数理解析研究所講究録 1456 巻 2005 年 40-48

Symbolic-Numeric Optimization for Kinetic Models - An application to bioinformatics field -

折居 茂夫
SHIGEO ORII
富士通株式会社
FUJITSU LTD *

穴井 宏和
HIROKAZU ANAI
(株) 富士通研究所/(独) 科学技術振興機構
FUJITSU LABORATORIES LTD / CREST, JST †

堀本 勝久
KATSUHISA HORIMOTO
東京大学医科学研究所
INSTITUTE OF MEDICAL SCIENCE UNIVERSITY OF TOKYO ‡

Abstract

The sequencing of complete genomes allows analyses of interactions between various biological molecules on a genomic scale, which prompted us to simulate the global behaviors of biological phenomena on the molecular level. One of the basic mathematical problems in the simulation is the parameter optimization in the kinetic model for complex dynamics, and many optimization methods have been designed. We introduce a new approach to optimize the parameters in biological kinetic models by quantifier elimination (QE), in combination with numerical simulation methods. We also show an computational example for the inhibition kinetics of HIV proteinase in order to demonstrate the effectiveness of our method.

1 Introduction

Many methods for local and global optimization have been developed to model and simulate the global network of biological molecules in a cell [9, 11], and some simulators based on various optimization methods have also been designed (e.g. [10]). In the optimization methods, the estimation of kinetic parameters plays a key role in the development of kinetic models, which, in turn, promotes functional understanding at the system level, for example, in several biological pathways [7, 12]. In addition to the development of optimization methodology, the high performance of computers for numerical calculations also supports the optimization of the kinetic parameters in the complex dynamics within a reasonable amount of computational time. The high computer performance supports not only the numerical calculations based on calculus, but also the symbolic computations based on computer algebra (CA). Indeed symbolic computation is popular in software platforms such as Maple [6] and Mathematica [14], and the use of symbolic computation is increasing rapidly in biology [2]. The quantifier elimination (QE) is one

*orii@strad.ssg.fujitsu.com
†anai@jp.fujitsu.com
‡khorimot@ims.u-tokyo.ac.jp

of the main subjects in CA [3]. The QE was originally described in the mathematical proof by Tarski (1951) [13], which stated that the elementary theory for the reals is decidable. Although the original algorithm was too complex to solve the problem, a new method by Collins (1975) [5], cylindrical algebraic decomposition (CAD), allowed efficient implementation. Subsequent improvements in the algorithm provided feasible software platforms for symbolic computation (e.g., [1]). Although some applications of QE to biological issues by symbolic computation were reported [4], an amalgam of symbolic computation by QE and numerical calculation has not been designed. In this paper, we developed a novel optimization method in combination with symbolic computation and numerical simulation. A procedure for parameter optimization was designed to solve differential equations by QE in combination with numerical simulation. The performance of our procedure is illustrated by optimizing ten parameters between nine variables in a model for the inhibition kinetics of HIV proteinase [8]. As for the optimization performance, the goodness of fit to the observed data and the optimized parameters are compared with those from the previous studies [8, 9]. Furthermore, some characteristics of the symbolic-numeric method are discussed with the behaviors of the parameters and the variables in the model.

2 MATERIALS AND METHODS

2.1 Main tool: Quantifier elimination for uncertain input data

Our key tool for realizing symbolic-numeric optimization is quantifier elimination (QE) over the reals [3]. When we find feasible model parameters according to the observed data, we can solve some constraints derived by substituting observed data for the corresponding variables/parameters in the original constraints. Such constraints have some uncertainty, due to the inexact input data. Hence, if we apply QE directly to the constraints, then there is a real danger of arriving at an incorrect answer. Actually, we often obtain a "false" result for feasible cases. In order to extract the nontrivial information of feasible parameters, even for the incorrect cases, we propose the introduction of new variables into the constraints (see 2.2). We call them "error variables", which play a role in absorbing the uncertainty due to inaccurate input data. If we apply QE for the constraints including error variables, then we obtain possible ranges of error variables, so that the constraints are feasible. Then we obtain feasible regions of the model parameters by applying QE again to the constraints, where the error variables are substituted with the minimum value of their feasible ranges.

