

# Identification of Mathematical Model Parameters of Stationary Process

**Yuri L. Menshikov**

Department of Mechanics & Mathematics, Dnepropetrovsk University, Gagarina, Dnepropetrovsk, Ukraine

Email: [Menshikov2003@list.ru](mailto:Menshikov2003@list.ru)

Received January 2014

---

## Abstract

**In work the problem of synthesis of mathematical model of stationary process is examined in deterministic statement. It is supposed that the amount of measurements by each variable minimally and coincides with number of variable in model. The some possible variants of statement of such problem are considered. The calculations on real measurements were executed for comparison with known methods.**

## Keywords

**Parameters Identification, Different Statements, Regularization Method**

---

## 1. Introduction

The problem of parameters identification of mathematical models of physical processes in the form of algebraic relations between the initial and final states of the process arises when constructing mathematical models of these processes, in the investigation of optimal or nonstandard regimes of work, etc. [1,2]. If these parameters are considered as constant, it is implicit that the relationship between the initial and the final states of the process varies little during the time of measurement, of constructing a mathematical model and the syntheses of forecasts (and this change can be ignored). Such physical processes will be called stationary (in the sense they are being repeated).

Mathematical models of the same processes may have different structure and different degree compliance of reality. Linear mathematical models are the most popular, since any smooth function in a small neighborhood of parameters change is well approximated by a linear dependence. The number of variables of state, which are characterized the investigated process, may be different from one to infinity. However, any number of variables of the mathematical model can not be assumed accurate. Usually the number of variables in the model depends on the degree of their influence on the process under study and on the capabilities of the measuring equipment, which are measured by these variables.

In practice, the identification of parameters of mathematical models based on measurements of the characteristics selected a priori mathematical model, the overwhelming majority of currently represent methods that use statistical characteristics of the measured values [3-5]. These methods are successfully used in the study of the average characteristics of the physical process, as well as to build a long-term (average) forecast. However, if

the identification problems are considered in a stochastic statement, then the class of stationary physical processes should be expanded to include the class of possible physical processes in which the relation between the initial and final states of the process can be variable by averaging of the raw data. Forecast which obtained with help the averaged mathematical model will have only the averaged character. This quality of the forecast does not always satisfy the requirements of practice.

Furthermore, in these methods it is assumed that a mathematical model of communication between the averaged characteristics of the process is linear. This assumption is not founded.

In this paper the different approach to the problem of parameters identification of mathematical models of stationary processes is suggested. This method is suitable for the forecast using a small number of measurements (number of measurements of the characteristics of the process is assumed to be the number of these characteristics) [6]. In this approach, the measurement error has interval type and its value is known [7].

## 2. Statement of Problem

We shall present the problem of synthesis of linear mathematical model with  $n$  variables  $q_1, q_2, \dots, q_n$  relatively  $q_1$  for number of measurements  $n$ , as a problem of the solution of system [6]:

$$A_p(q_2, q_3, \dots, q_n)z = q_1, \tag{1}$$

where the operator  $A_p(q_2, q_3, \dots, q_n)z$  is determined as follows

$$A_p(q_2, q_3, \dots, q_n)z = z_1q_2 + z_2q_3 + \dots + z_{n-1}q_n + z_n e$$

$e$  is the unit vector of dimension  $n$ ,  $z$  is unknown vector of parameters of the mathematical model of the process.

As the measurements of variables  $q_1, q_2, \dots, q_n$  are received experimentally it is assumed that each measurement  $q_{ij}$ ,  $1 \leq i, j \leq n$  has some error the maximal size of which is known:

$$|q_{ij} - q_{ij}^{ex}| \leq \delta_i, \quad 1 \leq j \leq n, i = 1, 2, \dots, n \tag{2}$$

where  $q_{ij}^{ex}$  is exact measurements of variable  $q_{ij}$ .

The similar information of measurement errors, as a rule, is known a priori. The statistical characteristics of errors of measurements are unknown.

Let us denoted vector  $p$  as vector from space  $R^n \oplus R^n \oplus R^n \oplus \dots \oplus R^n = R^{n(n-1)}$ :

$$p^T = (q_{21}, \dots, q_{2n}, q_{31}, \dots, q_{3n}, \dots, q_{n1}, \dots, q_{nn})$$

where  $R^n$  is Euclidean vector space,  $(\cdot)^T$  is sign of transposition.

