Adequate mathematical models in algebraic form

Menshikov Yu.L.,
Department of Mechanics and Mathematics, University of Dnipro, Ukraine
Gagarina av., 72, Dnipro,49010
Ukraine

Abstract—The paper studies the possibilities of mathematical models of cyclic processes in algebraic form for forecasting purposes. It is shown that such mathematical models can be used for reliable prediction only in a small local neighborhood of the initial data which are necessary for the synthesis of mathematical model. Outside of this area the forecast is not stable and additional conditions is needed. These conditions can be obtained from the physical laws of the studied processes or from additional requirements of specialists.

Keywords— cyclic processes, neural network, adequacy of a quantitative type, forecast.

I. INTRODUCTION

It is hard to develop a modern science without a wide use of mathematical models. An important place in this approach is occupied by neural network methods for simulation of physical processes. One of the types of neural networks are mathematical models in the form of algebraic relations. They are used in econometrics, in problems of search regularities, in problems of cryptography.

In this case, the question of the adequacy of the model and the process under study is important [1]. If a mathematical model does not meet the conditions of adequacy, then further use of such a model (neural network) is problematic. In the practice of synthesizing neural networks, expert assessments are usually used for these purposes. In this paper, the questions of the synthesis of adequate mathematical models (neural networks) in algebraic form are discussed, as well as the questions of the use of mathematical models of this type for the purpose of predicting the characteristics of physical processes.

II. MATHEMATICAL MODELS (NEURAL NETWORKS) IN ALGEBRAIC FORM

Consider a mathematical model of a physical process in the form of algebraic relations [1]

$$q_n = z_1 q_1 + z_2 q_2 + \cdots z_{n-1} q_{n-1} + z_n, \quad (1)$$

where $q_1, q_2, \dots q_{n-1}$ are characteristics of a physical process (inputs of a neural network), q_n is an output characteristic of a process (output of a neural network), $z_1, z_2, \dots z_n$ are the required constant coefficients of the mathematical model (synoptic weights of the neuron). Equation (1) is written in dimensionless form.

It can be shown that the mathematical model (1) corresponds to a single-layer neural network with (n-1) inputs and one output with constant parameters [2]. Mathematical models of this type describe well the physical processes of the cyclic type, for example, the baking of a dish, the smelting of metal, and so on.

. In this paper, the model (neural network) (1) is synthesized on experimental data. Such a problem in the literature is called the problem of parameter identification [3].

The problem of identifying the parameters of the stationary process is to build some mathematical model of the relationship between the selected characteristics of the process. The problem of identifying the parameters of the stationary process is to build some mathematical model of the relationship between the selected characteristics of the process $q_1, q_2, q_3 \dots \dots q_n$. For simplicity, we consider only the problem of constructing a linear model (1).

Let $z^T = (z_1, z_2, z_3, \dots z_n), (.)^T$ is the transposition sign. It is assumed that for each variable q_i $(i = 1,2, \dots n)$ we have m measurements $q_{ik} = (k = 1,2, \dots m), n \le m$. Denote by $q_i^T = (q_{i1}, q_{i2}, q_{i3}, \dots . q_{im})$. The construction of a linear multidimensional model in the classical formulation can be reduced to finding the solution of a redefined linear inhomogeneous system of algebraic equations by the least squares (MLS) method, for example. It is assumed that each measurement is given its statistical characteristics.

The paper proposes a deterministic approach to the problem of constructing multiple regression parameters, which does not use information about the statistical properties of measurements. The number of measurements is minimal and equal to the number of variables of the studied process.

We present the problem of constructing a linear model with respect to q_n for the number of measurements $m = n$, as the problem of solving the system

$$A_p(q_1, q_2, \dots q_{n-1})z = q_n, \quad (2)$$

where the operator $A_p(q_1, q_2, \dots q_{n-1})z$ is defined as follows

$$A_p(q_1, q_2, \dots q_{n-1})z = z_1q_1 + z_2q_2 + \cdots + z_{n-1}q_{n-1} + z_ne,$$

e is a vector with unit components of dimension n; $z, q_i(i = 1,2, \dots . n)$ is vector with E^n. For the norm of the vector in E^n we take the value

$$\|q_i\|^2 = (q_i, q_i) = q_{i1}^2 + q_{i2}^2 + \cdots + q_{in}^2,$$

It is easy to see that the operators A_p are linear. Since the measurements of variables are performed experimentally, we will think that each dimension q_{ik} , $1 \le i \le n, 1 \le k \le n$ has some error, the maximum value of which is known a priori:

$$|q_{ik} - q_{ik}^{ex}| \le \delta_i;\ 1 \le k \le n;\ i = 1,2, \dots n; \qquad (3)$$

where q_{ik}^{ex} are the exact measurement of the variable $q_i, 1 \le k \le n$.