2.2 Mathematical Framework

Problem: In this paper, we consider the following fitting problem: the biological kinetic model analyzed here is of the form:

$$\dot{x}_i = v_i(X, K) \tag{1}$$

where $X = \{x_1, \cdots, x_{n_i}\}$ is a set of variables, and $K = \{k_1, \cdots, k_{n_j}\}$ is a set of parameters. The problem is to fit the parameters K of the model to the observed data $\tilde{X} = \{\tilde{x}_i^t\}$ for , $i = 1, \cdots, n_x$, $t = 0, 1, \cdots, n_{\tilde{x}_t}$ under the following additional conditions:

(i) Conservation laws: $h_i(X) = 0$

(ii)Variable ranges: $x_i \in D_i$, where $D_i = [a, b]$, $a, b \in \mathbb{R} \cup \{\infty\}$.

Basic Formula Here we set up the leading formula of this paper. As mentioned above, we have the following constraints Ψ with error variables ei from kinetic models: $\Psi \equiv \wedge_i \psi_i$, where $\psi_i = \dot{x}_i - v_i(X, K) + e_i = 0$. For the error variables we introduce a new variable, e_max, which means the magnitude of the error variables: $|e_i| \leq e_{max}$. Moreover, for the variables whose observed data is given, we consider the following objective conditions: $X_l^{(t)} - \tilde{X}_l^{(t)} = 0$, to achieve fitting. Then the "basic formula" is given as

$$F\dot{X}, X, K, e_{max}, e_i) \equiv (\Psi \wedge h_i(X) = 0 \wedge x_i \in D_i \wedge |e_i| \leq e_{max} \wedge X_l^{(t)} - \tilde{X}_l^{(t)} = 0). \tag{2}$$

We apply our symbolic-numeric approach to formulas derived by slightly modifying the basic formula according to various purposes.

2.3 Optimization Procedure

We explain the concrete procedure of symbolic-numeric optimization, which consists of six parts as illustrated in Figure 1.

Numerical simulation First we prepare simulation data for $\dot{x}_i$ and x_i, for which we lack observed data, by performing a numerical simulation of the kinetic models.

1. Set initial conditions $\tilde{X}^{(0)}$ and starting values for unknown parameters $\tilde{K}^{(0)}$ as follows: $\tilde{X}^{(0)} \equiv \{\tilde{x}^{(0)} | i = 1, \cdots, n_x\}$ and $\tilde{K}^{(0)} = K_1^{(0)} \cup K_2^{(0)}$, where $\tilde{K}_1^{(0)} \equiv \{k_1^{(0)}, \cdots, k_j^{(0)}\}$ are starting values, and $\tilde{K}_2^{(0)} \equiv \{k_{j+1}^{(0)}, \cdots, k_{n_j}^{(0)}\}$ are given fixed parameters.

2. By numerical simulation of the kinetic model (1), we obtain a time series for x_i and $\dot{x}_i$: $X_i^{(t)} = \{x_i^{(t)} | i = 1, \cdots, n_i, t = 0, 1, \cdots, n_t\}$ and $\tilde{\dot{X}}_i^{(t)} = \{\tilde{\dot{x}}_i^{(t)} | i = 1, \cdots, n_i, t = 0, 1, \cdots, n_t\}$.

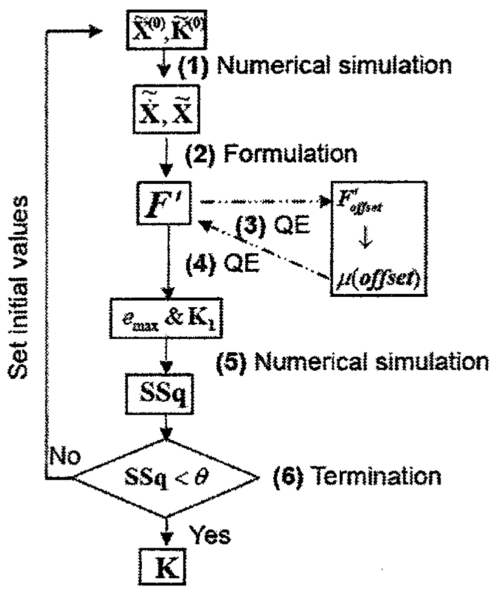

Figure 1: Flowchart of symbolic-numeric optimization. The variables and values are enclosed by the boxes, and the procedures are numbered corresponding to the description in the text.

