Numeric-analytical method of the investigation of non-uniqueness of the rate constants determination of non-stationary chemical kinetics.

R.M. Asadullin[1], A.B. Ryzhkov[2]

[1]*Department of Mathematics, Bashkir State University, 450074 Ufa, Frunze street 32, Russia*
[2]*Institute of Organic Chemistry of Ufa Scientific Center of Russian Academy of Science, 450054 Ufa, Prospect Oktyabra 71, Russia*

ABSTRACT: This article is concerned with the description and application of a new method to investigation the non-uniqueness of the rate constants determination in chemical kinetics. The method is based on reparametrization of initial model by elimination of the non-measurable variables. Then parameters that are functions of rate constants of transformed system can be determined uniquely. The method leads to an algorithm, which is demonstrated in two examples. Numerical results with the method indicate that unique evaluates of the parameters of the reparametrized system can be obtained.

KEYWORDS: inverse kinetic problem, non-uniqueness, reparametrization.

INTRODUCTION

It is well known [1,2] that determination of rate constants in mathematical modeling of complex chemical reactions (inverse problem) often results in the non-uniqueness of solution. It arises from incomplete experimental data since not all of the reagent concentrations that take part in reaction can be measured in the experiment. Two types of the non-uniqueness of solution can be discerned - the first results in the finite aggregate of isolated solutions [3] (global nonuniqueness or nonidentification), the second gives infinite aggregate of solutions [4] (local nonuniqueness or nonidentification). Both types of non-uniqueness are thoroughly considered in the literature for the stationary and quasi-stationary regimes of reactions. Little attention has been paid to the non-stationary regime of reactions in the literature - the main results for that case were obtained using

group analysis and considered only local non-uniqueness. Global non-uniqueness together with the problem of the determination of all numerical parameters in the case of non-uniqueness was not investigated practically.

As a rule, determination of model parameters is reduced to the determination of the minimum of the functional $F(k)$, that characterizes the extent of deviation between experimental and calculated parameters values :

$$F(k) = \| z^{exp} - z^{calc} \|,$$

where the norm $\|\bullet\|$ is taken according to the type of experimental data: continuous or discrete, either normal distribution or other type of distribution etc. Here z^{calc} - vector of calculated values of concentrations or rates of reactions according to the model; z^{exp} - is the vector of experimentally determined values. However, application of such methods is correct if there exists only single solution of the problem in question (the global minimum of functional), but as a rule this condition is not known to be fulfilled *a priory* for the models of chemical kinetics. The present work considers numeric-analytical approach that allows to reveal the number of solutions of inverse problem as well as to find out all the values of parameters that can be determined from the experiment.

ANALYTICAL APPROACH FOR THE DETERMINATION OF THE NON-UNIQUENESS OF SOLUTIONS.

Let the model of chemical reaction be written in the following form according to the mass-action-law:

$$\dot{x} = f(x, y, k) \tag{1}$$

$$\dot{y} = g(x, y, k) \tag{2}$$

$$x(t_0) = x_0, \quad y(t_0) = y_0 \tag{3}$$

Here k denotes the vector of the rate constants of elementary stages of the reaction, $k=(k_1,...,k_s)$; $x_i(t)$ – are measurable on the range $[t_0,T]$ variables, $i=1,...,n$; $y_j(t)$ - non-measurable variables, $j=1,...,m$, $m \geq 1$; functions $f_i(x,y,k)$ and $g_j(x,y,k)$ are polynomials of no higher than third order in the variables x_k, y_j and linear in k_i.

Understanding under the inverse problem an estimation of rate constants of elementary stages using experimentally obtained solutions $x(t)$, we divide the solution into two steps:

Step 1. *A priory* analysis of inverse problem to reveal the number of solutions (the structure identification). Here is supposed that identification is carried out in the conditions of ideal experiment, *i.e.* the solutions $x(t)$ of the system (1)-(3) are given without experimental errors and the sample size of experimental data is large enough.

Step 2. Numerical determination of constants or independent parametric functions using approximately given solutions $\tilde{x}(t)$ on $[t_0,T]$.

On the first stage we will show that initial model (1)-(3) can be transformed to the form that allows to estimate definitely its parameters, the parameters being a functions of rate constants. We will do the following.

Expand the system (1),(2) by differentiating the system (1) to the infinite system, replacing at each differentiation the derivatives of x_i , y_i with their right parts according to (1),(2).

