An algebraic approach to challenges on identification problems in systems biology

Mizuka Komatsu

Kobe University, 1-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan
m-komatsu@stu.kobe-u.ac.jp

Abstract

This talk is of twofold. A challenge appeared in identification problems, in particular unidentifiablility problems in parameter estimations, in the field of system biology is introduced with an example of PBPK models in the first section. In the second section, an algebraic and algebro-geometric approach to the challenge in a wider context is explained. More precisely, a method to extract a geometric structure that are uniquely determined by observed time series data and unidentifiable state-space models, followed by an underlying theory including, e.g., differential algebra, is introduced. An idea of model reduction technique that preserves the geometric structures of original models in a sense, is also discussed in the talk.
keywords: systems biology, parameter estimation, PBPK model, differential algebra, algebraic geometry, Gröbner basis, model reduction, generalized Lotka-Volterra equations

1 Introduction of challenges in identification problems in systems biology

In the field of systems biology, biological phenomena and their interactions are analysed integratedly. In such a field, biological systems are often modelled as and investigated through mathematical models. One of common modelling approaches is by ordinary differential equations, ODEs, which are constructed based on biological knowledge. Partial differential equations, PDEs, are also one of major modelling approaches in such a field. However, considering possible constructions of PDEs based on refining compartment models, a class of ODEs, analytical methods for ODE models are considered to be still useful to the ones for PDEs. Based on this, our study focuses on ODE models appeared in systems biology. In particular, we consider such models with unknown parameters to be estimated using observed time-series data, etc. Parameter estimation problems of such models, which is a type of identification problems of ODEs, are the main topic in this talk.

In order to illustrate challenges in parameter estimation problems in systems biology, an example based on physiologically-based pharmacokinetic models, so-called PBPK models is introduced. Roughly speaking, PBPK models describe drug responses in human bodies dynamically[1]. In the field of pharmacokinetics, such models are constructed and practically used for predicting the time course of concentration of drugs in plasma and other sites. Although PBPK models have their origin in such a field, their applications can be seen in other fields, e.g., [2], in which we newly applied the model into immunology. PBPK model is constructed by compartments each of which corresponds to an organ or a compound of organs. State variables of a PBPK model correspond to the amount of the interests, for example, drugs in compartments. Some of model parameters are assigned to the meaning of certain biological functions such as, e.g., clearance rates at organs. Thus, not only the behaviors of state variables but also parameter values are of importance, which tend to be unknown. Using relevant time-series data, e.g., drug concentrations in plasma measured at several time points from subjects, such parameters, i.e., biological functions, are quantitatively estimated for the corresponding subjects. In such estimations, there exist an intrinsic challenge; the parameters tend to be unidentifiable. In other words, parameters

cannot be uniquely determined from given data due to its insufficiency. Unidentifiablility of parameters causes difficulties in investigation of corresponding biological systems. For example, in the context of immunology, such situation makes detection of immunological abnormalities from estimated parameters difficult in spite of the desire. Besides, estimations conducted without considering the unidentifiable property may overlook possible important considerations on the systems as we pointed out in [4]. Due to experimental constraints, unidentifiability problems often appear in systems biology, suggesting that the importance of approaches to dealing with unidentifiable models.

2 Algebraic approaches to unidentifiable state-space models: the parameter variety

Based on the previous section, we consider an approach to unidentifiable models, especially described by ODE models with polynomial terms of the state and the input variables. In this study, we focus on the fact that the structures determined by observed data and unidentifiable models are uniquely determined as sets, which we call the parameter varieties, and then propose a method to describe them explicitly. In this method, a certain input–output relationship of the models are derived using the Gröbner basis for an ideal corresponding to the model at first. Then, the sets of parameters, each of which generates the given data, are described as sets of constraints in terms of parameters i.e. the algebraic varieties. Once the structure, i.e., the variety is extracted, overlooking of the feasible parameters, which may lead to insufficient or inappropriate system considerations, would never occur.

We applied the proposed approach in the analysis of viral dynamics, which reveals an important fact on the efficacy of the drug that was missed in a previous study [3]. Fig. 1 shows examples of estimated parameter varieties in parameter spaces of an unidentifiable model appeared in [3] given observed time-series data taken from two different subjects. See [4] for the details of the varieties. As can be seen in Fig. 1, our method captures feasible parameters exhaustively compared to conventional approaches [3], suggesting an applicability of our method to, for example, classifications of subjects in terms of their varieties.