Statistical characteristics of measurement errors are not used. Denote by p the vector with $E^n \otimes E^n \otimes E^n \otimes \dots \otimes E^n = E^{(n-1)n}$:

$$p = (q_{11}, q_{12}, \dots . q_{1n}, q_{21}, \dots . q_{2n}, \dots \dots q_{(n-1)1}, q_{(n-1)2}, \dots \dots q_{(n-1)n})^T .$$

Due to inequalities (3), the vector p can take values in some closed domain $D = D_n \otimes D_n \otimes D_n \otimes \dots \otimes D_n \subset E^{(n-1)n}$. The exact vector p^{ex} will correspond to the exact measurements $p^{ex} = (q_{11}^{ex}, q_{12}^{ex}, \dots q_{(n-1)n}^{ex})^T$ and the "exact" operator A^{ex} is defined as

$$A^{ex}z = z_1q_1^{ex} + z_2q_2^{ex} + \cdots + z_{n-1}q_{n-1}^{ex} + z_ne,$$

where $q_i^{ex} = (q_{i1}^{ex}, q_{i2}^{ex}, \dots, q_{(n-1)n}^{ex})^T, i = 1,2,3, \dots . , n.$. Rewrite (2) in the form

$$A_pz = u_{\delta_1} \qquad (4)$$

where $u_{q_1} = q_1;\ \|u_{\delta_1} - u_1^{ex}\| \le \delta_1;\ u_1^{ex} = q_n^{ex};\ u_{\delta_1}, u_1^{ex} \in R^n; z \in R^n$.

Each vector in the region corresponds to a certain operator A_p. The set $D \subset E^{(n-1)n}$ will correspond to the operator class $\{A_p\} = K_A$.

III. Synthesis of a Mathematical Model with the Adequacy of the Quantitative Type

We define an adequate mathematical model in form (2). Fix the operator A_p from the set K_A. Consider the set of possible solutions of equation (4) for the fixed operator $A_p \in K_A$:

$$\mathcal{Q}_{\delta_1,p} = \{z: \|A_pz - u_{\delta_1}\| \le \delta_1\}.$$

Definition. A mathematical model in the form (1) with a vector of synoptic weights of neuron z will be called adequate of quantitative type, if by substituting z in the left part of equation (4) we obtain the vector q_n, which differs from the exact vector q_n^{ex} by no more than δ_1 [1,4]. Comparison of the results of mathematical modelling with experimental data ensures the objectivity of the adequacy property of the mathematical model (neural network).

Comparison of the results of mathematical modelling with experimental data ensures the objectivity of the property of the adequacy of the mathematical model (neural network). Thus, any vector from the set $\mathcal{Q}_{\delta_1,p}$ gives an adequate mathematical model of a quantitative process.

If for a fixed $p \in D$ $\Delta = \det A_p \ne 0$, then the set $\mathcal{Q}_{\delta_1,p}$ is bounded. If for a fixed $p \in D$ $\Delta = \det A_p = 0$, then the set $\mathcal{Q}_{\delta_1,p}$ will be unbounded. If all errors are directed to zero $\delta_i \to 0$,, then $\Delta \to 0$. Thus, if the measurements coincide with the exact measurements or are close to each other, the determinant Δ will be equal to zero or close to zero. In addition, when performing numerical calculations with finite accuracy, some sets $\mathcal{Q}_{\delta_1,p}$ (with a small value of the determinant Δ) fall into the number of unbounded ones.