Finally, we will obtain the following infinite system of equations

$$\begin{aligned} &\dot{x} = f(x,y,k) \\ &\ddot{x} = f_1(x,y,k) \\ &... \\ &x^{(N+1)} = f_N(x,y,k) \\ &... \end{aligned} \qquad (4)$$

The next step is to exclude from the system (4) non-measurable variables y_i. In order to do it we will formulate (without proof) some statements.

Statement 1. Only the first m subsystems of equations system (where m – dimension of vector y) are necessary to exclude y_i variables from the system (4).

Suppose that unmeasurable variables y are excluded from system (4). We will get the equations system of the form

$$\begin{aligned} & x^{(m+1)} = \overline{F}_{m+1}(x^{(m)},...,x,k) \\ & x(t_0) = x_0, \\ & \dot{x}(t_0) = f(x_0,y_0,k) = x_{1,0} \\ & ..... \\ & x^{(m)}(t_0) = F_m(x_0,y_0,k) = x_{m,0} \end{aligned} \tag{5}$$

Statement 2. The system (5) is equivalent to the system (1)-(3) in the solutions $x(t)$.

Let term the system (5) as system of determining equations.

Corollary. The inverse problems for the systems (1)-(3) and (5) are equivalent.

The algorithm for the elimination of non-measurable variables can be built constructively, *i.e.* since the system (1)-(2) represents a system of polynomials, we may take advantage of the methods of computer algebra, where the technique of exclusion for similar systems is well developed [5]. It can be shown that in that case the system of determining equations is written in the following form:

$$\begin{aligned} & \psi^0_{l,0}(x^{(m_l+1)},...,x) = \sum_{i=1}^{\eta_i} \alpha_{l,i}(k)\psi_{l,i}(x^{(m_l)},...,x) \\ & x_l(t_0) = f_0(x_0,y_0,k), \\ & x_l^{(j)}(t_0) = \rho_l^j(x_0,y_0)\beta_l^j(k), \\ & i = \overline{1,n}, \quad j = \overline{1,m_l} \end{aligned} \tag{6}$$

where $m_l \le m$ and each equation (6) represents a polynomial in variables $x^{(m_l+1)},\ldots,x$; $\psi_{l,i}(x^{(m)}_l,\ldots,x)$, $i=1..\eta$ - linear independent monomials; $\alpha_{l,i}(k)$ - rational functions and $\beta_l^j(k)$ – polynomials in parameters. Let denote $\Psi_{l,1}(x^{(m)}_l,\ldots,x) = \Phi_{l,i}(t)$.

Statement 3. Let the initial data x_0, y_0 are such that for each j of (6) $\rho_l^j(x_0,y_0)\neq 0$. Moreover let for each equation of system (6) the points t_j be found on $[t_0,T]$, such that the vectors $\Phi_{l,i}(t_j)$, $j=1,..,\eta$ are linear independent. Then, there exists single solution of inverse problem of determination of the coefficients $\alpha_{l,i}$, β_l^j of the determining equations system.

From the corollary of statement 2 and statement 3 we obtain, that determination of the number of solutions of inverse problem for the model (1)-(3), is reduced to the analysis of the system of polynomial equations of the form:

$$\begin{aligned} \alpha_{l,i}(k) &= c_{l,i} \\ \beta_l^j(k) &= c_l^j, \end{aligned} \qquad (7)$$

where $c_{l,i}$, c_l^j - are numerical values of the coefficients of system (6). The system (7) can have single, finite or infinite number of solutions. Using the methods of polynomials algebra developed in the systems of analytical calculations such as Maple [6], the system (7) can be transformed to triangle or trapezoid form. In the first case the system (and inverse problem) has finite number of solutions and the number of possible solutions is easily estimated. In the second case, the inverse problem has infinite number of solutions and one obtains functionally independent combinations of parameters, that can be determined from experiment.

NUMERICAL METHOD.

The incorrectness of the numerical parameter estimation for the model class in question arises mainly when the inverse problem has infinite number of solutions. In the case of finite number of solutions, the problem is reduced to the determination of all minimums of functional. The

calculation features of this problem are discussed in [7]. The main difficulty in the determination of all minimums of functional is due to the great demand for calculation time. Thus, for the problem of big dimensionality determination of all solutions can not be carried out in practice since it is necessary to browse the region of the change of parameters together with simultaneous numerous solution of direct problem. Reparametrized models of the form (5) are considered rarely in the practice of numerical estimation. It is due to the fact that model contains the derivatives of high order (higher than the first) in variables that are measured in experiment.