Technically speaking, ODEs we consider may not be able to be dealt with commutative algebra naively since they contain derivatives of variables with respect to time. In fact, in order to derive input-output equations from such models, derivative operations, which are not allowed in commutative algebra, for the models are required. Thus, our method appears to fall under the umbrella of differential algebra, which is rather difficult compared to non-differential one. However, thanks to state-space representations of the models, it is guaranteed that input-output equations can be derived considering certain truncated differential ideals corresponding to the models, which can be regarded as non-differential ideals. Theoretical details of our method will be explained in the talk. In practice, the derivation of input-output equations relies on computation of the Gröbner basis, which implies its computational expensiveness. Considering this, in the talk, we discuss an idea of a model reduction method preserving the parameter variety, which is based on form-invariant transformations of generalized Lotka-Volterra equations.

Acknowledgments

We are grateful to Takaharu Yaguchi for collaboration throughout this work and Shinji Nakaoka for discussions on the generalized Lotka-Volterra equations. This work was supported by JST CREST JPMJCR1914 and JSPS KAKENHI Grant Numbers JP20J21185 and 20K11693.

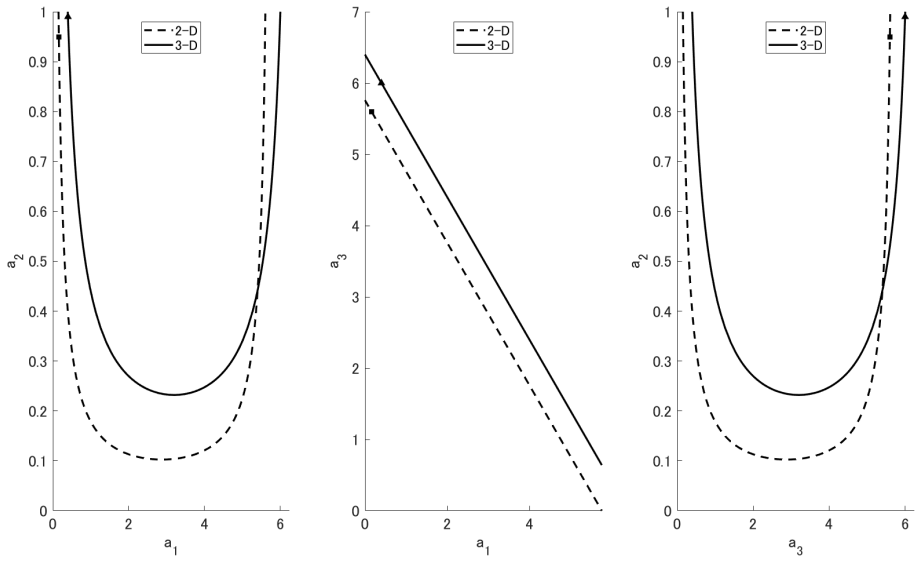

Figure 1: The parameter varieties of an unidentifiable model in two-dimensional parameter spaces that describes viral dynamics [3] given time-series data of viral load. $2-D$ and $3-D$ are labels that are assigned to subjects from which temporal measurements are taken. a_1, a_2, a_3 are parameters in the model. The symbols on the varieties correspond to estimated parameters in the previous study [3].

References

[1] S. A. Peters, "Phisiologically-based pharmacokinetic(PBPK) modeling and simulations: principles, methods, and applications in the pharamacoceutical industry," *Wiley* (2011).

[2] M. Komatsu and T. Yaguchi, "Mathematical modeling of kinetics of antigens and antibodies based on experiments toward analysis of the mechanism of allergies," *Transactions of the Japan Society for Industrial and Applied Mathematics*, Vol. 28, pp. 162–204 (2018).

[3] A. U. Neumann, N. P. Lam, H. Dahari, D. R. Gretch, T. E. Wiley, T. J. Layden and A. S. Perelson, "Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-α therapy," *Science*, Vol. 282, pp. 103–107 (1998).

[4] M. Komatsu and T. Yaguchi, "Method for estimating hidden structures determined by unidentifiable state-space models and time-series data based on the Gröbner basis," *arXiv*.