \in [2 \ 3]. \\ 0 & \text{otherwise} \end{cases} \quad (2.129)$$

Now, we can formulate an optimization problem:

$$\theta_N \rightarrow L_N(V_N, \theta_N) = \max_{\theta \in \Theta} L_N(V_N, \theta). \quad (2.130)$$

In order to find  $\theta_N$ , we have to solve the following equation:

$$\left. \frac{dL_N(V_N, \theta)}{d\theta} \right|_{\theta=\theta_N} = 0. \quad (2.131)$$

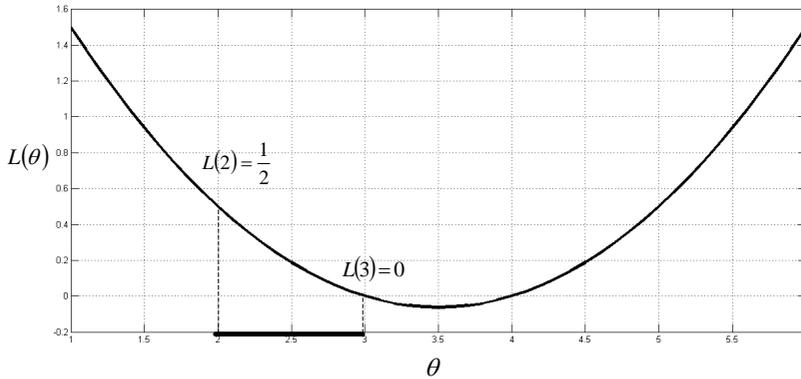
Substituting (2.129) in (2.131) we obtain:

$$\left. \frac{d}{d\theta} \left( \frac{\theta^2}{4} - \frac{7}{4}\theta + 3 \right) \right|_{\theta=\theta_N} = 0. \quad (2.132)$$

Before we solve the problem (2.132), we have to take into account that it is the optimization problem with some inequality constraints (2.129). Let us solve the problem graphically. The likelihood function, for example, is shown in Figure 2.18. It can be seen that this is a quadratic function, which has the minimal value (without taking constraints into account):

$$\begin{aligned} \frac{\theta_N}{2} - \frac{7}{4} &= 0 \\ \theta_N &= \frac{7}{2} \end{aligned} \quad (2.132)$$

It is worth stressing that the considered function is convex and monotonic decreasing at the left on the minimal point and monotonic increasing at the right side on minimal point (Figure 2.18).



**Figure 2.18. Likelihood function in solved example.**

Our problem is to find the maximum value of the likelihood function  $L(\theta_N)$  for given range  $\theta_N \in [2 \ 3]$ . Because the considered function is monotonic decreasing at the left on the minimal point and we know that minimal point is outside the range  $\theta_N \in [2 \ 3]$ , it means that we may find the proper value of  $\theta_N$  in point  $\theta_N = 2$  or  $\theta_N = 3$ . After the function's analysis in Figure 2.18, we can say that the solution of a given problem is the point  $\theta_N = 2$  where  $L(2) = \frac{1}{2}$ .

### Exercises

**Exercise 2.6.** Parameter  $\theta$  was measured with additive noise described by Gaussian function with mean  $m_z = 0$  and variance  $\sigma_z$ . **The problem:** determination of Bayesian estimator of  $\theta$  parameter for loss function  $L(\theta, \bar{\theta}) = -\delta(\theta - \bar{\theta})$ . Probability density function characterising  $\theta$  parameter population has the following form:  $f_\theta(\theta) = \lambda e^{-\lambda\theta}$ .

**Exercise 2.7. The problem:** To determine an algorithm based on the Maximum Likelihood Method, to determine estimation of the  $\theta$  parameter for  $N$  noised measurements which are additive:  $v_n = \theta + z_n$ , for  $n = 1, 2, \dots, N$ . It is assumed that the probability density function is Gaussian with mean equal  $m_z$  and variance  $\sigma_z$ .