Each vector  $q_i$  can accept meanings in some closed area  $D_i \subset R^n$  by virtue of inequalities (2). Vectors  $p$  can accept meanings in some closed area  $D = D_2 \oplus D_3 \oplus D_4 \oplus \dots \oplus D_n \subset R^{n(n-1)}$ . The certain operator  $A_p$  associates with each vector  $p$  from area  $D$ . The class of operators  $\{A_p\} = K_A$  will correspond to the set  $D \subset R^{n(n-1)}$ .

Shall we rewrite (1) as

$$A_p z = u_{\delta_1}, \tag{3}$$

where  $u_{\delta_1} = q_1$ ;  $u_{\delta_1} \in U = R^n$ ;  $z \in Z = R^n$ ;  $\|u_{\delta_1} - u_1^{ex}\| \leq \delta_1$ ,  $u_1^{ex}$  is exact right part of (3);  $\|u_{\delta_1} - u_1^{ex}\| \leq \delta_1$ ;  $\|\cdot\|$  is the norm of a vector in Euclidean space  $R^n$ .

Let us consider now the set of the solutions of the Equation (3) with the fixed operator  $A_p \in K_A$ :

$$Q_{\delta_1, p} = \{z : \|A_p z - u_{\delta_1}\| \leq \delta_1\}.$$

The set  $Q_{\delta_1, p}$  is limited if  $\Delta = \det A_p \neq 0$  and unlimited if  $\Delta = \det A_p = 0$ .

Any vector  $z$  from set  $Q_{\delta_1, p}$  is the good mathematical model of process so this vector after action of the operator  $A_p$  coincides with the given vector  $q_1$  with accuracy of measurement  $\delta_1$ . For choice of particular model from set  $Q_{\delta_1, p}$  it is necessary to use additional conditions. If such conditions are absent then it is possible to accept as the solution (3) the element  $z_p = Q_{\delta_1, p}$  for which the equality is carried out [6]:

$$\|z_p\|^2 = \inf_{z \in Q_{\delta_1, p}} \|z\|^2. \tag{4}$$

The vector  $z_p = Q_{\delta_1, p}$  is possible to interpret as a maximum steady element to the change of the factors not taken into account (most stable part), as the influence of these factors will increase norm of a vector  $z_p$  [8]. Such a property of the solution  $z_p$  is especially important if one takes into account that the vector  $z_p$  further will be used for forecasting real processes (parameter  $q_1$ ).

Consider now the set  $Q^* = \bigcup_{p \in D} Q_{\delta_1, p}$ .

Let us consider an extreme problem

$$\|z^*\|^2 = \inf_{p \in D} \inf_{z \in Q_{\delta_1, p}} \|z\|^2. \quad (5)$$

The vector  $z^* \in Q^*$  is an estimation from below of possible solutions of the Equation (3).

The statement of the following extreme problem is possible also:

$$\|z_{\text{sup}}^*\|^2 = \sup_{p \in D} \inf_{z \in Q_{\delta_1, p}} \|z\|^2 \quad (6)$$

The vector  $z_{\text{sup}}^* \in Q^*$  has the greatest norm among the solutions of a problem of synthesis on sets  $Q_{\delta_1, p}$ .

Models  $z^*, z_{\text{sup}}^*$  can be used for short-term forecasting of change of variable  $q_1$  as on the one hand models  $z^*, z_{\text{sup}}^*$  are received by a rapid way and on the other hand these models are steadiest to the change of the factors not taken into account.

Except (5), (6) it is possible to examine the following statements of problems:

$$\|z_{0,0,\dots,1}\|^2 = \inf_{q_2 \in D_2} \inf_{q_3 \in D_3} \dots \inf_{q_{n-1} \in D_{n-1}} \sup_{q_n \in D_n} \inf_{z \in Q_{\delta_1, p}} \|z\|^2 \quad (7)$$

$$\|z_{0,0,\dots,1,1}\|^2 = \inf_{q_2 \in D_2} \inf_{q_3 \in D_3} \dots \sup_{q_{n-1} \in D_{n-1}} \sup_{q_n \in D_n} \inf_{z \in Q_{\delta_1, p}} \|z\|^2 \quad (8)$$

$$\|z_{0,1,\dots,1,1}\|^2 = \sup_{q_2 \in D_2} \dots \sup_{q_{n-1} \in D_{n-1}} \sup_{q_n \in D_n} \inf_{z \in Q_{\delta_1, p}} \|z\|^2 \quad (9)$$

