Revised GMDH-type neural network algorithm self-selecting optimum neural network architecture

C. Kondo[1] and T. Kondo[2]

[1] *Graduate school of Design, Kyushu Univ., 4-9-1shiohara minamiku Fukuoka 815-8540, Japan*
[2] *Graduate School of Health Sci., The Univ. of Tokushima, 3-l8-15 Kuramoto-cho Tokushima 770-8509, Japan*
([2] *kondo@medsci.tokushima-u.ac.jp*)

***Abstract*:** In this study, the revised Group Method of Data Handling (GMDH)-type neural network algorithm self-selecting the optimum neural network architecture is applied to the identification of the nonlinear system. In this algorithm, the optimum neural network architecture is automatically organized using two kinds of the neuron architectures such as the polynomial and sigmoid function neurons. Many combinations of the input variables, in which high order effects of the input variables are contained, are generated using the polynomial neurons and useful combinations are selected using Prediction Sum of Squares (PSS) criterion. These calculations are iterated and the multi-layered architecture is organized. Furthermore, the structural parameters such as the number of layers, the number of neurons in the hidden layers and the useful input variables are automatically selected so as to minimize the prediction error criterion defined as PSS.

***Keywords*:** GMDH, Neural network, Medical image recognition

I. INTRODUCTION

The GMDH-type neural networks [1], [2] are automatically organized by using the heuristic self-organization method [3]. In the GMDH-type neural networks, the structural parameters such as the number of layers, the number of neurons in each layer, useful input variables and optimum neuron architectures are automatically determined. Furthermore, many types of neurons such as polynomial type, sigmoid function type and radial basis function type neurons, are used for organizing neural network architecture and therefore many types of nonlinear systems can be identified by the GMDH-type neural networks.

In this study, the revised GMDH-type neural network [1] is applied to the nonlinear system identification. In the revised GMDH-type neural network, the polynomial type and the sigmoid function type neurons are used for organizing the neural network architecture. A lot of complex nonlinear combinations of the input variables fitting the complexity of the nonlinear system are generated by using the polynomial type neurons and only useful combinations of the input variables are selected for organizing the neural network architecture. In the output layer, the sigmoid function type neurons are used for organizing the neural network and the output values of the neural network become between zero and one. It is shown that the revised GMDH-type neural network is accurate and useful method for the nonlinear system identification.

II. REVISED GMDH-TYPE NEURAL NETWRK ALGORITHM

In the conventional GMDH-type neural network [2], following neuron architectures are used for organizing neural network to fit the complexity of the nonlinear system.

1) First type neuron

$\sum$: (Nonlinear function)

$$z_k= w_1u_i+w_2u_j+w_3u_iu_j+w_4u_i^2+w_5u_j^2+w_6u_i^3+w_7u_i^2u_j +w_8u_iu_j^2+w_9u_j^3 - w_0\theta_1 \tag{1}$$

f : (Nonlinear function)

$$y_k= 1 / (1+exp (- z_k)) \tag{2}$$

2) Second type neuron

$\sum$: (Nonlinear function)

$$z_k= w_1u_i+w_2u_j+w_3u_iu_j+w_4u_i^2+w_5u_j^2+w_6u_i^3+w_7u_i^2u_j +w_8u_iu_j^2+w_9u_j^3 - w_0\theta_1 \tag{3}$$

f : (Linear function)

$$y_k= z_k \tag{4}$$

3) Third type neuron

$\sum$: (Linear function)

$$z_k= w_1u_1+w_2u_2+w_3u_3+ \cdots +w_ru_r - w_0\theta_1 \qquad (r<p) \tag{5}$$

f : (Nonlinear function)

$$y_k= 1 / (1+exp (- z_k)) \tag{6}$$

4) Fourth type neuron

$\sum$: (Linear function)

$$z_k= w_1u_1+w_2u_2+w_3u_3+ \cdots +w_ru_r - w_0\theta_1 \qquad (r<p) \tag{7}$$

f : (Linear function)

$$y_k= z_k \tag{8}$$

5) Fifth type neuron

Σ: (Nonlinear function)

$$z_k= w_1u_i+w_2u_j+w_3u_iu_j+w_4u_i^2+w_5u_j^2+w_6u_i^3+w_7u_i^2u_j+w_8u_iu_j^2+w_9u_j^3 - w_0\theta_1 \quad (9)$$

f : (Nonlinear function)