Formulation After choosing some variables from X, we call them "focusing variables", Y, and substitute observed/simulated data into the remaining variables:

1. Choose a subset Y of $X : Y \subseteq X$.
2. Substitute $\dot{X}$, $X \setminus Y$, in F by the values of $\tilde{\dot{X}}, \tilde{X}$ at a time point t: $\dot{X}_i \leftarrow \tilde{\dot{X}}_i^{(t)}, \dot{X}_i \leftarrow \dot{X}_i^{(t)}$, where $x_i \in \tilde{\dot{X}}, X \setminus Y$. Then we denote the new formula as $F'(Y, K_1, e_{max}, e_i)$. We note by performing a QE computation for the formula, $\exists Y \exists K_1 \exists e_{max} \exists e_i (F')$,

Computation of offset by QE Observed data often contain an offset. Therefore, we must first determine the offset value. Here we consider the case that the offset appears linearly. For the sake of simplicity, we assume that only $\tilde{x}_1$ has an offset. Let F'_{offset} be the formula obtained by putting $\tilde{x}'_1 - offset$ into $\tilde{x}_1^{(t)}$ of F', where $offset$ is a variable for offset. By performing QE for $\exists X \exists K_1 \exists e_{max} \exists e_i (F'_{offset})$, we obtain the quantifier-free formula $\pi(offset)$, which stands for the feasible ranges of $offset$. Then we substitute the minimum value of the offset for the variable offset in F', and we denote it again by $F'(Y, K_1, e_{max}, e_i)$.

Estimation of emax and K_1 by QE First, we use QE to make the magnitude of ei as small as possible, and then we estimate the parameters K_1 by QE:

1. Compute the feasible range of $\pi(emax)$: by computing QE for $F'(Y, K_1, e_i)$, we obtain a quantifier-free formula $\pi(emax)$ describing the feasible ranges of emax. Next, we put the minimum value of e_{max} into e_{max} in F', and denote the resulting formula as $F''(Y, K_1, e_i)$.
2. Compute K_1: by computing QE for $Y \exists e_i (F'')$, we obtain a quantifier-free formula $\tau(K_1)$ describing the feasible ranges of K_1 Actually, the feasible ranges of K_1 are usually sufficiently narrow intervals (e.g., about 10^{-6}) to choose an appropriate specific value of K_1.

Computation of sum of squares (SSq) We estimate the goodness-of-fit for the obtained parameter valuesK_1 from the feasible ranges ofK_1 in terms of SSq.

1. Set initial conditions $\tilde{X}^{(0)}$ and K_1.
2. Perform numerical simulation of kinetic model (1).
3. Compute $SSq : SSq = \sum_t (x_1^{(t)} - \tilde{x}_1^{(t)})^2$.

Termination If SSq is smaller than a specific level θ, output K. Otherwise, set new initial values and go to (1).

2.4 Biological Model

We analyzed a model for the inhibition kinetics of HIV proteinase [8], as shown in Figure 2. The proteinase monomer (M) is inactive, but the enzyme (E) is active in the dimeric form. The dimer catalyzes the conversion of the substrate (S) to the product (P). The inhibitor (I) is competitive for the substrate and the product, and the inhibitor-binding enzyme is irreversibly deactivated (EJ). In the model, there are ten parameters and nine variables. According to the previous studies [8, 9], five parameters

$(k_{11}, k_{12}, k_{21}, k_{41}, k_{51})$ are given, and the remaining five unknown parameters $(k_{22}, k_3, k_{42}, k_{52}, k_6)$, two initial values (E_{init}, S_{init}) and the offset of the fluorimeter are estimated by the present method. The experimental data of the product $[P]$, which are composed of 300 data points measured from 0 to 3600 seconds, were downloaded from a web site (http://www.gepasi.org/tutorials/opt/hivfit.html).

3 RESULTS

First, we will describe the practical procedure for parameter optimization in the kinetic model for HIV proteinase, and then we will evaluate the optimized parameters by the goodness of fit to the observed data.