The sets $\mathcal{Q}_{\delta_1,p}$ are convex and closed. In addition, the sets $\mathcal{Q}_{\delta_1,p}$ will be unbounded for a special choice of the vector $p \in D$. So, there is an infinite set of sets containing elements with any large deviations from one another. In addition, when performing numerical calculations with finite accuracy, some sets (with a small value of the determinant) fall into the number of unlimited ones.

Thus, in the general case, the set of possible solutions to system (4) $\mathcal{Q}_{\delta_1,p}$ (the set of adequate mathematical models) is unlimited.

Let us consider the formulation of the parameter identification problem as a synthesis problem. Each vector from the set $\mathcal{Q}_{\delta_1,p}$ is an adequate mathematical model of the process. The problem of finding a vector $z \in \mathcal{Q}_{\delta_1,p}$ can be called the problem of synthesizing the parameters of a mathematical model. To select a specific model z_p from sets $\mathcal{Q}_{\delta_1,p}$ (which may be unlimited), additional conditions must be used [5].

However, there is no reason to believe that the model z_p will be close to the exact mathematical model. The mathematical model obtained by the regularization method [5] is a model with a minimum norm, which is maximally resistant to changes in unaccounted factors. This property of the solution z_p is especially important for further use for forecasting processes (indicator q_n).

An algorithm for the numerical solution of the problem of synthesis of an adequate model (adequate neural network) is described in detail into [1].

IV. Numerical Calculations

As a test example, the problem of constructing a linear algebraic mathematical model of the steel smelting process is considered [1]. This process is cyclical and can be described using a linear mathematical model in a small area of the selected point of the data change area. The initial data for test calculations were selected from the work [6] on the chemical composition, the parameters of heat treatment and properties of steel strength are presented in Table 1 and in Table 2 (see

attached page). Here were accepted the following designations C (q_1) is the amount of carbon [kg]; Si (q_2) is amount of silicon [kg]; Mn (q_3) is amount of manganese [kg]; P (q_4) - amount of phosphorus [kg]; S (q_5) is the amount of sulfur [kg]; Cr (q_6) is amount of chromium [kg]; Ni (q_7) is amount of nickel [kg]; Al (q_8) is the amount of aluminium; Cu (q_9) is the amount of copper [kg]; Ti (q_{10}) is the amount of titanium [kg]; V(q_{11}) is the amount of vanadium [kg]; T_1 (q_{12}) is quenching temperature [^{o}C]; T_2 (q_{13}) is the quenching time [^{o}C]; w (q_{14}) is water consumption [m^3/hour]; T_{opt} (q_{15}) is the optimum temperature [^{o}C]; τ_{opt} (q_{16}) is the optimum time [^{o}C]; σ_b (q_{17}) is the strength of steel [MPa].

As a result of the calculations, an adequate mathematical model of the steel smelting process was obtained in algebraic form.

z = (15.2158, 73.1382, 43.3567, 0.3820, 0.2196, -0.8670, -0.2950, 1.0681, 4.5157, 0.0930, 1.2498, 0.0151, -1.2778, 10.2737, 1.2555, -1.3005, 16.9142)^T.

As follows from the calculation results, the parameters of an adequate mathematical model

have no reasonable physical meaning. For example, the composition of the alloy cannot contain components with a negative sign. However, this model gives the most accurate simulation results. But one expert, when building a neural network, is not able to offer such coefficients. The loss in physical interpretation is compensated for by the high accuracy of the simulation.

V. THE PROBLEM OF PREDICTING THE PROPERTIES OF STEEL

Adequate mathematical models (neural networks), first of all, are aimed at predicting the behavior of physical processes. In the general case, adequacy is understood as the fulfillment of two conditions: the adequacy of the quantitative type and the adequacy of the qualitative type. For mathematical models in algebraic form, the second criterion cannot be met.

Let us determine the limits of change in the initial data for $q_1, q_2, \ldots, q_{16}$ according to Table 1:

$q_1 \in [0.57; 0.61], q_2 \in [0.30; 1.43], q_3 \in [0.78; 1.46], q_4 \in [0.006; 0.012], q_5 \in [0.003; 0.009], q_6 \in [0.07; 0.09], q_7 \in [0.05; 0.06], q_8 \in [0.023; 0.035], q_9 \in [0.04; 0.09], q_{10} \in [0.006; 0.009], q_{11} \in [0.082; 0.088], q_{12} \in [860; 900], q_{13} \in [120; 220], q_{14} \in [70; 80], q_{15} \in [500; 520], q_{16} \in [130; 150].$

The vector product of these segments forms the neighborhood $Q \in E^{16}$, which we will call the *basic* neighborhood.

