Chapter 1

Introduction

The booklet at hand is a complement to the book "Systems biology: Mathematical Modeling and Model Analysis". In the book many examples are supported by numerical simulations and by additional exercises. Simulations were done with help of Matlab, a user friendly tool to do mathematical calculations. The intention here is **not** to give an introduction into Matlab; instead we describe here the general structure of our files in such a way that it might help students to become familiar with the model and allow to do their own simulation studies. Special toolboxes are not needed to perform the simulations. Although there are toolboxes for systems biology available we decided to use a simple .m file structure that can be started immediately on the computer. There is only one program we took from the web which was necessary to solve mixed integer optimization programs. Matlab itself offers no simple alternative in this case.

Exercises are used to deepen the methods introduced in the main text. Some solutions are already given; here one will find the complete calculations for all exercises supported again by numerical simulations. Most of the exercises provided were also used for exams during my lectures on modeling and model analysis in Munich. The set up of mathematical models for biochemical networks require that the student is familiar with different approaches and knows which approach can be applied for a given problem. Having meaningful models available, techniques to analyze the models can be used to better understand the system at hand.

The book follows the didactic idea starting from the basic structure of the equations, that is, in the deterministic case, a balance equation and then focusing on the individual processes like enzymatic conversion, polymerisation, and signaling. Having chosen a set of kinetic expressions and kinetic parameters the system is completely determined. In the deterministic case, the basic structure for intracellular metabolites looks like:

$$\dot{\underline{c}} \;=\; N\,\underline{r} \;-\; \mu\,\underline{c}, \tag{1.1}$$

with the specific growth rate μ. By putting in kinetic expressions in the rates $\underline{r}(\underline{c})$ in Equation (1.1) one obtains a non-linear differential equation system in the form:

$$\dot{\underline{c}} \;=\; \underline{f}(\underline{c})\,. \tag{1.2}$$

Based on this formal representation given in Equation (1.2) a number of

tools to analyze systems are described in the book. This part of the book covers time hierarchies, sensitivities, robustness, control coefficients within Metabolic Control Analysis, possibilities within Biochemical Systems Theory, structured kinetic model analysis, and control theoretical methods.

The last part deals with large networks and their properties. Here, kinetic aspects are not considered since information on a large number of reaction steps can hardly be found. Therefore, applications in biotechnology like metabolic flux analysis and topological aspects are considered in this part.

In chapter 2 an introduction into our Matlab files is given. In chapters 3 to 11 – corresponding to the chapter numbers of the main text – the solution of all exercises as well as numerical studies are provided. Although we have controlled all single steps of the calculations there may still be errors. If this is the case or if you have questions to the material provided please contact us.

Chapter 2

Matlab programs

Matlab (abbreviation of Matrix Laboratory) is a programming environment that allows data analysis, visualization and numeric computation. Matlab is widely used in academic field as a useful and powerful tool for research and development. There are some basic concepts one should know before start working with Matlab. These concepts comprise understanding what variables and functions are and how they are defined and can be used.

Variable: a variable allows basically to save data and perform operations. Defining a variable implies to assign data to a string:

```
variable1=4; variable2=3;.
```

In this case we created two variables named variable1 and variable2 which contain the numbers 34 and 0 respectively. You can define a new variable which stores the result of any operation between existing variables:

```
variable3=variable1*variable2;
```

In this case variable3 is equal to 12.

Scripts: Scripts are the simplest kind of program file because they have no input or output arguments. They contain series of commands, such as computations. Scripts are saved in m-Files.

Function: a function is an m-file which contains lines of commands. A function requires normally an input in order to generate an output. To declare a function one has to save the function code in a text file with an .m extension (m-file). The name of the file should match the name of the first function in the file.

function [y1,...,yN] = myfun(x1,...,xM) declares a function named myfun that accepts inputs x1,...,xM and returns outputs y1,...,yN. This declaration statement must be the first executable line of the function. You can declare multiple functions within the same file.

Variables within a function: Ordinarily, each function has its own local variables, which are separate from those of other functions, and from those of the base workspace. Variables of a function are not shown in the Workspace when running it. However variables of a function can be visible in the Workspace by saving them in a binary file (mat-file) and then loading the file in the Workspace. One can store only determinate variables or store all variables by using the command save(filename, variables) or save(filename) respectively.

Command Window

Lets you interactively enter data, execute commands and programs, and display results. Here you can execute m-files by typing its name and then pressing the enter key.

Structure of the provided m-files

For a better understanding of certain concepts explained in this book, some m-files are provided. These files are built with a common structure and contain three principal sections:

Parameters: In this section, the set of simulation conditions are defined. These parameters can be defined as individual variables, as a single vector or as a single variable containing all parameters. The two last options are preferred when working with a big set of parameters as input arguments of a function. By modifying the desired parameters, own simulation studies can be done.