3.1 Procedure for Optimizing Parameters in HIV inhibition Model

$$
\begin{array}{llll}
M+M & \rightleftharpoons & E & k_{11}(\rightarrow), k_{12}(\leftarrow) \\
S+E & \rightleftharpoons & ES & k_{21}, k_{22} \\
ES & \rightarrow & E+P & k_3 \\
E+P & \rightleftharpoons & EP & k_{41}, k_{42} \\
E+I & \rightleftharpoons & EI & k_{51}, k_{52} \\
EI & \rightarrow & EJ & k_6
\end{array}
$$

Figure 2: Kinetic model for the inhibitor of HIV proteinase. The start values for ten parameters and the initial values for nine variables [9] are as follows: $k_{11} = 0.1, k12 = 10^{-4}, k_{21} = 100, k_{22} = 300, k_3 = 10, k_{41} = 100, k_{42} = 500, k_{51} = 100$, $k_{52} = 0.1, and k_6 = 0.1; \tilde{x}_1 = 0, \tilde{x}_2 = 0.004, \tilde{x}_3 = 25.0, \tilde{x}_4 = 0, \tilde{x}_5 = 0, \tilde{x}_6 = 0, \tilde{x}_7 = 0.003, \tilde{x}_8 = 0, and \tilde{x}_9 = 0$.

To perform the numerical simulation (in 2.3.1),K_1 and K_2, are defined as the five unknown parameters and the five given parameters, and the nine variables are allocated to $[P]$, $[E]$, $[S]$, $[ES]$, $[M]$, $[EP]$, $[I]$, $[EI]$, and $[EJ]$. Then we set the start value $\tilde{K}^{(0)}$ and the initial value $\tilde{X}^{(0)}$. The start values for ten parameters and the initial values for nine variables are cited from the previous study [9] (see the legend in Figure 2). Also, the two initial values, Einit and Sinit, are changed within a limited range with reference to the previous studies [8, 9]: 31 discrete values for ($[E]$ =0.00350, 0.00355, $\cdots$, 0.00500) and 9 values for ($[S]$ =24.0, 24.5, $\cdots$, 28.0). The focusing variables Y (in 2.3.2) are simply obtained by the symbolic computation with QE from the relationship between X andK_1 in the model. In the inequality $v_i(X,K)\Delta t + x_i^t \geq 0$, the elimination of Δt by QE outputs five inequalities including five parameters: $100*[E]*[I]-k52*[EI]-k6*[EI] > 0, 100*[E]*[I]-k52*[EI] > 0, 100*[E]*[P]-k42*[EP]-k3*[ES] < 0, 100*[E]*[P]-k42*[EP] > 0, and 100*[E]*[S]-k22*[ES]-k3*[ES] > 0$. Among the five unknown parameters in the above five inequalities, $[P]$ is included in the objective function, and $[S]$ is a large value relative to the other variables in the reaction molecules. Except for the last three inequalities including $[P]$ and $[S]$, only $[EI]$ appears in the terms related to the unknown parameters in the first two inequalities. Thus, the focusing variables Y are defined as $[P]$, $[S]$, and $[EI]$ in the present model. All symbolic computations by QE in this study are performed by REDUCE (ver. 3.8) (http://www.uni-koeln.de/REDUCE/). In addition, the conservation laws in the present model are obtained by Gepasi [10], a tool for estimating the kinetic flux in a given model, as follows: $h_1(X) = [S]+[ES]+[P]+[EP]-S_{init} = 0$ and $h_2(X) = [M]+2[E]-2[S]-2[P]+2[EI]+2[EJ]-(2E_{init}-2S_{init}) = 0$.

The computation of offset by QE (in 2.3.3) is realized by eliminating all of the variables by QE, except for $offset$ in F'_{offset}. Then we obtain $\mu(offset)$, depending on $\tilde{x}_1$, and the values of six variables and five parameters, as the following inequalities: $\mu(offset) = c_1 + c_2 offset > 0$ and $c_3 + c_4 offset \geq 0$, where c_1, c_2, c_3 and c_4 are constants. The constants in the above equations are estimated in each optimization.