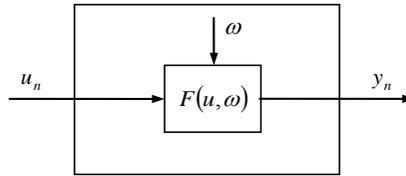
**Exercise 2.8.** For some fixed sequence of inputs  $u_1, u_2, \dots, u_N$  for linear system with characteristic  $y = \theta u$ , it was given following sequence of outputs  $w_1, w_2, \dots, w_N$ .

**The problem:** to determine the estimator of  $\theta$  parameter for the additive noise ( $v_n = \theta + z_n$ ) and

$$\text{the probability density function } f_z(z) = \begin{cases} 1 & \text{for } z \in [0 \ 1] \\ 0 & \text{for } z \notin [0 \ 1] \end{cases}$$

## 2.4 The choice of the Best Model: Probabilistic Case

In this section the choice of the best model problem in probabilistic conditions is considered. We start with the explanation why for the same value of input  $u_n$  different values of output  $y_n$  are collected. It can be caused by immeasurable variable  $\omega$  which is changing randomly. It is illustrated in Figure 2.19:



**Figure 2.19. System identification with random immeasurable variable.**

Consequently, it can be assumed that the investigated system is described by a set of characteristics indexed by random value  $\omega$  (Bub80). These problem is pictured in Figure 2.20.

**The problem formulation:**

The system's characteristic:  $\underline{y} = F(\underline{u}, \underline{\omega})$ ;

$\omega_n$  – value of random variable – system parameter,  $\underline{\omega}$  ( $\omega \in \Omega \subseteq \mathcal{R}^L$ ),  $f_\omega(\omega)$ ;

$y_n$  – value of random variable  $y$  – transformation of variable  $\underline{\omega}$ ;  $y \in \mathcal{Y} \subseteq \mathcal{R}^L$ ;

$F$  – invertible function with respect  $\omega$ :  $\underline{\omega} = F_\omega^{-1}(\underline{u}, \underline{y})$ ;

$u_n$  – value of random variable  $\underline{u}$ ,  $f_u(u)$ ;  $u \in \mathcal{U} \subseteq \mathcal{R}^S$ ;

$f_y(y | u)$  – Probability density function  $\underline{y}$  on condition that  $\underline{u} = u$ :  $f_y(y | u) = f_\omega(F_\omega^{-1}(u, y)) \cdot |J_F|$ ,  
 where  $J_F = \frac{\partial F_\omega^{-1}(u, y)}{\partial \underline{y}}$ .

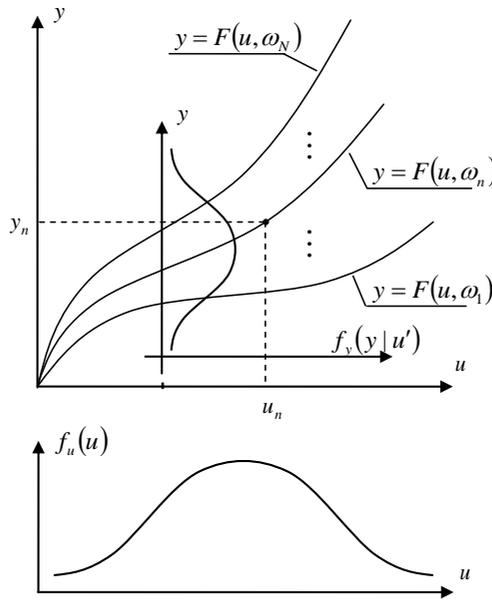


Figure 2.20. Static characteristic of SISO system with immeasurable variable.

**Example 2.10 – Regression of the type I.** The problem: to determine regression of the type I, for

given probabilistic density function:  $f(u, y) = \begin{cases} \frac{y}{u^2} & \text{for } u \in [0, 2], y \in [0, u] \\ 0 & \text{otherwise} \end{cases}$ .