$$y_k= exp\ (- z_k^2) \quad (10)$$

6) Sixth type neuron

Σ: (Linear function)

$$z_k= w_1u_1+w_2u_2+w_3u_3+\cdots+w_ru_r - w_0\theta_1 \quad (r<p) \quad (11)$$

f : (Nonlinear function)

$$y_k= exp\ (- z_k^2) \quad (12)$$

7) Seventh type neuron

Σ: (Linear function)

$$z_k= w_1u_1+w_2u_2+w_3u_3+\cdots+w_ru_r - w_0\theta_1 \quad (r<p) \quad (13)$$

f : (Nonlinear function)

$$y_k= a_0+a_1z_k+a_2z_k^2+\cdots+a_mz_k^m \quad (14)$$

Here, $\theta_1=1$ and w_i $(i=0,1,2,\cdots)$ are weights between the neurons. The optimum neuron architectures fitting the complexity of the nonlinear system are automatically selected by using the PSS [4]. Therefore, many kinds of nonlinear systems can be automatically identified by using the conventional GMDH-type neural network.

In the revised GMDH-type neural network, many kinds of nonlinear combinations of the input variables are generated by using the polynomial type neurons and only useful nonlinear combinations of the input variables are selected. Optimum neural network architectures are organized by using selected useful combinations of the input variables.

The revised GMDH-type neural network is shown in Fig.1. Here, nonlinear function g_i is described by the following Kolmogorov-Gabor polynomial:

$$g_i\,(x_1,x_2,\cdots,x_p)= a_0+\sum_i a_ix_i+\sum_i\sum_j a_{ij}x_ix_j+\cdots \quad (15)$$

This nonlinear function is automatically organized by using the second type neuron of the conventional GMDH-type neural network. The architectures of the revised GMDH-type neural network is produced as follows:

First, the architecture of the first layer is organized.

1. The first layer

$$u_j=x_j \qquad (j=1,2,\cdots,p) \quad (16)$$

where x_j $(j=1,2,\cdots,p)$ are the input variables of the system, and p is the number of input variables. In the first layer, input variables are set to the output variables.

2. The second layer

Many combinations of two variables (u_i,u_j) are generated. For each combination, the neuron architecture is described by the following equations:

Σ: (Nonlinear function)

$$z_k= w_1u_i+w_2u_j+w_3u_iu_j+w_4u_i^2+w_5u_j^2+w_6u_i^3+w_7u_i^2u_j+w_8u_iu_j^2+w_9u_j^3 - w_0\theta_1 \quad (17)$$

f : (Linear function)

$$y_k= z_k \quad (18)$$

where $\theta_1=1$ and w_i $(i=0,1,2,\cdots,9)$ are weights between the first and second layer. This neuron is equal to the second type neuron of the conventional GMDH-type neural network. The weights w_i $(i=0,1,2,\cdots,9)$ are estimated by using the revised regression analysis [5]. This procedure is as follows:

First, the values of z_k are calculated by using the following equation:

$$z_k=log_e(\phi'/(1-\phi')) \quad (19)$$

where ϕ' is the normalized output variable. Then the weights w_i $(i=0,1,2,\cdots,9)$ are estimated by using the stepwise regression analysis [5] which selects useful input variables by using the PSS[4]. Therefore, only useful terms in (17) are selected and neuron architecture can be organized by these selected useful terms.

From these generated neurons, L neurons which minimize the PSS are selected. The output values (y_k) of L selected neurons are set to the input values of the neurons in the third layer.

3. The third and succeeding layers

In the third and succeeding layers, the same computation of the second layer is continued until the PSS values of L neurons are not decreased. When the iterative computation is terminated, the following calculation of the output layer is carried out.

4. The output layer

In the output layer, the output values of the neural network are calculated from z_k as follows:

$$\phi=1/(1+exp(-z_k)) \quad (20)$$

So, in the output layer, the neuron architecture becomes as follows:

Σ: (Nonlinear function)

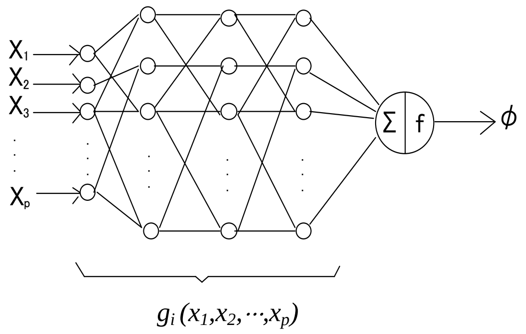
$$z_k= w_1u_i+w_2u_j+w_3u_iu_j+w_4u_i^2+w_5u_j^2+w_6u_i^3+w_7u_i^2u_j+w_8u_iu_j^2+w_9u_j^3 - w_0\theta_1 \quad (21)$$

f : (Nonlinear function)

$$\phi=1/(1+exp(-z_k)) \quad (22)$$

This neuron architecture is the same of the first type neuron of the conventional GMDH-type neural network. At last, the complete neural network architecture is produced by selected neurons in each layer.

