ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 5, Issue 4, April 2016

Self-organization of Polynomial Regression Models in Neural Structures of Geometric Transformations

Roman Tkachenko, Sergii Demianchuk

Abstract – In this paper presented methods for building self-organized polynomial regression models using input signal's functional expansion. The aim is to produces better prediction results using new developed prediction methods, compared to the known algorithms, prediction based on auto-associative method, the KNN method (k-nearest neighbors), and the group method of data handling (GMDH). Input signal's functional expansion is performed using the set of Kolmogorov-Gabor polynomials. Principal components are used to construct polynomial of the Kolmogorov-Gabor. The results of using newly developed methods prove effectiveness of the forecasting ability for the samples of a large size.

Index Terms – polynomial regression models, neural network, auto associative neural network, group method of data handling, model of geometric transformation.

I. INTRODUCTION AND LITERATURE ANALYSIS

Self-organization of polynomial regression models based on the GMDH

Theory of the models self-organization and the group method of data handling (GMDH) that is based on it represent the regression analysis based on the small amount of experimental data. The GMDH actually implements the integration of known regression analysis methods and various ways of regularization, ensuring the unity principle of mathematical model that is obtained by the GMDH for the given training data [1]. Theory of the GMDH ensures improved results of regression analysis by adapting it to the so-called direct modeling of complex systems, when short samples are used. The majority of approximation problems that are solved by the help of the GMDH algorithms are brought to the direct recovering of functional dependency by a small number of the given points (interpolation nodes). Thus, for example, the problems of identification require recovering of the object characteristics according to the experimental points (its "response hypersurface"). The peculiarity of this problem consists in ensuring "protection" of all the arguments where the GMDH algorithm [2] does not exclude ineffective arguments. At the same time, in the prediction and pattern recognition problems, conversely, any protection of the arguments can decrease efficiency of prediction. In the GMDH algorithms, the regulating influence must be saved in the equation for the optimal management and normative prediction. Consequently, the actual differences in the problem statement lead to certain differences in the appropriate GMDH algorithms. Simultaneously, all these algorithms have the common base – the principle of selection and self-organization.

The efficiency of the GMDH methods is substantially decreased for the multidimensional problems and large data samples due to considerable time delays. The use of the GMDH methods requires highly qualified users, especially for the cases of the almost degenerate multi-parameter prediction problems, or classification problems that are represented by short and voluminous samples of noisy data, if the correlation exists between the parameters, which impel toward the search of new approaches and methods of building the corresponding regression models.

Peculiarities of the neural networks of geometric transformations

Reliable tool for solving complex, almost degenerate problems is artificial neural networks (ANN). Taking into consideration certain difficulties in training and setting up traditional ANN, namely multilayer perceptrons [3], we will look into the ability to utilize the models of geometric transformations (MGT) [4] as the basis of the method for building polynomial models of neural-like structures. The models of geometric transformations possess properties that ensure both the high accuracy and speed comparing with traditional tools of neural networks, and the algorithms uniformity when there are differences in problem statements. Particularly, the MGT is characterized by a speedy non-iterative training ability if the number of steps for calculations is defined beforehand [5]. This provides the ability to reach full repeatability of training results and offers an opportunity to solve large problems. The high accuracy and improved generalized properties of the MGT [6] structures also provide the ability to obtain solutions for training samples of the smaller volume. The training process

Manuscript received April, 2016
***Roman Tkachenko**, Department of Computer Science and Information Technologies, Lviv Polytechnic National University, Lviv, Ukraine*
***Sergii Demianchuk**, Department of Computer Science and Information Technologies, Lviv Polytechnic National University, Lviv, Ukraine.*

of such structures can be complemented by optimizing adaptation procedures for the additional training criteria.

I. THE AIM AND THE TASK OF THE RESEARCH

The aim of the research is to synthesize algorithms for automated construction of poly-parametric polynomial models for almost degenerate problems, which could be represented by large and small volumes of data.

The main objective of the work is to develop certain methods for building self-organizational polynomial regression models using signal's functional expansion. Signal's functional expansion is based on the model of the geometric transformations. Functional expansion of input signals is implemented using Kolmogorov-Gabor polynomials.

