Prof. Jörg Stelling
D-BSSE
ETH Center UNG H14
Universitätstrasse 19
CH-8092 Zürich
https://www.csb.ethz.ch

Computational Systems Biology
636-0007-00 U, Autumn 2025

Assignment 8

(Issue: 14-Nov-2025)

1 Sensitivity Analysis and Parameter Estimation Accuracy in a Simple Biochemical System

We want to deal with a simple biochemical system, namely the conversion of a substrate S into a product P:

$$S \xrightarrow{r} P$$

The reaction is assumed to follow irreversible Michaelis-Menten kinetics

$$r = r^{max} \cdot \frac{c_S}{K_m + c_S} \tag{1}$$

where the optimal parameters values determined from experimental data are $r^{max} = 1.0\,mM/s$ and $K_m = 100\,mM$.

For the given system, we want to perform the following analyses using MATLAB or Python:

a) Simulate the system behavior for a duration of $t = 1000\,s$ using $c_S = 200\,mM$ and $c_P = 0\,mM$ as initial values.

b) Calculate the local state sensitivities for this scenario. If we write the system of mass balance equations as

$$\frac{\partial \mathbf{c}}{\partial t} = f(\mathbf{c}, \mathbf{p}, t) \tag{2}$$

where $\mathbf{c} = (c_S, c_P)^T$ is the vector of substrate and product concentrations and $\mathbf{p} = (r^{max}, K_m)^T$ is the vector of parameters, the corresponding state sensitivities (direct method) are given by (See lecture 9 slide 35)

$$\frac{\partial}{\partial t}\frac{\partial \mathbf{c}}{\partial \mathbf{p}}\bigg|_{\mathbf{p}^*,t_i} = \mathbf{J_c} \cdot \frac{\partial \mathbf{c}}{\partial \mathbf{p}}\bigg|_{\mathbf{p}^*,t_i} + \mathbf{J_p} \tag{3}$$

with

$$\mathbf{J_c} = \frac{\partial f(\mathbf{c}, \mathbf{p}, t)}{\partial \mathbf{c}},\ \mathbf{J_p} = \frac{\partial f(\mathbf{c}, \mathbf{p}, t)}{\partial \mathbf{p}} \text{ and } \mathbf{p}^* \text{ as the vector of optimal parameter values.}$$

Equation (3) essentially defines a time-dependent ODE in $\frac{\partial \mathbf{c}}{\partial \mathbf{p}}$ that can be solved alongside the actual system of balance equations. This allows us to compute the sensitivity matrix $\mathbf{S}$ given by

$$s_{k,j} = \frac{\partial c_k}{\partial p_j}, \qquad k = 1 \ldots n_{states},\ j = 1 \ldots n_{pars} \tag{4}$$

in a single simulation run starting with initial values of zero for all $s_{k,j}$. Plot the dynamics of the $s_{k,j}$ into one graph alongside the concentration trajectories of c_S and c_P. In which region of c_S do the largest absolute sensitivity values occur for each parameter?

c) Using the resulting sensitivities from b), determine the Fisher Information Matrix

$$\mathbf{F}(\mathbf{p}^*) = \sum_{i=1}^{n_{meas}} \left[\frac{\partial \mathbf{c}}{\partial \mathbf{p}}\bigg|^T_{\mathbf{p}^*,t_i} \cdot \mathbf{C}^{-1}(t_i) \cdot \frac{\partial \mathbf{c}}{\partial \mathbf{p}}\bigg|_{\mathbf{p}^*,t_i} \right] \tag{5}$$

for the measured dataset provided in Table 1 below. We assume that the covariance matrix of the measurements, $\mathbf{C}(t_i)$, only consists of diagonal elements

$$c_{m,n}(t_i) = \begin{cases} \sigma_m^2(t_i) & \forall m = n \\ 0 & \forall m \neq n \end{cases} \tag{6}$$

with $m, n = 1 \ldots n_{states}$ and $\sigma_m^2(t_i)$ as the measurement variance of state m at time t_i. This implies that measurements for $c_S(t_i)$ and $c_P(t_i)$ are independent. For sake of simplicity, we assume that the measured standard deviations $\sigma_m(t_i)$ are constant over time and amount to $\sigma_S(t_i) = 2.5\,mM$ and $\sigma_P(t_i) = 5\,mM$, respectively.

d) The Fisher information matrix $\mathbf{F}(\mathbf{p}^*)$ ($dim(\mathbf{F}(\mathbf{p}^*)) = n_{pars} \times n_{pars}$) can be used to calculate a lower bound of the parameter estimation error $\sigma_{P,j}$ based on the Cramér-Rao inequality according to

$$\sigma_{P,j} \geq \sqrt{cr_{j,j}} \tag{7}$$

where $cr_{j,j}$ denotes the diagonal elements of the inverted Fisher Information Matrix

$$\mathbf{CR} = \mathbf{F}^{-1}(\mathbf{p}^*) \tag{8}$$

Please calculate the vectors σ_P for the dataset provided in Table 1. What results do you obtain?

e) How do the parameter estimation accuracies change when instead of the experimental setup chosen in dataset 1 (Table 1) you select the alternative datasets (Table 2 and 3) for parameter identification? Use the nominal concentrations provided in the table legends as initial values for the simulation and assume that the same standard deviations apply to the measurements as in part c). Which dataset apparently provides the most accurate parameter estimate?

Note:

All the data in the tables below are available as plain text files on the course homepage.

Some useful MATLAB commands and help entries for this exercise:

`inv, eig, ode45, odeset, function handles (@), plot`

Some useful Python commands and help entries for this exercise:

`scipy.linalg (inv, eigvals), scipy.integrate (odeint, solve_ivp), matplotlib.pyplot (plot, show)`

Submission:

Please address any questions to

alix.moawad@bsse.ethz.ch

Table 1

t_i	$c_S(t_i)$	$c_P(t_i)$
0	201.6	0.1
200	91.2	115.7
400	16.0	177.0
600	0.0	191.3
800	0.0	194.5
1000	0.0	205.9

Table 2

t_i	$c_S(t_i)$	$c_P(t_i)$
0	203.5	0.0
125	120.7	70.1
250	53.5	133.6
375	18.8	182.5
500	2.6	191.2
625	1.4	183.0
750	0.0	193.6
875	0.0	200.9
1000	0.1	211.9

Table 3

t_i	$c_S(t_i)$	$c_P(t_i)$
0	100.3	5.0
100	60.5	38.7
200	23.8	69.4
300	14.5	95.2
400	0.0	99.4
500	0.0	89.8

Table 1: Measured concentration time courses t_i (s) of c_S and c_P (mM). The nominal starting concentrations were $200\,mM$ S and $0\,mM$ P (Tables 1 and 2) and $100\,mM$ S and $0\,mM$ P (Table 3).