

Symbolic-Numeric Optimization for Biological Kinetics by Quantifier Elimination

Shigeo Orii¹
orii@strad.ssg.fujitsu.com

Hirokazu Anai²
anai@jp.fujitsu.com

Katsuhisa Horimoto³
khorimot@ims.u-tokyo.ac.jp

¹ Science Solutions Group, FUJITSU LIMITED, 9-3, Nakase 1-Chome, Mihama-ku, Chiba, 261-8588, Japan

² Information Technology Core Laboratories, FUJITSU LABORATORIES LTD., Kamikodanaka 4-1-1, Nakahara-ku, Kawasaki 211-8588, Japan

³ Laboratory of Biostatistics, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan

Keywords: symbolic-numeric optimization, quantifier elimination, fitting, biological kinetics, HIV proteinase

1 Introduction

We introduce a new approach to optimization for biological kinetics that deals with numerical data by symbolic quantifier elimination (QE). In this study, we illustrate the feasibility of the symbolic-numeric method in comparison with previous numerical methods.

2 Method and Results

The symbolic-numeric approach is applied to an optimization problem estimating five reaction parameters to fit a simulated signal with the five parameters to observed one, in the model described by ODE for the mechanism of irreversible inhibition of HIV proteinase[1].

2.1 Method of Symbolic-Numeric Optimization

The outline of our main tool, QE, in combination with numerical simulation is shown in Fig. 1.

QE deals with the first-order formulas, which consists of polynomial equations, inequalities, quantifiers (\exists, \forall) and Boolean operators ($\wedge, \vee, \neg, \Rightarrow$ etc). QE procedure is an algorithm to compute equivalent quantifier-free formula for a given first-order formula over the real closed field. For example, for the input $\forall x(x^2 + bx + c > 0)$, QE outputs the equivalent quantifier-free formula $b^2 - 4ac < 0$. This implies that we can obtain a condition for unquantified free variables that make the input formula true by QE. Moreover if all variables are quantified, QE decides whether it is true or false. For the details about QE, see [2].

In general, the constraints in parameters to be determined occurring in Fig. 1 have some uncertainty caused by inexact data. Hence if we apply QE directly to the constraints there is a real danger of arriving at an incorrect answer. Actually we often get “false” for feasible cases. So we introduce new variables EMAX into the constraints, which is a variable for obtaining information of feasible parameters even for such incorrect cases by absorbing uncertainty due to inexact data. Applying QE for the constraints including EMAX then we get possible range of EMAX so that they have feasible parameters. Then we obtain feasible regions of parameters by applying QE again the constraints where EMAX is specialized by a feasible value.

2.2 Procedure

1ST Stage: choose target variables to be optimized among all variables based on their algebraic relations, which is obtained by applying QE to the constraints derived from static model.

2nd Stage: estimate the five reaction parameters $\bar{k}''(k''_{22}, k''_3, k''_{42}, k''_{52}, k''_6)$. The symbolic-numeric approach of Fig.1 is applied to each the following:

- 1- find OFFSET (baseline) of the observed signal, i.e. the curve 3 on Table 5 of [1], with starting parameters \bar{k} ($k_{22}=300, k_3=10, k_{42}=500, k_{52}=0.1, k_6=0.1$) and a few points of the observed signal.
- 2- compute $\bar{k}'(i, j)$, i.e. candidates of \bar{k}'' , with \bar{k} and the above-mentioned OFFSET and two points of the observed signal.
- 3- compute $\bar{k}''(i, J)$ with $\bar{k}'(i, j)$, OFFSET and the two points of observed signal and find \bar{k}'' with minimum SSq(Sum of Squares of residual between the observed and simulated signal).

DO {for initial values ((EI=0.0035, 0.0055, 0.0001), (SI=26.5, 28, 0.5))

1st : Numerical Simulation

- *INPUT*: EI, SI, (\bar{k} or $\bar{k}'(i, j)$)
- *OUTPUT*: time course signal

2nd : Symbolic QE

- *INPUT*: EI, SI, the simulated and the observed signal
- *OUTPUT*: (offset or $\bar{k}'(i, j)$ or $\bar{k}''(i, j)$) and EMAX }

ENDDO

Figure 1: The symbolic- numeric approach.

2.3 Numerical Results

A result of the reaction parameter $k''(k_{22}'', k_3'', k_{42}'', k_{52}'', k_6'')$ with minimum SSq is shown on Table 1. The table 1 indicates that the SSq of the symbolic-numeric approach is same order of the one of the best effort of numerical approach [1]. Fig. 2 shows the observed signal and the simulated one with k'' of the symbolic-numeric approach on Table 1.

Table 1: SSq and reaction parameters of Symbolic-Numeric and numerical optimization approaches.

Method	SSq	k_{22}''	k_3''	k_{42}''	k_{52}''	k_6''
Symbolic-Numeric	0.0095767	211.6	9.719	1140	0.01	1e8
Simulated annealing [1]	0.0051282	201.1	7.352	1171	1.314e4	3e4

2.4 Merits

- a. The model parameters $\bar{k}'(i, j)$ and \bar{k}'' are estimated with a few points (e.g. two points) of the observed signal.
- b. Feasible ranges of $\bar{k}'(i, j)$ and $\bar{k}''(i, J)$ are selected because unfeasible region can be confirmed exactly by the result "false" obtained by QE.
- c. Our method enables us to estimate exactly how much the uncertainties of numerical simulation and observation should be so that the constraints become feasible.
- d. The symbolic-numeric approach provides feasible ranges of reaction parameters, e.g. k_{22}'' on Table 1 has range between 211.6 and 221.3.

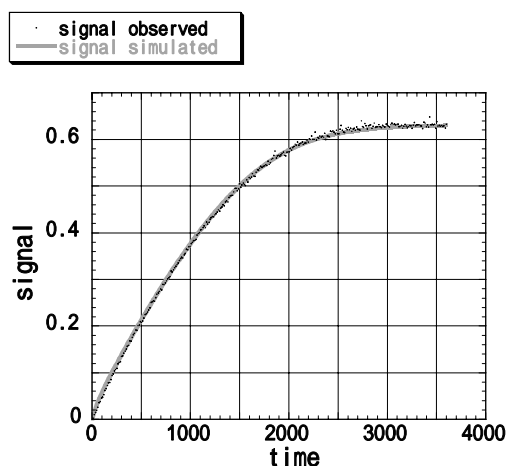


Figure 2: Observed and simulated signals.

3 Discussion

Our symbolic-numeric method of QE has distinct advantages in numerical analyses, and promises to give a new direction to the filed of bioinformatics. The details on the merits and falls of our approach will be discussed on the spot.

References

- [1] Mendes, P., and Kell, B.D., Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation, *Bioinformatics*, 14(10): 869-883, 1998
- [2] Caviness. B., and Johnson. J., editors., *Quantifier Elimination and Cylindrical Algebraic Decomposition, Texts and monographs in symbolic computation*, Springer-Verlag, 1998.