**The solution:**

In order to solve regression of problem I, we have to find the best model by use the following equation:

$$\bar{y} = \Phi^*(u) = E[\underline{y} | \underline{u} = u] = \int_{\mathcal{Y}} y f(y | u) dy. \quad (2.137)$$

The problem is illustrated in the Figure 2.21.

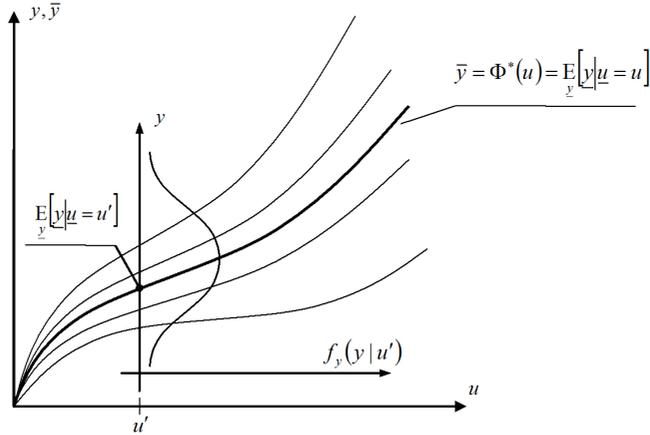


Figure 2.21. Regression of the type I.

First, we have to find the probability density function  $f_u(u)$ :

$$f_u(u) = \int_0^u f(u, y) dy = \int_0^u \frac{y}{u^2} dy = \frac{y^2}{2u^2} \Big|_0^u = \frac{1}{2}. \quad (2.138)$$

Then, we determine the conditional probability density function  $f_y(y|u)$  from the connection:

$$f_y(y|u) = \frac{f(u, y)}{f_u(u)} = \frac{\frac{y}{u^2}}{\frac{1}{2}} = \frac{2y}{u^2}. \quad (2.139)$$

Finally, we can solve some regression of the type I problem applied equation (2.137):

$$\bar{y} = \Phi^*(u) = E_y[y|u = u] = \int_{\mathcal{Y}} y f_y(y|u) dy = \int_0^u \frac{2y}{u^2} y dy = \frac{2y^3}{3u^2} \Big|_0^u = \frac{2u^3}{3u^2} = \frac{2}{3} u. \quad (2.140)$$

It means that the best model for given conditions is  $\bar{y} = \frac{2}{3} u$  for  $u \in [0 \quad 2]$ .

**Example 2.11 – Regression of the type II. The problem:** to determine regression of the type II for

given conditional probability density function:  $f(u, y) = \begin{cases} \frac{y}{u^3} & \text{for } u \in [0, 2], y \in [0, u^2] \\ 0 & \text{otherwise} \end{cases}$

Assume: model  $\bar{y} = \Phi(u, \theta) = \theta^{(1)}u + \theta^{(0)}$  and the quadratic performance index

$$Q(\theta) = \int \int_{\mathcal{U} \times \mathcal{Y}} [y - \Phi(u, \theta)]^2 f(u, y) dy du.$$

**The solution:**

Taking:

$$f(u, y) = f_y(y | u) f_u(u), \quad (2.141)$$

we can rewrite proposed quadratic performance index as:

$$Q(\theta) = \int_{\mathcal{U}} [y - \Phi(u, \theta)]^2 f_y(y | u) f_u(u) du. \quad (2.142)$$

In order to find an optimal value of the  $\theta$ , we have to solve the following optimization problem (see lecture 10):

$$\theta^* \rightarrow \min_{\theta} Q(\theta) = \min_{\theta} \int_{\mathcal{U}} (\Phi^*(u) - \Phi(u, \theta))^2 f_u(u) du. \quad (2.143)$$

We have to start with solving regression of the type I problem ( $\Phi^*(u)$ ). To be patterned on calculations from example 2.10, first we have to find the probability density function:

$$f_u(u) = \int_0^{u^2} f(u, y) dy = \int_0^{u^2} \frac{y}{u^3} dy = \left. \frac{y^2}{2u^3} \right|_0^{u^2} = \frac{u}{2}. \quad (2.144)$$