At the same time, the quantitative theory of differential equations gives the method of numerical investigation of such equations.

Assume for simplicity that system (6) is solved relative to highest derivatives and initial data do not contain parameters. These assumptions do not narrow the generality of the problem since the solvability relative to highest derivative is present initially or can be obtained by dividing by the coefficient of the highest derivative. In the case of presence of parameters in the initial data, the problem is solved similarly. Write the system (6) in the form:

$$\begin{aligned} & x^{(m_l+1)} = \sum_{i=1}^{\eta_l} \alpha_{l,i}(k)\varphi_{l,i}(x^{(m_l)},...,x) \\ & x_l(t_0) = x_0 \\ & x_l^{(j)}(t_0) = x_0^j,\ l = \overline{1,n} \end{aligned} \tag{8}$$

and put the following problem: to estimate the parameters $\alpha_i(k)$ of the system (8) and to find the values of rate constants of chemical reactions using approximately given solutions $\tilde{x}_i(t)$ on $[t_0,T]$.

Suppose that system (8) has total order in derivatives equal to $(n+m)$. Write the system in the form:

$$x^{(r_l+1)} = \sum_{i=1}^{\eta_l} \alpha_{l,i}(k)\varphi_{l,i}(x^{(r_l)},...,\ x) \tag{9}$$

where $\sum_{l=1}^{n} r_l = m, l = \overline{1,n}$. Introduce the following abbreviations:

$$
\begin{array}{llll}
z_1 = x_1; & z_{r_1+2} = x_2; & \ldots & z_{n+m-r_n+1} = x_n \\
z_2 = \dot{x}_1; & z_{r_1+3} = \dot{x}_2; & \ldots & z_{n+m-r_n+2} = \dot{x}_n \\
\ldots & & & \\
z_{r_1+1} = x_1^{(r_1)}; & z_{r_1+r_2+2} = x_1^{(r_2)}; & \ldots & z_{n+m-r_n+1} = x_n^{(r_n)}
\end{array}
\tag{10}
$$

Then the system (8) is written as (11) using new abbreviations.

$$
\begin{array}{llll}
\dot{z}_1 = z_2; & & \ldots & \dot{z}_{n+m-r_n+1} = z_{n+m-r_n+2} \\
\dot{z}_2 = z_3; & & \ldots & \dot{z}_{n+m-r_n+2} = z_{n+m-r_n+3} \\
\ldots & & & \\
z_{r_1+1} = \sum_{i=1}^{\eta_1} \alpha_{r_1 i}(k)\varphi_{1_1 i}(z); & & \ldots & \dot{z}_{n+m} = \sum_{i=1}^{\eta_n} \alpha_{r_n i}(k)\varphi_{r_1 i}(z)
\end{array}
\tag{11}
$$

System (11) represents normal system of (n+m) differential equations with (n+m) variables z_i. By writing the initial data in the form:

$$
\begin{array}{l}
z_1(t_0) = z_{1,0} \equiv x_1(t_0) \\
z_2(t_0) = z_{2,0} \equiv \dot{x}_1(t_0) \\
\ldots \\
z_{r_1+1}(t_0) = z_{r_1+1,0} \equiv x_1^{(r_1)}(t_0) \\
\ldots \\
z_{n+m}(t_0) = z_{n+m,0} \equiv x_n^{(r_n)}(t_0)
\end{array}
\tag{12}
$$

one can obtains the Cauchy problem for the determination of variables $z_i(t)$, which is equivalent to the Cauchy problem (8), and the inverse problem of the determination of the parameters $\alpha_{i,j}$ has single solution.

For the numerical solution of this problem, we can use any classic method developed for similar systems. It should be noted that the question considering the stability of such approach is not considered in the present investigation.

The program written in the lab of Chemical Kinetics of IOC USC RAS was used in the present work for the determination of the parameters $\alpha_i(k)$ of the system (11),(12).

The program uses the variant of configuration method with random choice of the direction of parametric optimization with the solution of direct problem at each step by Novikov's (m,k)-method of 3rd order [8]. Without giving the thorough analysis of the advantages and disadvantages of such estimation method, it should be noted that configuration method in our case is more efficient then any other gradient method. It is mainly due to the fact that in the case of finite number of solutions of inverse problem the configuration method allows to find all of the solutions in the range of determination of parameters. It is understood that in the case of problem of larger dimension the application of configuration method becomes less effective. Moreover, the solution of direct problem itself is an effort taking procedure due to the rigidity of initial model. Nevertheless, for problems of low dimension the method gives rather illustrative results.