Table 1: Goodness of fit with optimized parameters by symbolic-numeric method.

time	Iteration	k_{22}	k_3	k_{42}	k_{52}	k_6	SSq	E_{init}	S_{init}	*offset*
336	1	250.2	9.776	1306	0.103	0.0969	0.00962	0.00495	27.5	-0.028591
984	1	155.5	9.982	1127	0.102	0.0980	0.00825	0.00470	28.0	-0.040591
1848	2	134.5	9.817	1128	0.110	0.0900	0.00963	0.00460	28.0	-0.040591
		40.58	9.990	1211	0.102	0.0983	0.06380	0.00385	28.0	-0.040591
336, 984	1	225.9	9.970	1239	1759	1759	0.00856	0.00490	27.5	-0.028591
336, 1848	1	191.9	9.980	1304	4991	4991	0.00666	0.00475	27.5	-0.028591
984, 1848	2	211.6	9.719	1140	0.010	1.00E+08	0.00659	0.00495	27.0	-0.016591
		99.6	9.990	1254	1000	1000	0.03334	0.00430	27.5	-0.028591
Mendes & Kell	-	201.1	7.352	1171	1.31E+04	3.00E+04	0.00513	0.00547	26.79	-0.008962
Kuzmic	-	179.7	9.46	1117	0.0831	0.1224	-	0.00387	24.63	-0.01

For reference, the values related to the present optimization are also cited from previous studies [8, 9].

Using F' of Y and the offset obtained above, we can estimate emax andK_1 by QE (in 2.3.4). Note that 279 sets of emax andK_1 are obtained by the corresponding sets of E_{init} and S_{init}. Since the fitting of simulated data strongly depends on the initial values, we further simulate numerically E_{init} and S_{init} within the above ranges of E_{init} and S_{init}; by a standard technique of the bisection method, E_{init} and S_{init} for each set of emax andK_1 are estimated to minimize the SSq that is calculated for 300 values of $[P]$ (in 2.3.5). Finally, we obtain a set of e_{max},K_1, E_{init} and S_{init} by selecting a minimum SSq among the 279 SSq's.

To judge whether the loop in Figure 1 terminates or not (in 2.3.6), the minimum of SSq's is compared with the threshold θ. In the present study, the threshold is set to 0.01 to attain the same magnitude as that in the previous study [9]. If the SSq is smaller than θ, then we terminate the optimization process. If the SSq is larger than θ, then we start the loop by substituting 279 sets ofK_1 into $\tilde{K}^{(0)}$ with the same initial value sets of $\tilde{x}_2^{(0)}([E_{init}])$ and $\tilde{x}_3^{(0)}([S_{init}])$. Although the number of starting values $\tilde{K}^{(0)}$ increases as 279^n with the n-th iteration, the restriction of the parameter and the variable spaces prevents multiple iterations. Indeed, only one or two iterations were sufficient to attain the threshold in the present model.

3.2 Observed Data Fitting with the Optimized Parameters

The optimized parameters with the six sets of observed data are listed in Table 1, together with the iteration number, the goodness of fit measured by SSq, the initial values of E_{init} and S_{init}, and the offset. In addition, the fittings of simulated values to the observed data in six cases are described in Figure 3.

One of the remarkable features of the present fitting is that only one or two points of the observed data are sufficient to fit 300 data points with an SSq value of less than 0.01. The data point for the optimization is randomly chosen from 300 points of data, and all fittings attain the threshold by one or two iterations of the loop. In two of the six cases, two rounds of iterations were required, but the first fitting in each case agreed well with the observed data. This is partly because QE powerfully restricts the possible ranges of the parameters and the variables, and partly because the present model is simpler than that expected from the complex kinetics of ten parameters and nine variables. These points will be discussed in the following section.