Then, the conditional probability density function (see 2.141) must be determined:

$$f_y(y | u) = \frac{f(u, y)}{f_u(u)} = \frac{\frac{y}{u^3}}{\frac{u}{2}} = \frac{2y}{u^4}, \quad (2.145)$$

and

$$\bar{y} = \Phi^*(u) = E[y | u = u] = \int_{\mathcal{Y}} y f(y | u) dy = \int_0^{u^2} \frac{2y}{u^4} y dy = \left. \frac{2y^3}{3u^4} \right|_0^{u^2} = \frac{2u^6}{3u^4} = \frac{2}{3} u^2. \quad (2.146)$$

Then, we can rewrite the performance index  $Q(\theta)$  in the form:

$$Q(\theta) = \int_{\mathcal{U}} \left( \frac{2}{3}u^2 - \Phi(u, \theta) \right)^2 f_u(u) du. \quad (2.147)$$

Taking into account, that  $f_u(u) = \frac{u}{2}$  and  $\Phi(u, \theta) = \theta^{(1)}u + \theta^{(0)}$ , we can rewrite (2.147) in the form:

$$Q(\theta) = \int_{\mathcal{U}} \left( \frac{2}{3}u^2 - \theta^{(1)}u - \theta^{(0)} \right)^2 \frac{u}{2} du. \quad (2.148)$$

In order to solve the optimization problem (2.147) we have to consider the following system of equations:

$$\nabla_{\theta} Q(\theta) \Big|_{\theta=\theta^*} = 0_2. \quad (2.149)$$

Substituting (2.148) in (2.149) and taking into account that  $\mathcal{U} = \{u \in \mathcal{R} : 0 \leq u \leq 2\}$ , we obtain:

$$\nabla_{\theta} Q(\theta) = \nabla_{\theta} \int_0^2 \left( \frac{2}{3}u^2 - \theta^{(1)}u - \theta^{(0)} \right)^2 \frac{u}{2} du \Big|_{\substack{\theta^{(0)}=\theta^{*(0)} \\ \theta^{(1)}=\theta^{*(1)}}} = 0_2. \quad (2.150)$$

In order to find the optimal value of parameters vector  $\theta$ , we have to solve the following set of equations (2.150):

$$\frac{\partial Q(\theta)}{\partial \theta^{(0)}} = 2 \int_0^2 \left( \frac{2}{3}u^2 - \theta^{*(1)}u - \theta^{*(0)} \right) (u)(-1) du = 0 \quad (2.151)$$

$$\frac{\partial Q(\theta)}{\partial \theta^{(1)}} = 2 \int_0^2 \left( \frac{2}{3}u^2 - \theta^{*(1)}u - \theta^{*(0)} \right) (u)(-u) du = 0$$

After a few transformations, we obtain the following equations:

$$\int_0^2 \left( -\frac{2}{3}u^3 + \theta^{*(1)}u^2 + \theta^{*(0)}u \right) du = 0, \quad (2.152)$$

And

$$\int_0^2 \left( -\frac{2}{3}u^4 + \theta^{*(1)}u^3 + \theta^{*(0)}u^2 \right) du = 0. \quad (2.153)$$

Let us now focus on a connection (2.152):

$$\begin{aligned} \int_0^2 \left( -\frac{2}{3}u^3 + \theta^{*(1)}u^2 + \theta^{*(0)}u \right) du &= 0 \\ -\frac{2}{3} \frac{u^4}{4} \Big|_0^2 + \theta^{*(1)} \frac{u^3}{3} \Big|_0^2 + \theta^{*(0)} \frac{u^2}{2} \Big|_0^2 &= 0 \end{aligned} \quad (2.154)$$