The calculation experiments were carried out for several mechanisms with different types of non-uniqueness. The results of such calculations are given below.

NUMERICAL EXAMPLES

Example (1).

As the first example, we consider the system of consecutive reactions of the type:

$$X_1 \xrightarrow{k_1} Y_1 \xrightarrow{k_2} Y_2 \xrightarrow{k_3} X_2 \qquad \text{(A)}$$

It is known [3], that two solutions of inverse problem exist for this mechanism. We will give the complete scheme for the determination of both solutions using the analytic approach described above for this example.

1) Firstly, one should build “experimental” data.

For initial model

$$\begin{aligned}
&\dot{x}_1 = -k_1 x_1 \\
&\dot{x}_2 = k_3 y_2 \\
&\dot{y}_1 = k_1 x_1 - k_2 y_1 \\
&\dot{y}_2 = k_2 y_1 - k_3 y_2 \\
&x_1(0) = 1,\ x_2(0) = y_1(0) = y_2(0) = 0
\end{aligned} \tag{13}$$

at given parameters k_1=2, k_2=3; k_3=6 , the direct problem was solved, i.e. the solutions $x_2(t)$ were found on the range [0;3] with the step h=0,2.

Using the experimental data $x_2(t)$ (“precise” experiment) the estimations of the parameters k_i, i=2,3 were made. Configuration method for the parameters k_2 and k_3 neatly gives two solutions: (k_2,k_3) = (3;6) and (6;3). Figure 1 shows the existence of two minimums.

In the next step the noise was added to experimental curve x_2=$x_2(t)$ with the normal distribution law with σ=0,1 to give the “experimental” data. An estimation of parameters k_2 and k_3 gave (3.03,5.75) and (5.75,3.03).

2) Transform the initial model to the following system of determining equations:

$$\begin{aligned}
&\dot{x}_1 = -k_1 x_1 \\
&\dddot{x}_2 + (k_3 + k_3)\ddot{x}_2 + k_2 k_3 \dot{x}_2 - k_1 k_2 k_3 x_1 = 0 \\
&x_1(0) = 1 \\
&x_2(0) = \dot{x}_2(0) = \ddot{x}_2(0) = 0
\end{aligned} \tag{14}$$

3) The coefficients of model (14) are the following functions on the rate constants: $\alpha_1 = k_1, \alpha_2 = k_2 + k_3, \alpha_3 = k_2 k_3, \alpha_4 = k_1 k_2 k_3$, the basis for which (the system of generators) are polynomials

$$\bar{\alpha}_1 = k_1, \bar{\alpha}_2 = k_2 + k_3, \bar{\alpha}_3 = k_2 k_3 \tag{15}$$

According to (15) the inverse problem has two solutions determined as roots of quadratic equation:

$$k_3^2 - \bar{\alpha}_2 k_3 - \bar{\alpha}_3 = 0 \tag{16}$$

4) Transform the system of determining equations (14) to the normal system of equations:

$$\begin{aligned} &\dot{z}_1 = -\alpha_1 z_1,\ \dot{z}_2 = z_3,\ \dot{z}_3 = z_4,\ \dot{z}_4 = -\alpha_2 z_3 - \alpha_3 z_2 + \alpha_1 \alpha_3 z_1 \\ &z_1(0) = 1,\ z_2(0) = 0,\ z_3(0) = 0,\ z_4(0) = 0 \end{aligned} \tag{17}$$

The coefficients of reparametrized model (17) are determined uniquely. We get single solution: $\bar{\alpha}_1 = 2$, $\bar{\alpha}_2 = 8.77$, $\bar{\alpha}_3 = 18.06$ by solving the inverse problem for model (17)

Figure 2 shows the lines of surface level $\Phi = \Phi(\alpha_2, \alpha_3)$, with one minimum.

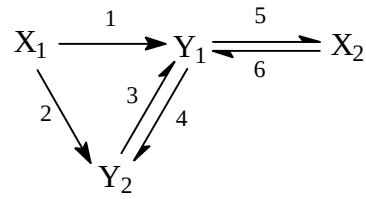
5) Two solutions of inverse problem are found from the quadratic equation (16): $k' \approx 3.033$ and $k'' \approx 5.74$. These values are practically identical to those found in “noise-added” experiment according to model (14).