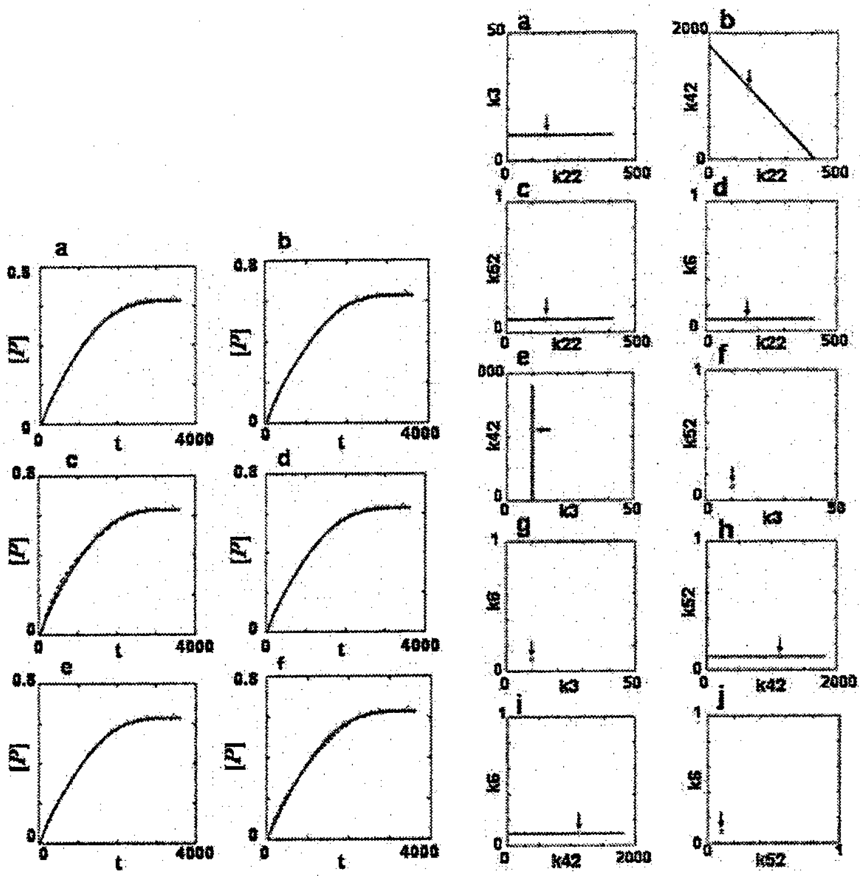

Figure 3: [left] Fitting to observed data with optimized parameters. The amount of product $[P]$ is multiplied by a coefficient (0.024), according to [9]. The experimental data are denoted by the jagged curve. The simulated curves are denoted by the solid curves (finally optimized) and if the loop iterates twice, by the broken curves (first optimized): **a**, t =336; **b**, t =984; **c**, t =1848; **d**, t =336 and 984; **e**, t =336 and 1848; **f**, t =984 and 1848. [right] Relationships between the five optimized parameters. The parameter ranges were estimated by the given values at t =984. The open circles indicated by arrows are the optimized parameter values at t =984.

Another feature is that the values of the parameters agree well with those in the previous studies [8, 9]. In particular, the highlighted parameters in this model, the inhibitor binding constant (k_{52}) and the deactivation rate constant (k_6), are about 0.10 and 0.097 in three of the six cases, which are similar values to the constants in one previous study [8]. In contrast, the constants in the three remaining cases are enormously large, except for k_{52} at t =984 and 1848, which are also similar to the magnitude of the rate obtained in the other previous study [9]. In comparison with both cases, the value in the latter case is unreasonably large for the analysis to be acceptable. Thus, the large dissociation and deactivation rate constants suggest that the potency of the inhibitor is overestimated in terms of the inhibitor reaction.

Two problems in the present optimization remain: one is the choice of the observed data for the optimization, and the other is the confidence intervals for the parameters. As for the data choice, the data showing a flat slope in the kinetic curve seem intuitively inadequate for the simulation. Indeed, by using the data of more than t =2500 in Figure 3, QE frequently outputs 'false'; this means no parameter and variable spaces for the initial conditions in F'. Any data, except for those in the steady states, may possibly output 'true' for the optimization by QE. As for the confidence intervals for the parameters, we will discuss them in terms of the parameters and the variable spaces in the following section.

4 DISCUSSION

4.1 Parameter Spaces Estimated by QE

The relationship between the parameters can be easily estimated by QE. Indeed, the parameter relationship in arbitrary ranges of variables is obtained by only the exclusion of the object function in F', and is useful to elucidate visually the parameter optimization by QE in terms of the confidence intervals of the optimized parameters.

Figure 4 shows the relationships between five optimized parameters. Given E_{init} and S_{init}, the F' obtained by excluding the object function consists of five known and five unknown parameters and three focusing and six remaining variables. In F', the three focusing variables and three of the five unknown parameters can be eliminated by QE, given numerical values at t =984 for six known variables, and then we can obtain the relationship between the remaining pair of parameters (a quantifier-free formula for the parameters).