In some cases it is expedient to consider the following problems of identification of parameters:

$$\|z^{0,0,\dots,0}\|^2 = \inf_{z \in Q_{\delta_1, p^{0,0,\dots,0}}} \|z\|^2, \quad (10)$$

$$\|z^{0,0,\dots,1}\|^2 = \inf_{z \in Q_{\delta_1, p^{0,0,\dots,1}}} \|z\|^2, \quad (11)$$

$$\|z^{1,1,\dots,1}\|^2 = \inf_{z \in Q_{\delta_1, p^{1,1,\dots,1}}} \|z\|^2, \quad (12)$$

where vector  $p^{0,0,\dots,0}$  has the minimal possible size of all components of vector  $p$ ,  $p^{0,0,\dots,1}$  has the minimal possible size of components  $q_1, q_2, \dots, q_{n-1}$  and has the maximal size of  $q_n$ ;  $p^{1,1,\dots,1}$  has the maximal possible size of all components of vector  $p$ .

It is possible to consider the following extreme problem

$$\|A_{p^{opt}} z_{\delta_1}^{pl} - u_{\delta_1}\|^2 = \inf_{z_a \in Q^*} \sup_{A_p \in K_A} \|A_p z_a - u_{\delta_1}\|^2 \quad (13)$$

where  $z_a$  is the solution of extreme problem

$$\|z_a\|^2 = \inf_{z \in Q_{\delta_1, a}} \|z\|^2. \quad (14)$$

Let's called solution  $z_{\delta_1}^{pl}$  as *more plausible mathematical model*.

Use of such model with the purpose of the forecast allows to receive the characteristic  $q_1$  with the least maximal deviation from experiment with possible variations of variables  $q_2, q_3, \dots, q_n$  within the given errors.

### 3. Methods of Solution and Test Calculations

For solution of above extreme problem (4) is used Tikhonov regularization method [9]. Extreme problem (4) is

replaced by the equivalent problem of minimizing the smoothing functional for more efficient use of numerical methods [9]:

$$M^\alpha[z, A_p, u_{\delta_1}] = \|A_p z - u_{\delta_1}\|^2 + \alpha \|z\|^2. \tag{15}$$

Euler equation for the functional (5) has the form:

$$A_p^* A_p z + \alpha z = A_p^* u_{\delta_1}, \tag{16}$$

where  $A_p^*$  is conjugate operator to  $A_p$ .

Regularization parameter  $\alpha$  was obtained by discrepancy method [9]:

$$\|A_p z_p - u_{\delta_1}\|^2 = \delta^2,$$

where  $z_p$  is the vector for which the minimum of the functional (16) on the set of possible solutions  $Q_{\delta_1, p}$  for a fixed operator  $A_p$ .

As initial data for the test calculations were selected economic characteristics of development of Ukraine in the period from 1999-2008 (**Table 1**).

In **Table 1**, the following notations:  $q_1$  is debt of state administration (billion UAH),  $q_2$  is Ukraine's GDP at constant prices (billion UAH),  $q_3$  is Ukraine's GDP in current prices (billion UAH),  $q_4$  is Inflation, average consumer prices (percent changes),  $q_5$  is the unemployment rate (percentage of total manpower),  $q_6$  is population (million),  $q_7$  is government revenues (billion UAH),  $q_8$  is government expenditure (billion UAH). The size of error in the initial data is taken equal to  $0.05 \|u_{\delta_1}\|$ .

The ultimate goal of parameters identification in this case is the construction of a mathematical model suitable for constructing a stable forecasting the selected stationary process and checking for its adequacy. For this purpose, the data are selected from the **Table 1** for 1999 to 2006 and is performed the solution of (6). Regularization parameter is obtained by discrepancy method:  $\alpha = 0.01$  [9].