This paper explores the possibilities of forecasting without using a qualitative criterion of adequacy.

For this purpose, retrospective calculations of steel strength were performed for the cases of steel smelting specified in the original table. As an example of such a check, the calculations of steel strength for the all experiments are given (lines 1-17 in Table 1). Calculations have shown good agreement between the calculated value q_{17} and the tabulated value (see Table 1). The maximum deviation of the calculated value of steel strength and the experimental value does not exceed 7%.

Numerical calculations of the prediction of steel properties have shown that if the initial parameters (chemical composition) are within the basic neighborhood of initial data Q, for which an adequate mathematical model of steel smelting was built, then the prediction of steel strength practically coincides with the tabular values.

Other results were obtained for experiments in the initial data [6], which are not included in matrix A. For example, a retrospective calculation of the strength of steel for the 25th line in Table 1 [6] showed that the calculated value of steel strength is 1170.0, and the experimental the magnitude is 960.4. The forecast error is 22%. Similar calculations for the 26th row are as follows: the calculated value of the steel strength is 1151.4, and the experimental value is 970.2. The margin of error is 19%. The calculations for the 38th row are as follows: the calculated value of the steel strength is 1189.6, and the experimental value is 1009.4, the error is 18%. The calculations for row 32 are as follows: the calculated value of the steel strength is 1395.1, and the experimental value is 970.2. The margin of error is 44%.

For all these cases, the initial data $q_1, q_2, \ldots, q_{16}$ went beyond the base neighborhood Q. In this case, a large forecast error corresponds to a larger deviation of the initial forecast data from the base neighborhood.

In addition, the analysis showed that the input data from the basic neighborhood forms an area with a sufficiently deformed shape, for example, in comparison with a multidimensional ellipsoid. Let's depict in Figure 1. - Figure 8. three-dimensional intersections of the basic neighborhood of the input data (Table 1) in order to analyze the forecast errors.

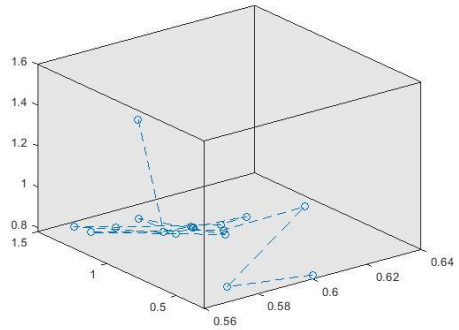

Fig. 1. Data dependence in space $q_1 q_2 q_3$.

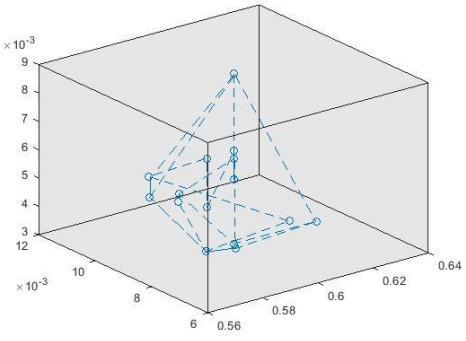

Fig. 2. Data dependence in space $q_1q_4q_5$.

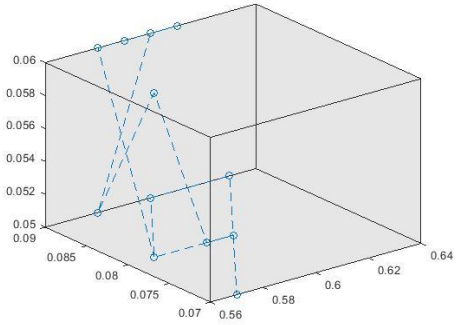

Fig. 3. Data dependence in space $q_1q_6q_7$.