One of the striking features of the elimination by QE is that the parameter spaces are highly restricted in all parameter pairs. The parameter space of each pair emerges in a very narrow range between two boundaries, which is a visual representation of the uncertainty of the parameters. For example, the narrow range for the parameter space between k_{22} and k_{42} is obtained with the following inequalities: $155.49657522 < k_{22} < 155.49657525$ and $1127.2733904 < k_{42} < 1127.2733905$. Furthermore, the total parameter ranges are shown visually in the restricted ranges; the maximum ranges of the parameter pairs in Fig. 4b are $0 \leq k_{22} \leq 411$ and $0 \leq k_{42} \leq 1813$, and the dissociation constant of $[EP]$ (k_{42}) varies in a wider range than that of $[ES]$ (k_{22}).

As seen in the figure, the relationships between the two parameters are divided into independent and dependent relationships. The parameter range with a slope indicates that the two parameters change mutually, and are dependent on each other, and the range with no slope indicates that the parameters are independent of each other. Only the relationship between k_{22} and k_{42} (b in Figure 4) is dependent, and the remaining relationships are independent. Thus, the present method reveals the mutual dependency of the parameters with their feasible ranges.

4.2 Concluding Remarks

The present study is the first application of QE to the parameter optimization problem in conjunction with a numerical simulation. Our symbolic-numeric method by QE shows the same magnitude of goodness of fit as the previous numerical optimization. Furthermore, the present method has the distinct potential to elucidate the relationships between the parameters in the kinetic model. Thus, our method provides a new direction for the analysis of kinetic models in the field of computational biology.

Acknowledgements

We would like to express our gratitude to Mr. Taku Takeshima for his kind assistance. One of the authors (K. H.) was partly supported by a Grant-in-Aid for Scientific Research on Priority Areas

"Genome Information Science" (grant 15014208) and for Scientific Research (B) (grant 15310134), from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] H. Anai, and H. Yanami, "SyNRAC: A maple-package for solving real algebraic constraints", In Sloot, P.M.A. et al. (eds.), Proceedings of the International Workshop on Computer Algebra Systems and Their Applications: CASA'2003, volume 2657 of LNCS. Springer-Verlag, 2003.

[2] M. Barnett, "Computer algebra in the life sciences", ACM SIGSAM Bulletin, Vol. 36, 2002, pp. 5-32.

[3] B. Caviness, and J. Johnson (ed.), Quantifier Elimination and Cylindrical Algebraic Decomposition, Springer-Verlag, 1998.

[4] C. Chauvin, M. Muller, and A. Weber, "An application of quantifier elimination to mathematical biology", In J. Fleischer, J. Grabmeier, F.W. Hehl, and W. Kuchlin (eds.), Computer Algebra in Science and Engineering, pp. 287-298, World Scientific, 1995.

[5] G.E. Collins, "Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition", In Lecture Notes in Computer Science, Vol. 33, pp. 134-183, Berlin: Springer-Verlag, 1975.

[6] F. Garvan, The Maple Book, London: Chapman and Hall, 2001.

[7] M.H. Hoefnagel, M.J.C. Starrenburg, D.E. Martens, J. Hugenholtz, M. Kleerebezem, I.I. Van Swam, R. Bongers, H.V. Westerhoff, and J. Snoep, "Metabolic engineering of lactic acid bacteria, the combined approach: kinetic modeling, metabolic control and experimental analysis", Microbiology, Vol. 148, 2002, pp. 1003-1013.

[8] P. Kuzmic, "Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase", Anal. Biochem., Vol. 237, 1996, pp. 260-273.

[9] P. Mendes, and D.B. Kell, "Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation", Bioinformatics, Vol. 14, 1998, pp. 869-883.

[10] P. Mendes, and D.B. Kell, "MEG (Model Extender for Gepasi): a program for the modelling of complex, heterogeneous, cellular systems", Bioinformatics, Vol. 17, 2001, pp. 288-289.

[11] C.G. Moles, P. Mendes, and J.R. Banga, "Parameter estimation in biochemical pathways: a comparison of global optimization methods", Genome Res., Vol. 13, 2003, pp. 2467-2474.

[12] I. Swameye, T.G. Muller, J. Timmer, O. Sandra, and U. Lingmuller, "Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling", Proc. Natl. Acad. Sci. USA., Vol. 100, 2003, pp. 1028-1033.

[13] A. Tarski, Decision Methods for Elementary Algebra and Geometry, Berkeley University of California Press, 1951.

[14] S. Wolfram, The Mathematica Book, 4th ed. Cambridge University Press, 1999.