II. STATEMENT OF THE REGRESSION PROBLEM

Basics of the proposed method can be illustrated based on the following example of problem solution "Predicting material breakdown voltage (dielectric strength) based on the known independent characteristics of dielectric layer MgO" (Table 1). Partial view of the input parameters for the example with dielectric layer MgO is presented in fig. 3.

The total data sample includes 31 vectors containing 3 inputs and 1 output. Each input vector corresponds to one experiment with different characteristics of the dielectric layer. Provided the right complex is used, the main goal of building a neural network is to ensure maximum accuracy in real utilization, where a model of optimal complexity in the sense of the GMDH should be created. To solve the problem of setting up the structure of the complex, the available data sample will be divided into 3 subsets.

1. The data sample for neural network training.
2. Test sample for testing and debugging.
3. The sample for the final inspection of neural network properties of the complex.

It is important to note that the number of experiments is limited and the trained structure of the MGT must be efficient in learning on a limited number of training data. It should be noted that the following method procedures can be applied universally and transported intact to other prediction problems.

Description of the input parameters of the example with dielectric layer MgO

t – Process duration of the layer receiving;

a – The thickness of the resulting layer;

I1 – Deficiency;

Ubr – Layer breakdown voltage (dielectric strength)

Specification of the mathematical model construction approach

The basis of the approach is the method of building the Wiener series based on the MGT networks. This series is also known as the "Kolmogorov-Gabor polynomial" [7]. Sometimes, in the literature, it is called Ito's expansion [8]. In general, the Wiener series can be viewed as a discrete analogue of the Volterra series [9].

Generally, the Wiener series states the following:

$$Y(x_1,...,x_n) = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} x_i x_j x_k + ... \quad (1)$$

where: x_i, x_j, x_k – dependent input signals

The choice of the Wiener series as an effective approximation can be theoretically justified by the Weierstrass theorem [10]. The accuracy can be improved by increasing the degree of the polynomial. Usage of the Wiener series (1) for high degrees from the set of tabular data is limited to the following circumstances:

1) Too many members of the series;
2) Large generalization errors (when predictions are made based on the unknown data), which also occur as small error by training (based on the data used to calculate the coefficients of the series).

As noted above, the removal of the given limitations can be achieved by using the GMDH based on a genetic algorithm for synthesis of optimal model complexity.

III. FUNDAMENTALS OF THE METHOD FOR SYNTHESIS WIENER POLYNOMIAL BASED ON NEURAL NETWORKS MGT:

- Construction of the polynomial is performed not by the input data attributes, but by their principal components (PC) that were outlined by the dedicated auto-associative neural network on the combined sample of the input data (for training and use);
- Dispersion of PC according to the MGT neuro-paradigm decreases from the first and to the last one;
- PC with a very small dispersion are discarded (**only the dispersion calculated for the given training sample is considered**), it is logical since the output is formed as a linear polynomial of PC and if the curvature of the outputs for the last PC is small (i.e. the coefficients are small for them), their discard will not worsen the accuracy noticeably because of a minor contribution of small summands, but if the curvature is large and such PC are not discarded, then the accuracy in the utilization mode will be dramatically decreased (small deviations in these PC will give large emissions to their summands – the case of near degeneracy) [11];
- The discard threshold of PC in the existing library of the MGT structures has been established and is implemented automatically, but it is very small, so we propose to make it a controlled value – Δ, which would be established for different problems, but simultaneously maintaining the existing threshold;

- The mean-square deviation is calculated – σ for each PC, only those PC for which $\sigma \geq \Delta$ are considered.

Neural network synthesis of Wiener polynomial implemented as a neural network complex is made by a cascade of two artificial neural MGT networks [12] where the first one is ANN in the auto-associative mode of utilization [13, 14]. The Auto-associative neural network (AANN) is built and trained each time in a new corresponding cycle, using the results of the previous cycle as the input data for next cycle. In the first cycle, the ANN sorts out the set of vectors $\overline{PC}^1$ based on the set of input vectors $\overline{X}$ after discarding PC with a small standard deviation (SD), and the set of vectors $\overline{PC}^1$ is considered as a set of new input vectors. For new inputs, full quadratic extension is performed, new inputs and their extensions are submitted to the new AANN (trained beforehand), and a new set of vectors $\overline{PC}^1$ is defined; after discarding PC with a small standard deviation, new PC are treated as newer inputs. For these inputs, full quadratic extension is performed [15], and new inputs and their extensions are submitted to the AANN and new inputs are defined over again.