Then, we obtain:

$$\theta^{*(0)} = \frac{4}{3} - \frac{4}{3}\theta^{*(1)}. \quad (2.155)$$

In the next step, we determine  $\theta^{*(1)}$  from the equation (2.153):

$$-\frac{2}{3} \frac{u^5}{5} \Big|_0^2 + \theta^{*(1)} \frac{u^4}{4} \Big|_0^2 + \theta^{*(0)} \frac{u^3}{3} \Big|_0^2 = 0, \quad (2.156)$$

and then we obtain:

$$\theta^{*(1)} = \frac{16}{15} - \frac{2}{3}\theta^{*(0)}. \quad (2.157)$$

Finally, taking (2.155) and (2.157), we obtain such optimal values of parameters vector  $\theta$  as:

$$\theta^{*(0)} = -\frac{4}{5}, \theta^{*(1)} = \frac{8}{5}. \quad (2.158)$$

It means, that the best model as regression of the type II for given system is  $\bar{y} = \Phi(u, \theta) = \frac{8}{5}u - \frac{4}{5}$ .

Verification of the solution:

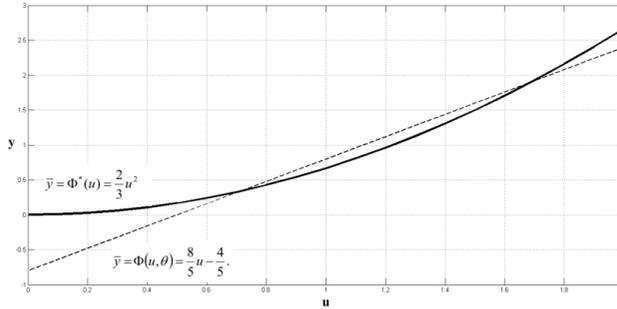


Figure 2.22. Results for determination of regression of the type II.

### Exercises

**Exercise 2.9.** Determine regression of the type I for given probabilistic density function:

$$f(u, y) = \begin{cases} 1 & \text{for } u \in [0, 2], y \in [u^2, u^2 + 1] \\ 0 & \text{otherwise} \end{cases}$$

**Exercise 2.10.** Determine regression of the type II for given conditional probability density function:

$$f(y, u) = \begin{cases} 2 & \text{for } u \in [0, 1], y \in [1-u, 1] \\ 0 & \text{otherwise} \end{cases}. \text{ Assume: model } \bar{y} = \Phi(u, \theta) = \theta^{(1)}u + \theta^{(0)} \text{ and}$$

$$\text{quadratic performance index } Q(\theta) = \int \int_{\mathcal{U} \times \mathcal{Y}} [y - \Phi(u, \theta)]^2 f(u, y) dy du.$$

## 2.5 Example Application of the System Identification for Real Life Problem

One of the most popular model in the system identification is Takagi-Sugeno system (T-S). In this system we have a sequence of *If...Then* sentences (*antecedents*) and affine linear functions (*consequents*). For example, one rule of T-S can be written as follows (Nel01, Bab03):

$$R_k : \text{IF } u \text{ is } A_k \text{ THEN } y_k = \theta_k^T u + \theta_k^{(0)}, \quad (2.159)$$

where  $A_k$  are some antecedent linguistic terms such as “small”, “large” and  $k = 1, 2, \dots, K$  is the number of rules. Takagi-Sugeno model combines the linguistic description with some regression functions. In order to determine the input of the system, we usually use the following formula:

$$y = \frac{\sum_{k=1}^K \beta_k(u) y_k}{\sum_{k=1}^K \beta_k(u)}, \quad (2.160)$$

where  $\beta_k(u)$  is the degree of fulfillment for  $k$ -th rule. Taking into account that  $y_k = \theta_k^T u + \theta_k^{(0)}$ , we can rewrite above formula as:

$$y = \frac{\sum_{k=1}^K \beta_k(u) (\theta_k^T u + \theta_k^{(0)})}{\sum_{k=1}^K \beta_k(u)}. \quad (2.161)$$