Considered example allows to conclude that accuracy of the estimation of parameters from reparametrized models is not lower than that for direct methods. In general, the question considering the accuracy of method should be viewed specially.

Example (2).

In this example, we will consider the scheme of reaction that results in infinite number of solutions of inverse problem at given experiment.

Let there be given a linear scheme of reaction



(B)

The Cauchy problem for scheme (B) has the form:

$$\begin{aligned} \dot{x}_1 &= -(k_1 + k_2)x_1 \\ \dot{x}_2 &= k_5 y_1 - k_6 x_2 \\ \dot{y}_1 &= k_1 x_1 + k_3 y_2 - k_4 y_1 - k_5 y_1 \\ \dot{y}_2 &= k_2 x_1 + k_4 y_1 - k_3 y_2 \end{aligned} \tag{18}$$

Write the reparametrized model:

$$\begin{aligned} \dot{z}_1 &= -\alpha_5 z_1 \\ \dot{z}_2 &= z_3 \\ \dot{z}_3 &= -\alpha_1 - \alpha_2 z_2 + \alpha_3 z_1 + \alpha_4 \end{aligned} \tag{19}$$

where $z_1 \equiv x_1, z_2 \equiv x_2, z_3 \equiv \dot{x}_2$ and system parameters of (19) are given by conditions:

$$\begin{aligned} \alpha_1 &= k_3+k_4+k_5+k_6 \\ \alpha_2 &= k_4k_6+k_3k_6+k_3k_5 \\ \alpha_3 &= k_1k_5-k_3k_5 \\ \alpha_4 &= k_3k_5 \\ \alpha_5 &= k_1+k_2 \end{aligned} \tag{20}$$

The inverse problem has infinite number of solutions - the number of parametric functions is less than the number of unknown parameters. Determination of the parameters of initial model by any direct optimization method should result in the insensitivity of criteria towards the changes of model parameters. Criteria minimum in that case is represented not by a single or numerous sets of solutions but by a line obtained at intersection of the surfaces given by equations (19) in the space R^6 and any small value of mismatch criteria of experimental and calculated according to the model values is reached in the surroundings of this line.

Let illustrate it by numerical calculations. Firstly with a given constants K=(1;2;3;4;5;6) we will solve the direct problem, *i.e.* we will find the kinetic curves $x_1(t)$ and $x_2(t)$, and further we will

consider them as an experiment. Solve the inverse problem several times. The calculations are given in table 1.

Table. 1. Results of solving of inverse problem of the initial system of equation.

#	k_1	k_2	k_3	k_4	k_5	k_6
1	1.6111512	1.3888467	4.8262238	6.2693687	3.1022047	3.7781942
2	1.0308063	1.9691910	3.0896740	2.3885249	4.8496286	7.6584490
3	2.6984542	0.3015443	8.0994905	4.7964360	1.8531258	3.2588802
4	2.7391096	0.2608888	8.2166380	4.7072388	1.8253812	3.2494731
5	1.0518574	1.9481408	3.1502187	2.1222152	4.7516027	7.9492612

The coefficient values of the system of determining equations in the form (20) for these sets of constants gives the following results:

Table. 2. Coefficient values of the system of determining equations for sets of constants from table 1.

#	α_1	α_2	α_3	α_4	α_5
1	17.975991	56.893237	-9.9738133	14.971934	2.9999979
2	17.986277	56.938278	-9.9847437	14.983771	2.9999973
3	18.007933	57.035654	-10.008799	15.009375	2.9999985
4	17.998731	56.994287	-9.9985774	14.998497	2.9999984
5	17.973298	56.880542	-9.9705792	14.968588	2.9999982

As can be seen from table 2, the values of parameters α_i for different sets of k_6 are almost identical.

By solving the inverse problem for reparametrized system (19) using the same kinetic curves for x_1 and x_2, we obtain the following values for parameters α_i:

Table. 3. Results of solving of inverse problem of the reparametrized system of equations.