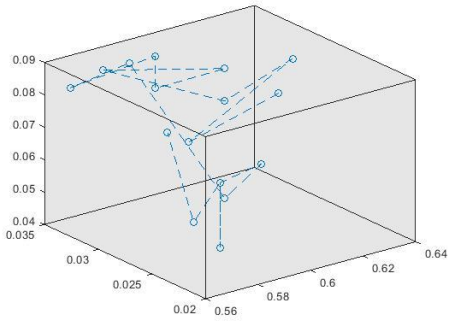

Fig. 4. Data dependence in space $q_1q_8q_9$.

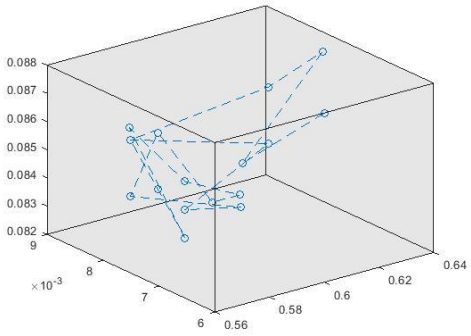

Fig. 5. Data dependence in space $q_1q_{10}q_{11}$.

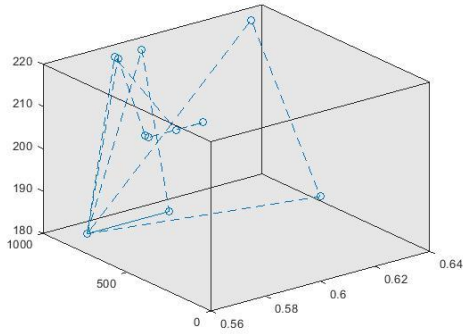

Fig. 6. Data dependence in space $q_1q_{12}q_{13}$.

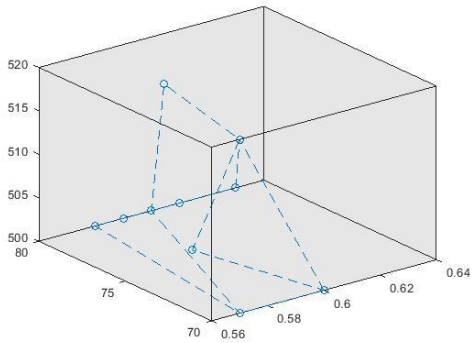

Fig. 7. Data dependence in space $q_1q_{14}q_{15}$.

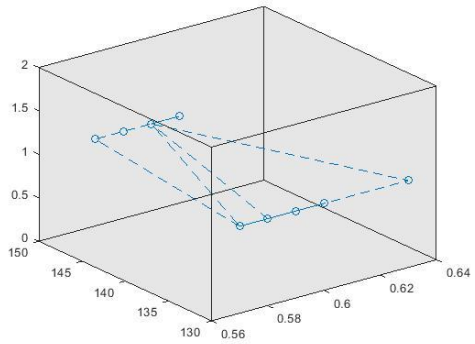

Fig. 8. Data dependence in space $q_1q_{16}1$.

Unstable forecast results were obtained if the initial parameters go beyond the boundary of the basic neighborhood.

Obviously, in order to obtain good forecast results, it is necessary to have a more correct basic neighborhood. In other words, for forecasting purposes it is necessary to have specially organized experimental measurements.

To replace the property of qualitative type adequacy, it is possible to use some other characteristics of the physical process for forecasting. For example, the proportions of the characteristics of a mixture of alloy components must obey a certain ratio. For example, a soup cannot contain salt in the size of cabbage or potatoes. So, to replace the property of of qualitative type adequacy, it is necessary to involve the ratio and dependences from the general physical laws of nature. It is also possible to form special conditions for forecasting purposes.

The presence of an adequate mathematical model makes it possible also to make some forecast on the value of the coefficients of this model. For example, a large positive coefficient for a physical variable will indicate a large increase in the strength of the steel with an increase in this physical parameter. Confirmation of this property can be observed inside the area of Q.

VI. Conclusion

For the cyclic processes under consideration, an algorithm for constructing an adequate mathematical model in algebraic form is proposed. Such models describe well the physical process only in a small neighborhood of the initial data. Several possible directions for improving the quality of forecasting using adequate models in algebraic form are proposed.