By using above procedures, the revised GMDH-type neural network can be organized.



Σ: *(Nonlinear function)*

$z = \sum w_i\, g_i(x_1,x_2,\cdots,x_p)$

f : *(Nonlinear function)*

$\phi = 1 / (1+\exp(-z))$

Fig.1 Architecture of revised GMDH-type neural network [1]

III. APPLICTION TO NONLINEAR SYSTEM IDENTIFICATION

Nonlinear system is assumed to be described by the following equations:

$$\phi = f_1(x_1, x_2, x_3) / f_2(x_1, x_2, x_3)+\varepsilon \quad (23)$$

$$f_1(x_1, x_2, x_3)=1.0+2.0\, x_1^2x_2+3.0\, x_2^2x_3 \quad (24)$$

$$f_2(x_1, x, x_3)=1.0+2.0\exp(x_1)+3.0\exp(x_1x_2)+4.0\exp(x_3) \quad (25)$$

Here, ϕ is output variable and $x_1 \sim x_3$ are input variables and ε is Gaussian white noise which is N(0, 0.005^2). An additional input, x_4, is added as input variable of neural network to check that revised GMDH-type neural network can detect and eliminate useless input variables. Neural network is organized by twenty training data. Twenty other data are used to check prediction and generalization ability.

1. Identification results of revised GMDH-type neural network

A. Input variables

Four input variables were used, but useless input variable x_4 was automatically eliminated.

B. Number of selected neurons in each layer

Four neurons were selected in each hidden layer.

C. Variation of PSS

Variation of PSS is shown in Fig.2. PSS values converged at the fifth layer.

D. Architecture of neural network

Calculation of revised GMDH-type neural network was terminated at the fifth layer because PSS values were not decreased. Five layered neural network architecture was organized. The first layer is input layer and the second, the third, the fourth layer are hidden layers and the fifth layer is output layer.

E. Estimation accuracy

Estimation accuracy was evaluated by the following equation:

$$J_1 = \frac{1}{20}\sum_{i=1}^{20}\left|\phi_i - \phi_i^*\right| \quad (26)$$

where ϕ_i $(i=1,2,\cdots,20)$ are actual values with Gaussian white noise ε and ϕ_i^* $(i=1,2,\cdots,20)$ are estimated values by revised GMDH-type neural network. ϕ_i $(i=1,2,\cdots,20)$ were used to organize revised GMDH-type neural network. Value of J_1 is shown in Table1. In this table, Revised GMDH-type NN shows revised GMDH-type neural network developed in this paper and GMDH shows conventional GMDH algorithm.

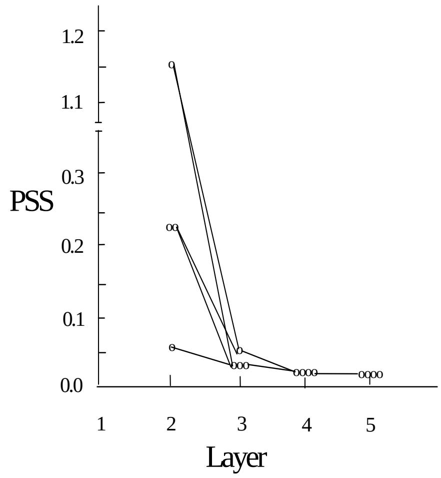


Fig.2 Variation of PSS in revised GMDH-type neural network

Table 1 Identification results of two methods

Method	J1	J2	Layer
Revised GMDH-type NN	0.02274	0.02942	5
GMDH	0.04094	0.03829	5

F. Prediction accuracy

Prediction accuracy was evaluated by using the following equation:

$$J_2 = \frac{1}{20}\sum_{i=21}^{40}\left|\phi_i - \phi_i^*\right| \quad (27)$$

where ϕ_i $(i=21,22,\cdots,40)$ are actual values with Gaussian white noise ε and ϕ_i^* $(i=21,22,\cdots,40)$ are predicted values by revised GMDH-type neural network.