#	α_1	α_2	α_3	α_4	α_5
1	18.000137	57.000585	-10.000133	15.000153	2.9999988
2	17.999514	56.997793	-9.9994444	14.999419	2.9999988
3	17.999842	56.999265	-9.9998073	14.999806	2.9999988
4	18.000748	57.003324	-10.000808	15.000873	2.9999988
5	18.000761	57.003383	-10.000822	15.000889	2.9999988

The values of α_i given in tables 2 and 3, are not differing significantly from each other as should have been expected.

Thus, the calculations show that direct determination of model parameters in the case of infinite number of solutions leads to the arbitrary sets of rate constants that describe experiment equally good. The reparametrization results in the uniqueness of inverse problem. The recommendation in that case is to fix one of the reliable parameters on a certain value and to determine the other constants. It should be noted that additional check for global non-uniqueness is necessary in that case too. Thus by fixing k_5 at precise value we will obtain only two solutions of inverse problem K^*=(1;2;3;4;5;6) и K^{**}=(1;2;3;3;5;7).

Finally, it should be noted that *a priory* analysis of uniqueness could be easily carried out in the systems of analytical calculations such as Maple V, Reduce, Derive and other.

A.B. Ryzhkov would like to thank Prof. E.A. Novikov for a copy of his program NODE.

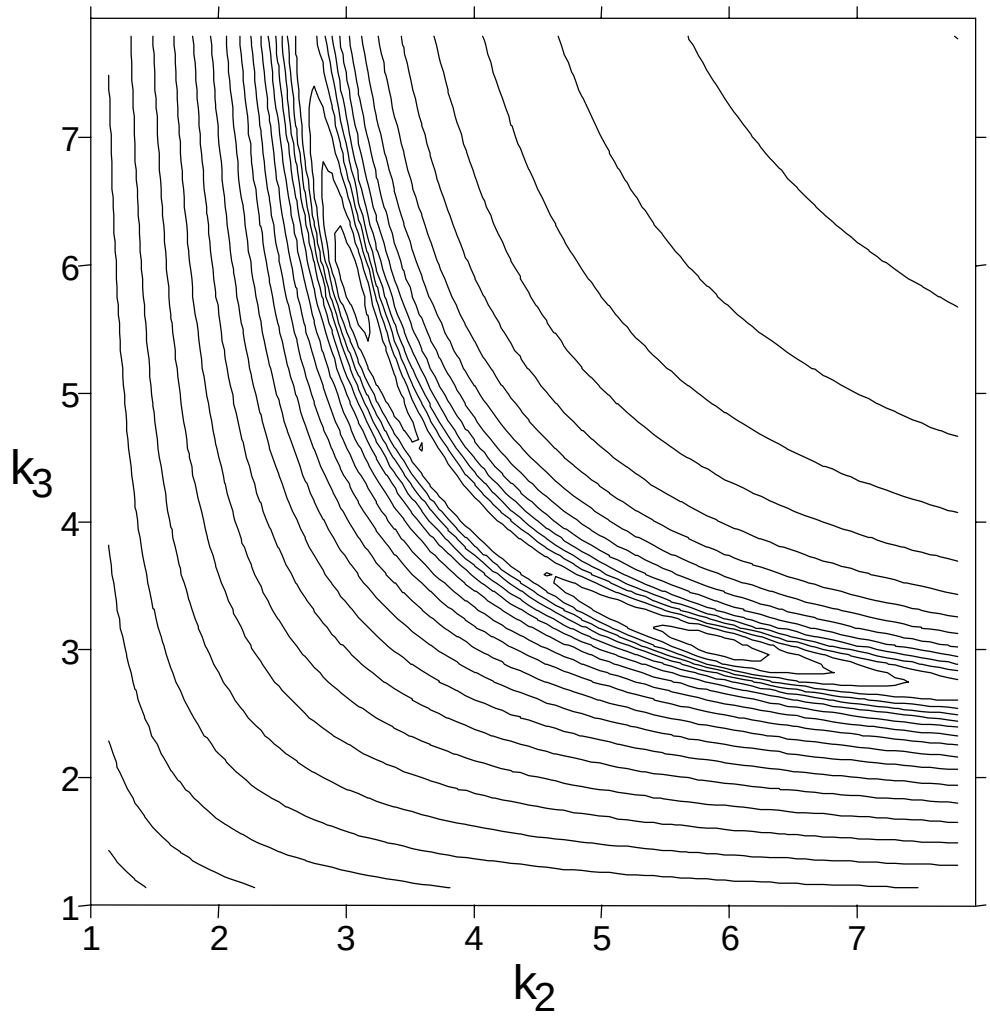


Figure 1 Contour plot of $\Phi = \Phi(k_2,k_3)$.