IV. SOLUTION RESULT OF THE DESCRIBED PROBLEM USING DEVELOPED SOFTWARE PRODUCT

Based on the proposed approach, a program in C# on the .Net Framework 4.0 software platform was developed.
Program execution:

1. User starts the program
2. Provided that program was successfully started, the user can use one active option “Download training sample” on the “File” menu and the “Exit” option.
3. Provided that training sample was successfully downloaded, created and trained network of the fast-transversal-filter (FTF) type [16], and updated tables with the data for “Training Sample” tab are shown in Fig. 1.

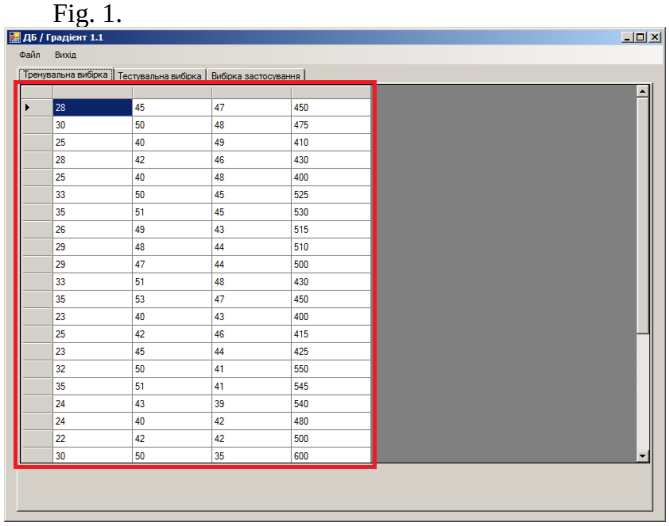

Fig.1. Partial view of the input parameters for the example with dielectric layer MgO

3.1. The options in the “File” menu become active:
“Download testing sample”
“Download the application sample”

4. Provided that training set was successfully loaded, we have next results:

Usage of the network created in step 3.1 for the testing sample. Generated table and graph in the “Testing sample” tab. The graph is built using Windows Form control ZedGraph [17].
Quantitative values of prediction errors (Fig. 2)

In the fig. 2 date in columns represent the following values: t – process duration of the layer receiving; a – the thickness of the resulting layer; I1 – deficiency; Ubr – layer breakdown voltage (dielectric strength); estimated layer breakdown voltage Ubr, for each input vector in the testing sample.

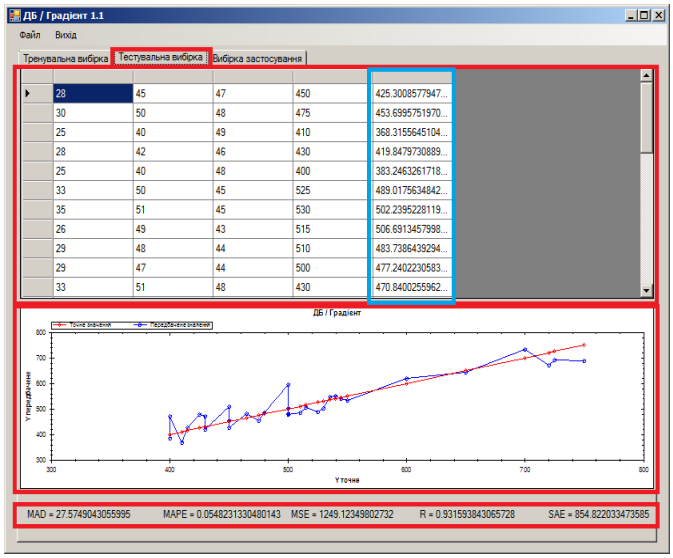

Fig.2. Quantitative values of prediction errors

5. After creating and testing the new FTF network [18, 19], the trained network can be applied using the “Download the application sample” menu item.
6. The results of downloading the application sample are:
   a. Table results on the “Application Sample” tab (Fig. 3)
   b. Prediction results in the last column (Fig. 3)