**Example 2.12 – Simple example.**

Let us propose T-S system with the following set of rules:

$$\begin{aligned} R_1 : \text{IF } u \text{ is } \textit{Large} \text{ THEN } y_1 &= \theta_1 u + \theta_1^{(0)} \\ R_2 : \text{IF } u \text{ is } \textit{Normal} \text{ THEN } y_2 &= \theta_2 u + \theta_2^{(0)}. \\ R_3 : \text{IF } u \text{ is } \textit{Small} \text{ THEN } y_3 &= \theta_3 u + \theta_3^{(0)} \end{aligned} \quad (2.162)$$

and

$$y = \frac{\beta_{\textit{Large}}(u)(\theta_1 u + \theta_1^{(0)}) + \beta_{\textit{Normal}}(u)(\theta_2 u + \theta_2^{(0)}) + \beta_{\textit{Small}}(u)(\theta_3 u + \theta_3^{(0)})}{\beta_{\textit{Large}}(u) + \beta_{\textit{Normal}}(u) + \beta_{\textit{Small}}(u)}. \quad (2.163)$$

At the computational level, rules in the form (2.163) are not easy to implement. In a real life applications, in order to describe the antecedent linguistic terms, we usually use some smooth function, such as Gaussian function. This function in Takagi-Sugeno models is called *membership function*. It has the following form:

$$\mu_k^{(s)}(u^{(s)}; m_k^{(s)}, \sigma_k^{(s)}) = \exp\left(-\frac{(u^{(s)} - m_k^{(s)})^2}{(\sigma_k^{(s)})^2}\right), \quad (2.164)$$

where  $m_k^{(s)}$  is the mean value for  $s$ -th input and  $k$ -th rule and  $\sigma_k^{(s)}$  is the standard deviation for  $s$ -th input and  $k$ -th rule of Gaussian function. Taking into account (2.164), we can rewrite (2.161) as:

$$y = \frac{\sum_{k=1}^K \exp\left(-\sum_{s=1}^S \frac{(u^{(s)} - m_k^{(s)})^2}{(\sigma_k^{(s)})^2}\right)}{\sum_{k=1}^K \exp\left(-\sum_{s=1}^S \frac{(u^{(s)} - m_k^{(s)})^2}{(\sigma_k^{(s)})^2}\right)} (\theta_k^T u + \theta_k^{(0)}). \quad (2.165)$$

**Example 2.13 – Takagi-Sugeno model.** In the figure below some data set is shown. Our task is to find an estimation of the Takagi-Sugeno model's parameters, based on given measurements.

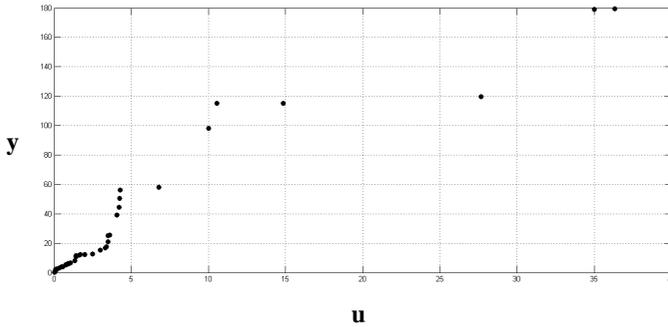


Figure 2.23. Data set for example 2.13.

At the beginning, let us set Takagi-Sugeno system with only 2 rules. The results are shown in the figure below:

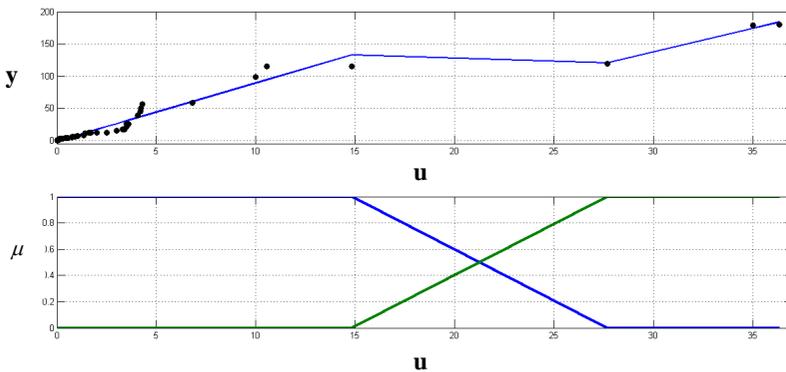


Figure 2.24. The results for 2 rules.