Calculations: Contains command lines that enable calculations between different variables. Commands typically used here include logical operators, control flow statements (if/elseif/else, for, while, switch/case/otherwise, etc) and arithmetic Operators ($+ \; - \; * \; /$ (division) and $\wedge$ (for power)). This part of the code contains also explanatory text in form of comments that guides you through the calculations. Comments are preceded by the percent character and are ignored by Matlab.

Plotting: This section contains the necessary command lines that allow a graphical representation of the calculations performed by the m-file. It usually contains the command plot(x,y) and other command lines that specify line styles and colors, legend and axis labels.

Running an m-file: There are basically two ways of running an m-file: by simply typing filename at the Matlab command window and pressing the enter key or by clicking the “Run”-button. If you don’t want to run the whole m-file, you can just select the part of the m-file that you want to run and press the F9 key or click the “Run section”-button.

Chapter 3

Exercise chapter 3

EXERCISE 3.1 *Master equation.*

Consider a simple model of DNA binding where an inactive binding site D_i or the active D_a form is available. The RNA polymerase can change the state from inactive to active. As the polymerase is not explicitly modeled, one obtains the following reaction scheme for which a master equation is to be determined and analyzed:

$$D_i \overset{k_1,k_2}{\rightleftharpoons} D_a\,, \tag{3.1}$$

a For the process for which a component is increased and respectively decreased the driving force is the component itself. For the given reaction equation, determine the transitions for the variable D_i. Note that the total amount is constant: $n_T = D_i + D_a$. Complete the table the corresponding driving forces and the kinetic parameter for the individual cases.

	Case 1	Case 2	Case 3	Case 4
State D_i	$\#D_i = n - 1 \to n$			
D_a	$\#D_a = m + 1 \to m$			
Driving force	$\#D_a = m + 1$			
Parameter	k_2			

Using the entries of the table, a master equation for $P_n(D_i)$ can be determined.

b Simplify the master equation for the case $n_T = 1$. Eliminate the corresponding terms from the master equation.

c Determine the solutions of the master equation for $P_0(D_i)$ with the initial conditions: (i) $P_0(D_i(t = 0)) = 1$ and (ii) $P_0(D_i(t = 0)) = 0$. Illustrate graphically the significance of initial conditions.

Show that the solutions for a large time are stationary and independent of the initial conditions.

d Determine the mean value and variance of the steady state. Compare with the deterministic solution.

SOLUTION:

Due to the fact that the sum of both variables is constant the system is

a one dimensional one. Considering $\#D_i$ as our variable we have to consider four cases: two cases (1 and 3) that generate new molecules and two cases (2 and 4) were the amount is degraded by one molecule.

	Case 1	Case 2
State D_i	$\#D_i = n-1 \to n$	$\#D_i = n \to n+1$
D_a	$\#D_a = m+1 \to m$	$\#D_a = m \to m-1$
Driving force	$\#D_a = m+1$	$\#D_a = m$
Parameter	k_2	k_2
	Case 3	Case 4
State D_i	$\#D_i = n+1 \to n$	$\#D_i = n \to n-1$
D_a	$\#D_a = m-1 \to m$	$\#D_a = m \to m+1$
Driving force	$\#D_i = n+1$	$\#D_i = n$
Parameter	k_1	k_1

For the individual propensities we get after taking into account the conservation equation $n_T = n + m$:

$$\frac{P_{n-1,n}}{\Delta t} = k_2\,(m+1) = k_2\,(n_T - n + 1); \quad \frac{P_{n+1,n}}{\Delta t} = k_1\,(n+1)\,; \tag{3.2}$$

$$\frac{P_{n,n+1}}{\Delta t} = k_2\,m = k_2\,(n_T - n); \qquad \frac{P_{n,n-1}}{\Delta t} = k_1\,n\,. \tag{3.3}$$

b The following two differential equations are obtained:

$$\dot{P}_1(D_i) = k_2\,P_0 - k_1\,P_1 \tag{3.4}$$

$$\dot{P}_0(D_i) = k_1\,P_1 - k_2\,P_0 \tag{3.5}$$

c The initial conditions must be interpreted as follows: for $t = 0$ the system is completely in one state, either inactive or active. Taking this into account we get one differential equation for D_i:

$$\dot{P}_0(D_i) = k_1 - (k_1 + k_2)\,P_0 \tag{3.6}$$

with the solution:

$$P_0(D_i) = \frac{k_1}{k_1 + k_2} + C_0\,e^{-(k_1+k_2)\,t} \tag{3.7}$$

with C_0 is obtained from initial conditions. For the two cases we get:

$$\text{case 1:} \quad P_0(D_i) = \frac{k_1}{k_1 + k_2} + \frac{k_2}{k_1 + k_2}\,e^{-(k_1+k_2)\,t} \tag{3.8}$$

$$\text{case 2:} \quad P_0(D_i) = \frac{k_1}{k_1 + k_2} + \frac{k_1}{k_1 + k_2}\,e^{-(k_1+k_2)\,t}\,. \tag{3.9}$$