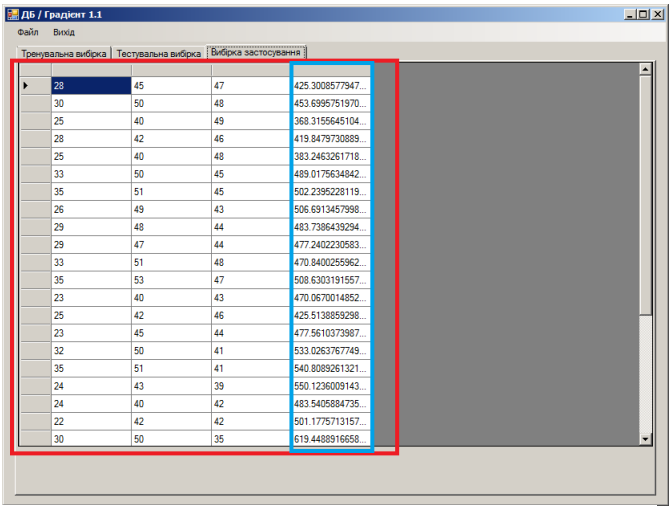

Fig.3. Trained network results

7. By clicking "Save the application sample", the results of the program can be saved to a file for the later use.

The example shows that the relative error of prediction is 5.48%, which demonstrates very high quality of prediction, even with a short training sample [20, 21].

V. ADDITIONAL EXAMPLE OF THE SOLUTION FOR PROBLEM WITH THE LARGE AMOUNT OF TRAINING SAMPLE

As a demonstration example, experimental data, that describe the load power grid based system, is taken. Training sample consists of 365 input vectors that contain 11 input columns and one output. The test sample contains 214 vectors, respectively. Developed program demonstrates high prediction accuracy, when solving problems with large volumes of training data. Therefore, relative error of prediction equals to 3.52%. Figure 4 shows program use result.

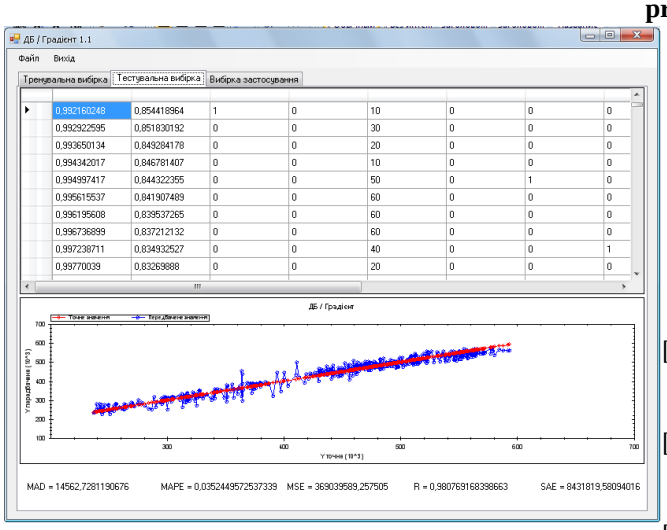

Fig.4. Result of applying the developed program for large volumes of training data

VI. CONCLUSIONS

Developed method allows to shape polynomials of high degrees in the structure of neural network, which provides, in contrast to existing methods, accurate modeling of dependencies with significant nonlinearity. The peculiarity of the proposed method is that it is formed in polynomial neural structure and can be applied implicitly, as a result of neural structure operation, or explicitly, as a result of neural structure training. The last statement is appropriate for a relatively small number of variables and small degree of the polynomial.

The proposed algorithm produces better prediction results compared to the known algorithms, prediction based on auto-associative method, the KNN method (k-nearest neighbors), and the group method of data handling (GMDH).

The basic idea of the proposed algorithm is that the method of predicting is not done by the attributes of the input data but by principal components. This allows separating data that has the biggest influence on the outcome prediction in the beginning, and using the quadratic extension especially for it; the data that has the least impact on the result should be applied in the linear form only or should be discarded according to the threshold. This approach reduces the likelihood of very noisy data in the outcome prediction.

The idea of using a threshold for discarding principal components allows adjusting the proposed complex for problems with various inputs. The main components that can only increase the prediction error are discarded on the each iteration before the process of quadratic extension. Thus, the influence factor of the condition when small input values give large emissions in the approximator output can be effectively controlled by the developed program.

