Bohl-Marek decomposition applied to a class of biochemical networks with conservation properties

Š. Papáček[1], *C. Matonoha*[2], *J. Duintjer Tebbens*[2,3]

[1] Institute of Information Theory and Automation of the CAS, Prague
[2] Institute of Computer Science of the CAS, Prague
[3] Faculty of Pharmacy, Charles University, Hradec Králové

1 Introduction

This study presents an application of one special technique, further called as Bohl-Marek decomposition, related to the mathematical modeling of biochemical networks with mass conservation properties. We continue in direction of papers devoted to inverse problems of parameter estimation for mathematical models describing the drug-induced enzyme production networks [3]. However, being aware of the complexity of general physiologically based pharmacokinetic (PBPK) models, here we focus on the case of enzyme-catalyzed reactions with a substrate transport chain [5]. Although our ultimate goal is to develop a reliable method for fitting the model parameters to given experimental data, here we study certain numerical issues within the framework of optimal experimental design [6]. Before starting an experiment on a real biochemical network, we formulate an optimization problem aiming to maximize the information content of the corresponding experiment. For the above-sketched optimization problem, the computational costs related to the *two formulations of the same biochemical network*, being (i) the classical formulation $\dot{x}(t) = Ax(t) + b(t)$ and (ii) the 'quasi-linear' Bohl-Marek formulation $\dot{x}_M(t) = M(x(t))\, x_M(t)$, can be determined and compared.

2 Problem formulation

The system of differential equations describing the processes under study is described in Tab. 1. It can be systematically derived using the so-called stoichiometric matrix $S \in \mathbb{R}^{n\times q}$, where q is the number of reactions (including the transport of species).

Table 1: Description of the transport and reaction processes defining the network.

Description of the related process	**Chem. notation**	**Param.**
R_0: Substrate X_{ext} dosing (model input)	$\emptyset \to X_{ext}$	$u(t)$
R_1: Substrate transport between compartments	$X_{ext} \rightleftharpoons X_{int}$	k_0
R_2: Enzyme E binds to substrate,	$X_{int} + E \rightleftharpoons C$	k_1
formation of a complex C		
R_3: Reverse reaction to R_2		k_{-1}
R_4: Complex breaks down into E plus	$C \to E + P$	k_2
a product P – altered substrate molecule		

The vector of changes in species concentrations $x \in \mathbb{R}^n$ is then described as a linear transformation of the reaction rate vector $\nu \in \mathbb{R}^q$:

$$\dot{x}(t) = S\, \nu(x, p), \tag{1}$$

where

$$S = \begin{array}{c} \begin{matrix} R_1 & R_2 & R_3 & R_4 \end{matrix} \\ \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \qquad \nu = \begin{pmatrix} k_0\ (x_1 - x_2) \\ k_1\ x_2\ x_3 \\ k_{-1}\ x_4 \\ k_2\ x_4 \end{pmatrix}, \qquad p = \begin{pmatrix} k_0 \\ k_1 \\ k_{-1} \\ k_2 \end{pmatrix}. \tag{2}$$

Reaction networks frequently possess subsets of reactants that remain constant at all times, i.e., they are referred to conserved species. Generally, there exists a conservation matrix Γ (with dimension $h \times n$), where the rows represent the linear combination of species (reactants), which are constant in time. It can be solved explicitly for large systems ($0 = \Gamma\ S$). For our case of S in form (2), the conservation property reads

$$x_3 + x_4 = e_0, \quad x_1 + x_2 + x_4 + x_5 = u_0. \tag{3}$$

Consequently, here

$$\Gamma = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{pmatrix}. \tag{4}$$

The existence of two relations (3) signifies not only the possibility to reduce the number of state variables, but also induces the reformulation of the governing equations for species concentration using negative M-matrices, see (9). For instance, using (2), we get the resulting ODE system in the usual form

$$\dot{x}(t) = Ax(t) + b(x(t)), \tag{5}$$

with the constant matrix (the linear part of the system)

$$A = \begin{pmatrix} -k_0 & k_0 & 0 & 0 & 0 \\ k_0 & -k_0 & 0 & k_{-1} & 0 \\ 0 & 0 & 0 & k_{-1} + k_2 & 0 \\ 0 & 0 & 0 & -(k_{-1} + k_2) & 0 \\ 0 & 0 & 0 & k_2 & 0 \end{pmatrix} \tag{6}$$

and the vector representing nonlinear, e.g. bilinear, parts

$$b(x(t)) = \begin{pmatrix} u(t) \\ -k_1 \cdot x_2(t) \cdot x_3(t) \\ -k_1 \cdot x_2(t) \cdot x_3(t) \\ k_1 \cdot x_2(t) \cdot x_3(t) \\ 0 \end{pmatrix}. \tag{7}$$

The initial conditions are

$$x(0) = \begin{pmatrix} u(t_0) & 0 & e_0 & 0 & 0 \end{pmatrix}^T. \tag{8}$$

The ODE system (5) is nonlinear because of the bilinear terms and time-varying dosing function $u(t)$. Nevertheless, thanks to the conservation properties (3), there exists an alternative, a quasi-linear approach representing (in some sense) linearization of originally non-linear system (5) with the block diagonal system matrix of a special form (negative M-matrix). However, the system matrix dimension (order) has to be bigger because of the repeated presence of some state

variables (as it is shown in the next section). To the best of our knowledge, this approach was proposed by Bohl and Marek [1, 2] and further extended into the control theory framework by Marek [4].

Theorem (Bohl-Marek decomposition): When the conservation equations of a system of ODEs contain all variables, then the system can be decomposed into coupled, quasi-linear subproblems.

