

# Multistage GMDH Algorithm for Classification Purposes of Biological Objects Defined by a Set of Measurements

Volodymyr Pavlov, Olexandra Konoval

National technical university of Ukraine «Kyiv polytechnic institute»,  
37 Peremogy Ave, Kyiv, Ukraine, 03056

vpavlo@bk.ru, alexandra.konoval@gmail.com

**Abstract.** *To solve the classification problem of objects, given a set of measurements by transferring tasks in the space of parameters of their models is proposed. To this end, we developed a multistage GMDH algorithm that finds the models overall structure and optimal complexity at the same time for all the set of classification objects. An iterative procedure for finding the inverse matrix in the calculation of the model parameters is used. This reduces the time of searching and selecting models. Each new model is created on base of the best model of the previous stage, and the inverse informational matrix for finding model parameters is calculated using “framing” formula for the inverse informational matrix of previous model.*

## Keywords

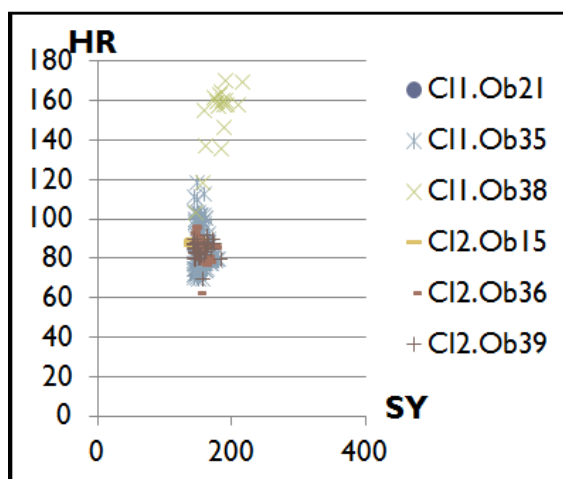
Multistage GMDH algorithm, classification,  
set of measurements, class, object,  
feature space, space of model parameters.

## 1 Introduction

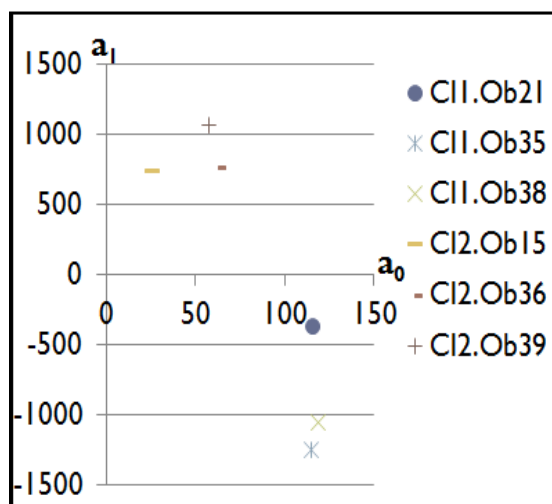
The object is defined not by the single measurement in a features space, but by their subsets, and the partial intersection of subsets is possible for each of them from different classes, is an actual case of classification solution problem. Set of load characteristics of patients' cardiovascular system is an example of mentioned objects. This problem formulation requires the development of special approach to be able to deal with sets of measurements, or with some characteristics, obtained on these sets, or with some parameters of different expedient decompositions of available characteristics in model range [1]. To solve this problem, the transfer the classification solution problem from the initial feature space to the space of model parameters of classification objects is proposed.

## 2 Theoretical Part

Each of six objects is represented with a set of points in feature space, which could not be classified, on Figure 1. Each object is represented with only one point in the space of model parameters on Figure 2 and the objects could easily be classified.



**Fig.1.** Input data: HR-heart rate, SY- systolic blood pressure (objects location in the original feature space)



**Fig. 2.** Input data ( after transfer to the space of model parameters objects)

The multistage GMDH (group method of data handling) algorithm is suggested to determine the model structure for which the objects classification could be done the best way in its space of model parameters. Note that GMDH algorithm is well-known as an effective tool for solving modeling problems in different application areas [2,3]. The method features, the possible it's types, properties and features, borders application of the method described here [4].

The task of the suggested algorithm - is to identify the optimal single (for all classification objects) model structure and their individual parameters that will provide the best objects separability in the space of model parameters. Here is a brief description of the algorithm:

The input data is represented by the matrixes of measurement for each objects that are grouped in classes. Each line of such matrix is a multidimensional point that describes the object in specific conditions (Figure 1).

The results of optimization of GMDH algorithms mentioned in [5] are used in this paper: the «framing» method allows to find inverse matrix by recursive procedure [6] and was used to reduce the time of calculating model parameters. The partial models are the subsets of complete polynomial that made from initial features, the maximum degree of which is set by us.. The selection algorithm consists of generation and analysis of all possible structures that can be obtained from a given maximum degree polynomial with the rejection of those branches of structures tree that do not improve

models according to the used selection criterion. The calculation of parameters of generated structures is executed then. For the received models, the selection criterion value should be found out. The criterion of separability of classes in the space of model parameters, that is calculated on this stage of algorithm, is used as a selection criterion. The best model is chosen by the minimum value of selection criterion.

### 3 Conclusions

The only optimal structure for all initial objects and respective coefficients of received structure for each object are obtained in algorithm execution result. Each object is defined only by single point in the space of model parameters now (Figure 2). The next task is to construct the classifiers but it is not a problem now since a lot of effective multidimensional classification methods are already developed.

### References

- [1] Knyshov G.V., Nastenko Ie.A., Kondrashova N.V., Nosovets E.K., Pavlov V.A., The combinatorial algorithm for constructing a parametric feature space for classification of multidimensional models // *Inductive modeling of complex systems*. Col. of Sc. works. // Kyiv: IRTC IT&S, 2014. (in russian)
- [2] Fainzilberg L.C, Algorithms for self-organization in the problem of synthesis of information technology signal processing / Application of GMDH in solving problems // *Control systems and machines. International Scientific Journal*. – 01/04/2003 . - №2 – pp.112-120. (in russian)
- [3] Bojarinov U.G, Dli M.I., Stojanova O.V., The application of neural-fuzzy group method of data handling for the construction of models of social-economic systems// *Software and systems. International Journal*. – 22/09/2006/. - №3. (in russian)
- [4] Ivakhnenko A. G., Stepashko V. S. Noise Immunity of Modelling Kiev: «*Naukova dumka*», 1985. - 216 p. (in russian)
- [5] Stepashko V.S. Combinatorial GMDH algorithm with the optimal scheme of models sorting-out, *Soviet Automatic Control*, No. 3, pp. 31-36, 1981. (in russian)
- [6] D. K. Faddeev, V. N. Faddeeva, “Computational methods of linear algebra”, Computational methods of linear algebra. Parallel computing, Zap. Nauchn. Sem. LOMI, 54, “*Nauka*”, Leningrad. Otdel., Leningrad, 1975, 3–228. (in russian)