In the next figure, results obtained from 5 rules are presented:

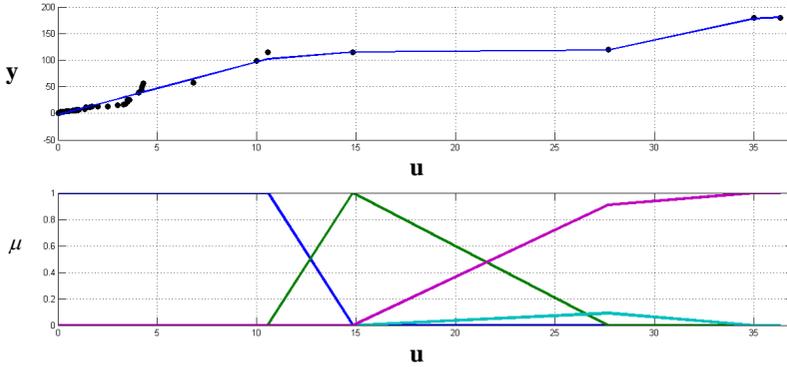


Figure 2.25. The results for 5 rules.

In the last figure, the results obtained from 11 rules are illustrated:

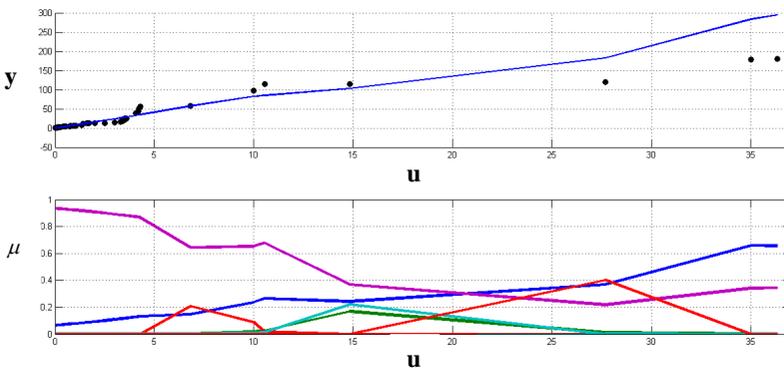


Figure 2.26. The results for 11 rules.

## Bibliography

[Bab03] BABUSKA R., VERBRUGGEN H. „Neuro-fuzzy methods for nonlinear system identification, Annual Reviews in Control 27, 2003, pp. 73–85.

[Bis07] BISHOP C.M. „Pattern Recognition and Machine Learning, Springer 2007.

[Bub80] BUBNICKI Z. „Identification of Control Plants, PWN-Elsevier, Warszawa-Amsterdam-Oxford-New York 1980.

[Bro06] BRONSON R., COSTA G.B. „Differential equations, Schaum’s Outline Series, McGraw-Hill Companies, Inc. 2006.

[Che07] CHERKASSKY V., MULIER F. „Learning from data: concepts, theory, and methods, Wiley, New Jersey 2007.

[Eyk74] EYKHOFF P. „System Identification, Parameter and State Estimation, Wiley, New York-London-Sydney 1974.

[Lju99] LJUNG L. „System identification: Theory for the user, Prince Hall, Englewood Cliffs, New Jersey 1999.

[Nel01] NELLES O. „Nonlinear system identification. From classical approach to neural networks and fuzzy models, Springer, Berlin, Heidelberg, New York 2001.

[Oga02] OGATA K. „Modern control engineering, Upper Saddle River, Prentice Hall 2002.

[Vuk03] VUKIC Z., KULJACA L., DONLAGIC D., TESNJAK S., Nonlinear Control Systems, Marcel Dekker, Inc. 2003.

[Świa09]