Sketch of the proof: Knowing that all state variables are involved in the conservation properties, the rate of change of the sum of certain variables (in the left hand side of a corresponding ODE) must be zero. Consequently the corresponding part of column sums also must be zero. Finally, the ODE can be reassembled in blocks with desired special structure of M-matrices.

Here in our case study, the state variables are listed in two subsets $\{x_3, x_4\}$ and $\{x_1, x_2, x_4, x_5\}$, and thus the non-linear ODEs (5) can be represented as a linear system with the system matrix of a special form, a negative M-matrix. Let these two subsets of state variables be assembled and merged together as follows

$$\tilde{x}(t) = \begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix}, \quad x^1(t) = \begin{pmatrix} x_3(t) \\ x_4(t) \end{pmatrix}, \quad x^2(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_4(t) \\ x_5(t) \end{pmatrix}.$$

Then the ODE system for modified state variable vector $\tilde{x}(t)$ is

$$\frac{\mathrm{d}\tilde{x}(t)}{\mathrm{d}t} = M\tilde{x}(t), \tag{9}$$

with the block diagonal system matrix M of a special form

$$M = \begin{pmatrix} -k_1 \cdot x_2 & k_{-1} + k_2 & 0 & 0 & 0 & 0 \\ k_1 \cdot x_2 & -(k_{-1} + k_2) & 0 & 0 & 0 & 0 \\ 0 & 0 & -k_0 & k_0 & 0 & 0 \\ 0 & 0 & k_0 & -k_0 - k_1 \cdot x_3 & k_{-1} & 0 \\ 0 & 0 & 0 & k_1 \cdot x_3 & -(k_{-1} + k_2) & 0 \\ 0 & 0 & 0 & 0 & k_2 & 0 \end{pmatrix}. \tag{10}$$

The initial conditions are

$$x(0) = \begin{pmatrix} e_0 & 0 & u(t_0) & 0 & 0 & 0 \end{pmatrix}^T.$$

3 Parameter estimation and experimental design

The quality of parameter estimation is usually measured by the squared error functional

$$J = \int_{t_0}^{t_f} \left(z_m(t) - z(p; u(t); t)\right)^2 \mathrm{d}t, \tag{11}$$

where $z(p; u(t); t) \in \mathbb{R}^{n_{out}}$ is the output vector, $z_m(t) \in \mathbb{R}^{n_{out}}$ are (continuous) measured data, $p \in \mathbb{R}^q$ is a parameter vector, e.g., $p = (k_0, k_1, k_{-1}, k_2)^T$, and $u(t)$ is the control input.

Here, in order to maximize the information content of the corresponding experiment, we formulate the optimal control problem, e.g., we look for an optimal impuls input $u(t_i)$

$$\max_{\text{admissible } u(t)} \|\mathcal{F}(p_0)\|. \tag{12}$$

If the quantity $\|\mathcal{F}(p_0)\|$, being evaluated at p_0, is the determinant of the Fisher information matrix, i.e., $\|\mathcal{F}(p_0)\| \equiv \det(\mathcal{F}(p_0))$, we speak about a D-criterion. Note that the key role in evaluation of $\mathcal{F}$ plays the sensitivity matrix $\chi = \frac{\partial z(p_0;u(t);t)}{\partial p} \in \mathbb{R}^{n_{out} \times q}$ because $\mathcal{F} = \chi^T \chi \in \mathbb{R}^{q \times q}$.

4 Conclusion

As a proof of concept, we took the case of enzyme-catalyzed reactions with a substrate transport chain, see [5] for parameter values. For two above introduced model formulations, i.e. the classical formulation (5) and the 'quasi-linear' Bohl-Marek formulation (9), and based on the different impuls controls $u(t_i)$ – the same dosis of substrate in different time instants t_i, one can calculate (numerically) parameter sensitivities, i.e. the partial derivatives of the output vector $z(p; u(t); t)$ with respect to individual model parameters. Afterwards, comparing $\|\mathcal{F}(p_0)\|$, the optimal control input maximizing the information content can be selected. Eventually, the computational costs related to both formulations (5) and (9) can be compared as well.

Acknowledgement: The work of Š. Papáček was supported by the Czech Science Foundation through the research grant No. 21-03689S. The work of Ctirad Matonoha was supported by the long-term strategic development financing of the Institute of Computer Science (RVO:67985807).

References

[1] E. Bohl, I. Marek: *Existence and Uniqueness Results for Nonlinear Cooperative Systems.* In: I. Gohberg, H. Langer (eds): Linear Operators and Matrices. Operator Theory: Advances and Applications, Vol. 130. Birkhäuser, Basel, 2002.

[2] E. Bohl, I. Marek: *Input-output systems in biology and chemistry and a class of mathematical models describing them.* Appl. Math. 50, 2005, pp. 219–245.

[3] J. Duintjer Tebbens, C. Matonoha, A. Matthios, Š. Papáček: *On parameter estimation in an* in vitro *compartmental model for drug-induced enzyme production in pharmacotherapy.* Applications of Mathematics, 64, 2019, pp. 253–277.

[4] I. Marek: *On a Class of Stochastic Models of Cell Biology: Periodicity and Controllability.* In: R. Bru, S. Romero-Vivó (eds.): Positive Systems. Lecture Notes in Control and Information Sciences, Vol. 389, Springer, Berlin, Heidelberg, 2009.

[5] C. Matonoha, Š. Papáček, V. Lynnyk: *On an optimal setting of constant delays for the D-QSSA model reduction method applied to a class of chemical reaction networks.* Applications of Mathematics, 67, 2022, pp. 831–857.

[6] A. Strouwen, B.M. Nicolaï, P. Goos: *Optimizing oxygen input profiles for efficient estimation of Michaelis-Menten respiration models.* Food and Bioprocess Technology 12, (5), 2019, pp. 769–780.