As a result of calculations the mathematical model of the process was obtained:

$$q_1 = 0.08q_2 + 0.41q_3 + 0.32q_4 - 0.4q_5 + 0.5q_6 - 0.19q_7 - 0.71q_8 - 0.53 \tag{17}$$

To test the adequacy of the results were made retrospective calculations of the size of government debt for 2007 and 2008. In the first case, we used data from a column at 2007 **Table 1** (2nd-8th rows), and in the second case, we used the data from column in 2008 **Table 1** (2nd-8th rows).

Calculations showed the following results:

$$q_1(2007) = 90.05; \quad q_1(2008) = 153.1$$

**Table 1.** Economic indicators of Ukraine in the period of 1999-2008.

	Years									
$q_i$	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
$q_1$	79.6	77.0	74.6	75.7	78.5	85.4	78.2	80.6	88.7	189.4
$q_2$	406.0	429.7	469.1	493.5	540.7	606.1	622.5	668.0	720.7	735.9
$q_3$	130.4	170.1	204.2	225.8	267.3	345.1	441.4	544.1	720.7	948.1
$q_4$	22.78	28.3	12.0	0.74	5.21	9.04	13.6	9.06	12.8	25.2
$q_5$	11.1	11.5	10.8	9.6	9.1	8.6	7.2	6.8	6.3	6.4
$q_6$	49.1	48.7	48.2	47.8	47.4	47.1	46.7	46.5	46.2	46.0
$q_7$	41.6	56.8	68.4	81.3	101.5	128.1	184.6	235.2	301.6	419.7
$q_8$	34.8	62.4	74.6	85.4	103.9	143.3	194.7	242.7	315.9	449.3

According **Table 1**, for 1999 to 2006 the linear mathematical model of the same process by the method of least squares (MLS) was obtained [4]. Then similar calculation of government debt for 2007 and 2008 for Ukraine was made. In this case as in the first case, we used data from the column in 2007 **Table 1** (2nd-8th rows), and in the second case, we used data from the column in 2008 **Table 1** (2nd-8th rows).

The results of calculations by method of MLS showed the following results:

$$\tilde{q}_1(2007) = 84,68 ; \tilde{q}_1(2008) = 95,79 .$$

Comparison of calculations results showed that suggested method gives more adequate mathematical model of process.

Choice of the certain mathematical model is being determined of the specificity of a concrete problem and final goal of use of mathematical model. However the best mathematical model for the forecast can not be determined a priori [8].

#### 4. Conclusion

The offered approach to a problem of parameters identification of mathematical models of stationary processes allows expanding a class of the possible solutions (linear mathematical models) up to maximal possible. Some variants of statement of a problem of identification of parameters are considered. The calculations of test examples were obtained.

#### References

- [1] Grop, D. (1979) *Methods of Identification of Systems*. World, Moscow.
- [2] Seidg, A.P. and Melsa, Dj.L. (1974) *Identification of control systems*. Science, Moscow.
- [3] Malenvo, E. (1975) *Statistics Methods in Econometrics*. Statistics, Moscow.
- [4] Douguerty, K. (1999) *Introduce to Econometric*. INFRA-M, M., XIV.
- [5] Wit, E. and McClure, J. (2004) *Statistics for Microarrays: Design, Analysis, and Inference*. 5th Edition, John Wiley & Sons Ltd., Chichester. <http://dx.doi.org/10.1002/0470011084>
- [6] Menshikov, Yu.L. (2006) The Fast Identification of Parameters. *Proc. 11th Int. Conf. Mathematical Methods in Electromagnetic Theory, MMET'06*, Kharkiv, Ukraine, 167-169.
- [7] Polajk, B.T. and Nazin, S.A. (2006) Evaluation of Parameters in Linear Many-Dimensional Systems with Interval Uncertainty. *Problems of Control and Informatics*, Kiev. Ukraine, 1-2, 103-115.
- [8] Menshikov, Yu.L. and Nakonechnij, A.G. (2003) Principle of the Maximal Stability in Inverse Problems with a Minimum of the a Priori Information. *Proc. of Inter. Conf. Problems of Decision making under Uncertainties (PDMU-2003)*, Kiev-Alushta, 8-12 September 2003, 80-82.
- [9] Tikhonov, A.N. and Arsenin, V.Ya. (1979) *Methods of Solutions of Incorrect Problems*. Science, Moscow.