The results of the work can be used also when constructing adequate neural networks of almost any physical processes with a continuous change in properties. In this case, the most accurate network can be obtained, however, with the loss of physical meaning for the network coefficients.

References

[1] Yu. L Menshikov, Synthesis of adequate mathematical descriptions of physical processes, Monograph, Cambridge Scholars Publishing, 2020.

[2] R. Hecht-Nielsen, Neurocomputing: Picking the Human Brain , IEEE Spectrum March, 1988, p 36.

[3] L. Ljung, System Identification - Theory for the User N.J.: PTR Prentice Hall, 1999.

[4] Yu. L. Menshikov, Criteria for estimation the adequacy of mathematical descriptions of physical processes Modeling and Simulation in Engineering, Intech-Open London, Chapter 6, 2019.

[5] A.N. Tikhonov and V.Ya. Arsenin, Solving Methods of ill-posed problems, Moscow Science, 1979.

[6] D.N. Togobitskaya, A.I. Babachenko and A.S. Kozachek, “Physicochemical criteria and models for assessing the effect of chemical composition on the properties of wheel steel” , Scientific news. Modern problems of metallurgy, Dnipropetrovsk: N Metau, n. 16, 2014.

Attached page

Table 1. Chemical components for steelmaking.

N / N	C (q_1)	Si (q_2)	Mn (q_3)	P (q_4)	S (q_5)	Cr (q_6)	Mi (q_7)	Al (q_8)	Cu (q_9)	Ti (q_{10})	V (q_{11})	T_1 (q_{12})	T_2 (q_{13})	w (q_{14})	T_{opt} (q_{15})	τ_{opt} (q_{16})	σ_b (q_{17})
1.	0.60	0.3	0.8	0.009	0.009	0.08	0.05	0.023	0.09	0.006	0.088	900	180	70	500	130	1048.4
2.	0.5	0.33	0.84	0.006	0.005	0.07	0.05	0.024	0.08	0.006	0.087	900	180	70	520	130	1039
3.	0.63	0.94	0.824	0.009	0.003	0.09	0.05	0.029	0.085	0.007	0.088	900	220	80	500	130	1159
4.	0.60	0.94	0.82	0.009	0.003	0.09	0.05	0.028	0.08	0.007	0.088	0.880	200	80	500	150	1136
5.	0.57	1.31	0.80	0.008	0.006	0.08	0.05	0.032	0.09	0.008	0.086	900	180	70	500	130	1137
6.	0.60	0.93	0.80	0.009	0.006	0.08	0.05	0.028	0.09	0.007	0.086	900	180	70	500	130	1127
7.	0.59	1.36	0.78	0.008	0.006	0.08	0.05	0.032	0.08	0.008	0.083	900	220	76	500	130	1166
8.	0.59	1.09	0.79	0.008	0.007	0.08	0.05	0.032	0.09	0.007	0.084	900	220	76	500	130	1127
9.	0.57	1.43	0.79	0.007	0.006	0.08	0.06	0.035	0.08	0.008	0.084	900	180	70	520	130	1156
10.	0.58	1.33	0.78	0.008	0.007	0.09	0.05	0.032	0.09	0.008	0.086	900	220	76	520	130	1078
11.	0.60	0.96	0.84	0.012	0.004	0.09	0.06	0.028	0.05	0.008	0.083	860	200	80	500	150	1137
12.	0.61	0.97	0.84	0.008	0.004	0.09	0.06	0.027	0.06	0.008	0.083	860	200	80	500	150	1137
13.	0.59	0.97	0.86	0.009	0.003	0.09	0.06	0.026	0.06	0.008	0.084	860	200	80	500	150	1123
14.	0.59	0.96	0.86	0.011	0.004	0.09	0.06	0.026	0.04	0.009	0.085	860	200	80	500	150	1139
15.	0.59	0.98	0.86	0.009	0.003	0.09	0.06	0.026	0.06	0.008	0.082	880	200	80	500	150	1147
16.	0.58	0.98	0.87	0.009	0.005	0.09	0.06	0.026	0.05	0.008	0.084	880	220	80	500	150	1137
17.	0.57	0.97	1.46	0.008	0.006	0.08	0.05	0.026	0.08	0.008	0.086	900	180	70	500	130	1137