The proposed method is implemented in the software product "DB/Gradient", which provides the following properties:

- Versatile usage for data of different volumes, both small-scale and high-scale;
- High accuracy when using trained predictor;
- No special requirements for mathematical qualifications and computer skills of the user.

REFERENCES

[1] Ivakhnenko A. G. Long-term forecasting and management of complex systems. - K .: "Tehnika" 1975. (in Russian)

[2] Ivakhnenko A.G. Polynomial Theory of Complex System. IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-1, No. 4, Oct. - 1971

[3] Widrow, B., Lehr, M.A., "30 years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation," Proc. IEEE, vol 78, no 9. - 1990.

[4] Hrytsyk V. V., Tkachenko R. O. New approaches to the study of artificial neural networks // Reports of the National Academy of Sciences of Ukraine.-

№ 11 - 2002. (in Ukrainian)

[5] Tkachenko R. O. The new paradigm of artificial neural networks of direct distribution // Bulletin of the State University "Lviv Polytechnic" Computer Engineering and Information Technology. - № 386. - 1999. (in Ukrainian)

[6] Tkachenko R. O., Doroshenko A. V. Model basis that on the top of the model of geometric transformations for decision support systems // Computer technology of printing. - № 17. - 2007. (in Ukrainian)

[7] Gabor D., Wilby W.R., Woodcock R.A. A universal nonlinear filter, predictor and simulator which optimizes itself by a learning process. Proc. Inst. Electr. Engrs., vol. 108., part B, № 40, 1961. — P. 85-98.

[8] Kolmogorov A. N. Interpolation and extrapolation of stationary random sequences // Math. USSR Academy of Sciences. Ser. Mat., vol. 5: 1, 1941. (in Russian)

[9] Weiner N. The Extrapolation Interpolation and Smoothing of Stationary Time-Series. I. Willey, N.Y., 1949. — 290 p.

[10] Weierstrass K., Abhandlungen aus der Funktionenlehre, B., 1860; Math. Werke, Bd 2, B., 1895.

[11] Hrytsyk V. V., Tkachenko R. O. New approaches to the study of artificial neural networks // Reports of the National Academy of Sciences of Ukraine.- № 11 - 2002. (in Ukrainian)

[12] Tkachenko R. O., Andriyetskiy B. Improving the accuracy of neural structures of geometric transformations. National University "Lviv Polytechnic" 2011. (in Ukrainian)

[13] Tkachenko R. O., Doroshenko A. V. Neural structures machinery geometric transformations in the tasks of data mining // International Conference "Intelligent Decision Support Systems and problems of computational intelligence» ISDMCI'2009: Coll. Science. pr. 2 vols., May 18-22. 2009, Evpatoria, Ukraine. -X. ; Kherson, 2009. - Vol 2. (in Ukrainian)

[14] Polishchuk U., Tkachenko P., Tkachenko R., Yurchak I. Features of the auto-associative neurolike structures of the Geometrical Transformation Machine (GTM) // Advanced technologies and methods for designing MEMS materials Fifth international.conf Young Scientists. MEMSTECH'2009, 2009. (in Ukrainian)

[15] Mladenov V., On the recurrent neural networks for solving general quadratic programming problems. // Neural Network Applications in Electrical Engineering, 2004.

[16] Haykin, Simon. Adaptive Filter Theory, 4th Ed. Upper Saddle River, NJ: Prentice Hall, 2002

[17] ZedGraph: A flexible charting library for .NET. Available online at: http://www.codeproject.com/Articles/5431/A-flexible-charting-library-for-NET

[18] Bodyanskiy E., Vynokurova O., Peleshko D. Information technology clustering of data in a short training sample on the basis of associative probability the neuro-fuzzy system. Control systems and machines. - 2014. - №4. - 2014. (in Russian)

[19] Bodyanskiy E., Vynokurova O., Pliss I., Peleshko D. Multilayer neuro-fuzzy system for solving on-line diagnostics tasks. Computation Models for Business and engeneering domains. ITHEA, Rzeszow, Sofia (INFOS 2014, 7th International conference on intelligent information and engeneering systems, September 16-20, Rzeszow-Krynica-Zdroj, Poland). 2014. (in Polish)