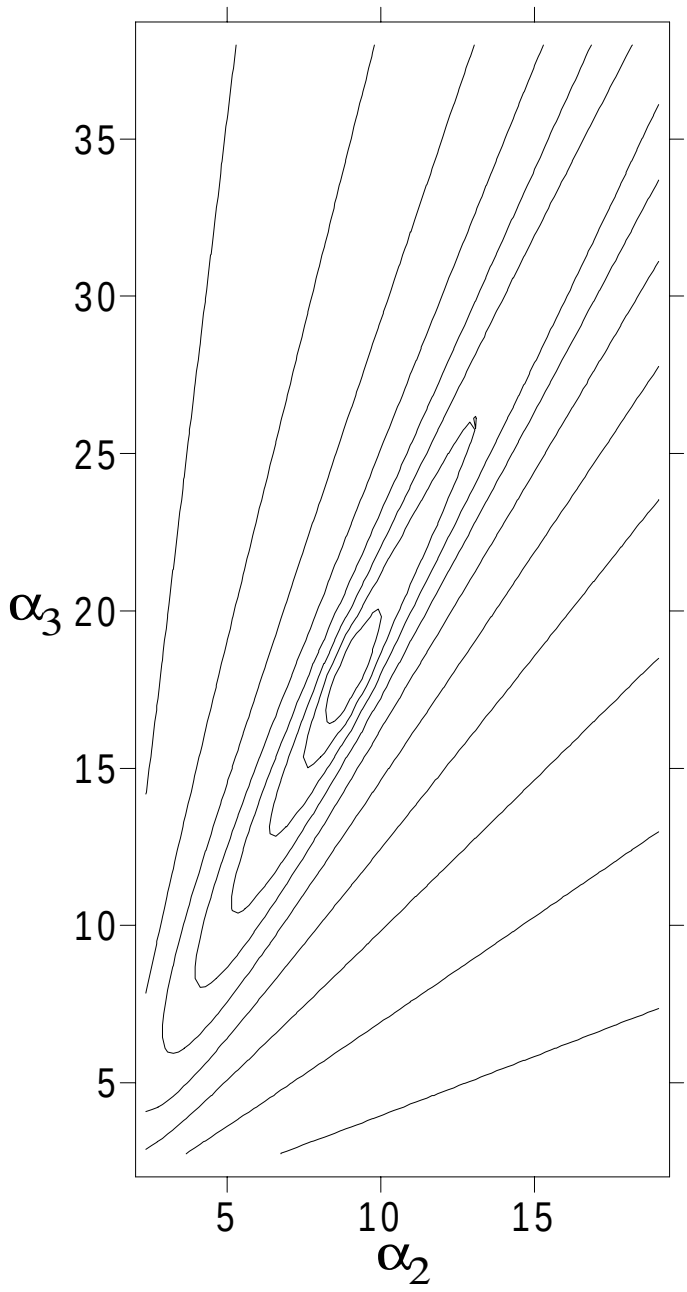


Figure 2 Contour plot of $\Phi = \Phi(\alpha_2, \alpha_3)$.

BIBLIOGRAPHY

1 V.G. Gorskii, S.I. Spivak. *Zhurnal Strukt. Khimii*, **29**, No. 6, (1988).
2 V.G. Gorskii, *Planning of Kinetic Experiments*, Nauka, Moskow, 1988
3 S.I. Spivak, Z.S. Ahmadshin, *React. Kinet. And Catal. Lett*, **10**, No. 3, (1979).
4 K. Glover, J.C. Willems, *IEEE Trans. Automat. Contr.*, **19**, No. 6, (1974)
5. B. Buchberger et al. (Ed.):*Computer Algebra. Symbolic and Algebraic Computation*, Springer-Verlag, New-York, 1982.
6. D. Redfern. *The Maple Handbook*, Springer-Verlag, New-York, 1993.
7 J.Milstein, in: *Modelling of Chemical Reaction Systems*, K.H. Ebert and P. Deuflhard, Ed. *The Inverse Problem: Estimation of Kinetic Parameters*, Springer-Verlag, New-York, 1981.
8 E.A. Novikov, M.I. Golushko, Yu.A Shitov, *Advances in Modeling & Analysis*, A, AMSE Press, **28**, No. 1 (1995).