ϕ_i *(i =21,22,···,40)* were not used to organize revised GMDH-type neural network and were used to check generalization ability. Value of J_2 is shown in Table1 and is very small. From this prediction result, we can see that revised GMDH-type neural network do not overfit training data and have good generalization ability.

G. Estimated and predicted values

Estimated and predicted values of ϕ by revised GMDH-type neural network are shown in Fig.3. Estimated values are shown for the data points between the first and 20-th data entities and predicted values are shown for the data points between the 21-th and 40-th data entities. We can see that estimated and predicted values are accurate.

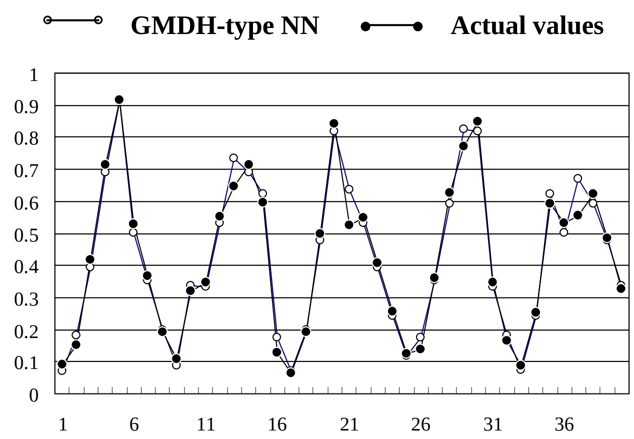


Fig.3 Estimated and predicted values by revised GMDH-type neural network

2. Identification Results of GMDH

Identification results of GMDH algorithm are shown as follows:

Four input variables were used, but again useless input variable x_4 was automatically eliminated. Four intermediate variables were selected in each selection layer. Variation of PSS is shown in Fig.4. PSS converged at the fifth layer. Calculation of GMDH converged at the fifth layer and neurons of the fifth layer had the minimum PSS value. Five layered polynomial network architecture was organized. The first layer is input layer and the second, the third and the fourth layer are hidden layers and the fifth layer is output layer. Estimation accuracy was evaluated by Eq.(26) and value of J_1 is shown in Table1. Prediction accuracy was evaluated by Eq.(27) and value of J_2 is shown in Table1.

VI. CONCLUSION

In this paper, a revised GMDH-type neural network algorithm self-selecting optimum neural network architecture was applied to the nonlinear system identification. In this algorithm, optimum neural network architecture is automatically organized using the polynomial and sigmoid function type neurons. It was shown that revised GMDH-type neural network algorithm was a useful method for the nonlinear system identification.

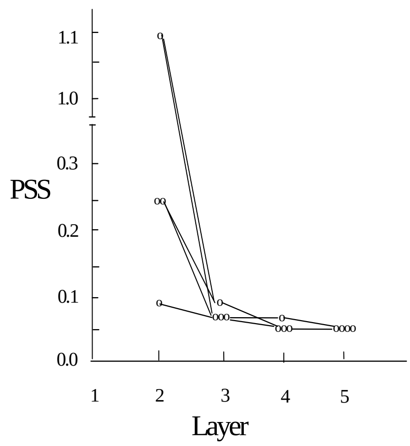


Fig.4 Variation of PSS in GMDH

REFERENCES

[1] Kondo T. and J.Ueno, Logistic GMDH-type Neural network and Its Application to identification of X-ray Film Characteristic Curve, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Voll.11, No.3, PP.312-318, 2007.

[2]. Kondo T., GMDH neural network algorithm using the heuristic self-organization method and its application to the pattern identification problem, *Proc. of the 37th SICE Annual Conference*, pp.1143-1148,1998.

[3] Ivakhnenko A.G., Heuristic self-organization in problems of engineering cybernetics, *Automatica*, Vo.6, No.2, pp.207-219, 1970.

[4] Tamura H. and T.Kondo, Heuristics free group method of data handling algorithm of generating optimum partial polynomials with application to air pollution prediction, *Int. J. System Sci.*, Vol.11, No.9, pp.1095-1111, 1980.

[5] Draper N.R. and H.Smith, *Applied Regression Analysis*, John Wiley and Sons, New York, 1981.