For large values of t the solution is independent from the initial value and the same result is obtained.

d To calculate mean and variance standard formulas can be used: For the mean we get:

$$< D_i > \quad = \quad \sum_{n=0}^{1} n\ P_n \ = \ 0\ P_0 \ + \ 1\, P_1 \ = \ \frac{k_2}{k_2 \ + \ k_1} \tag{3.10}$$

and for the variance:

$$<< D_i >> \quad = \quad < D_i^2 > \ - (< D_i >)^2 \tag{3.11}$$

with

$$< D_i^2 > \quad = \quad \sum_{n=0}^{1} n^2\ P_n \ = < D_i >; \tag{3.12}$$

using this part and inserting, we get:

$$<< D_i >> \quad = \quad \frac{k_2}{k_2 \ + \ k_1} \ - \ \left(\frac{k_2}{k_2 \ + \ k_1}\right)^2 \ = \ \frac{k_1\, k_2}{(k_2 \ + \ k_1)^2}\,. \tag{3.13}$$

The deterministic calculation leads to an equilibrium: $k_1\, D_i = k_2\, D_a$; with the conservation equation $D_0 = D_a + D_i = 1$: $D_i = D_0 k_2/(k_2 \ + \ k_1)$. This is the same result as above.

EXERCISE 3.2 *Balance equation.*

A biotechnological process is ongoing in a continuously running reactor with a volume V (Substrate S_i will be added at a rate of q_i^{in} and a concentration $c_{S,i}^{in}$). The biochemical network of the biomass shall be considered, where for every extracellular substrate an intracellular metabolite exists (M_i) and a single transport rate (r_{ai}) are taken into account (stoichiometry N_a). In total there are n substrates/metabolites and q reactions in the intracellular network (r_i, stoichiometry N_i) that taken into account. The metabolites M_i are precursors for the main components P_j of the biomass. The reaction equations for the synthesis are:

$$\text{Reaction j:} \quad \gamma_{1j}\ M_1 \ + \ \gamma_{2j}\ M_2 \ \cdots \quad \longrightarrow \quad P_j\,; \quad \text{rate: } r_j \tag{3.14}$$

a Determine the dimensions of the matrices and vectors of the system.

b Determine the corresponding differential equations for the individual components (substrate, biomass, intracellular metabolites and main components of the biomass).

c Determine the steady state of the whole system. Under what condition can the rates r_i in the network be directly read from the measurement data (uptake rates and biomass composition) ?

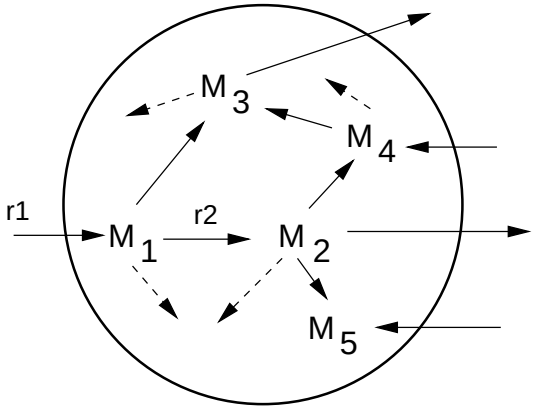

FIGURE 3.1: A network with uptake/production as well as intracellular rates (solid line) and drain (dashed line) to the main components of the biomass.

SOLUTION:

a Dimensions of vectors and matrices are as follows:

$$\begin{array}{lll} N_a \quad n \times n & N_i \quad n \times q & N_p \quad n \times p \\ \underline{r}_a \quad n \times 1 & \underline{r}_i \quad q \times 1 & \underline{r}_p \quad p \times 1 \end{array} \tag{3.15}$$

b We start be considering the individual components in the liquid phase: Balance equation for substrate i, $(i = 1, 2, \ldots n)$:

$$\dot{c}_{S,i} = \frac{q_i^{in}}{V} c_{S,i}^{in} - \frac{\sum_j q_j^{in}}{V} c_{S,i} - \gamma_{ai}\ r_{ai}\ mg_i\ c_B\,; \tag{3.16}$$

using vectors we get

$$\dot{\underline{c}}_S = \text{diag}\left(\frac{q_i^{in}}{V}\right) \underline{c}_S^{in} - \frac{\sum_j q_j^{in}}{V} \underline{c}_S - N_a\ \underline{r}_a^*\ c_B \tag{3.17}$$

with

$$\underline{r}_a^* = \underline{r}_a * \underline{mg} \qquad \text{(element-wise multiplication)} \tag{3.18}$$

Note that the transport rates r_{ai} have units $\frac{mol}{h\ g}$. They are also based on the cell dry mass.

Balance equation for biomass:

$$\dot{c}_B = \left(\mu - \frac{\sum q_i^{in}}{V}\right) c_B \tag{3.19}$$