Proceedings of the 36th Chinese Control Conference
July 26-28, 2017, Dalian, China

GMDH-type Neural Network for Remaining Useful Life Estimation of Equipment

Lin Zhao[1], Yipeng Wang[1], Yuan Liu[1], Yong Hao[1]

1. College of Automation, Harbin Engineering University, Harbin 150001, China
E-mail: spacead@163.com

Abstract: The Group Method of Data Handing (GMDH)-type neural network algorithm is proposed to solve the problem of network structure design when using traditional neural network to predict Remaining Useful Life (RUL) of equipment. The Principal Component Analysis (PCA) algorithm is used to process the initial input data, which reduces the computational burden of the network. Using the Prediction Error Sum of Square (PESS) to select the hidden layer neurons, and the PCA method to limit the number of hidden neurons. Using the actual motor operating data to validate this algorithm, the results show that this method can adaptively construct equipment failure network model, avoid network structure selection problem, and has strong generalization ability.

Key Words: Machine Learning, Prognostics, GMDH, Neural Network, Remaining Useful Life

1 Introduction

With the development of science and technology be applied to complex industrial and military fields, more and more attention has been paid to the health management of complex systems. In the field of aerospace, the technology of fault prediction and health management is an effective way to improve the affordability and safety of aircraft economy, which technology has been defined as one of the core technologies for breakthrough for advanced aircraft [1].

The most common and important application technology of predictive science in system health management is to estimation of the RUL of the system equipment. The data-driven approach for RUL estimation normally relies on the availability of run-to-failure data, based on which the RUL can be estimation followed by extrapolation to the damage progression [2]. Artificial neural network is a kind of advanced calculation method of driving data, which has been applied to many filed, such as pattern recognition, signal processing, economic forecasting, modeling, system identification, process monitoring and fault detection, fault-tolerant control and biomedical engineering [3]. Due to the good nonlinear approximation ability of artificial neural network, in recent years, the use of BP neural network on the prediction of residual life of the equipment has achieved good effect. However, the traditional neural network need to design the network topology, and the unreasonable network structure will result in the lack of fitting or the consequences of excessive training costs. In order to reduce the workload of modeling, a GMDH-type neural network is proposed [4]. The GMDH-type neural network can automatically organize the neural network architecture by using a heuristic self-organization method which is the basic premise of the GMDH algorithm [5]. Due to the GMDH-type neural network method does not need to know the prior knowledge of the input and output variables, the network structure is determined dynamically in training, GMDH-type network has been widely used in the modeling and prediction of complex systems.

*This work is supported by National Natural Science Foundation (NNSF) of China under Grant 61633008.

This paper focuses on using the algorithm of GMDH-type neural network to estimate the RUL of equipment. First of all, using classical GMDH-type neural network to predict the RUL of a motor, and find out the shortcomings of the classic method; Secondly proposed a method using PCA method to assisted selection hidden layer neurons; Finally, the algorithm of GMDH-type neural network based on PCA method is used to deal with the remaining life test data of the motor, and the remaining life of the motor is predicted and compared with the classical neural network method. The results show that the proposed method can effectively predict the RUL of the motor and avoid the structural design of the BP neural network.

2 RUL Prediction Using Classical GMDH-Type Neural Network

The concept of the GMDH approach relies on replacing the complex neural model by the set of hierarchically connected partial models. The model is obtained as a result of neural network structure synthesis with the application of the GMDH algorithm [6]. The synthesis process consists of a partial model structure selection and parameters estimation. In the next step of the process synthesis, the partial model are evaluated, selected and included to newly created neuron layers. During the network synthesis new layers are added to the network. The process of network synthesis leads to the evolution of the resulting model structure to obtain the best quality approximation of real system output signals [7].

In the process of self-organization of the network, the most important two step is to determine the parameters of neurons and to determine the rules of hidden layer neurons selection. The parameters of neurons can be estimated by using existing linear estimation methods. Estimation can reach an ideal result using the LMS method of polynomial parameters, then polynomial output as input of a nonlinear function and the output results which is multiple neurons nonlinear function's superposition can approach any nonlinear function curve. The selection rule of neurons is mainly based on the calculation of the process error. The calculation of the process error can be determined by FPE, AIC or F-test [8]. Based on the above method, we can choose the appropriate neurons as the input of next layer.

The neuron selection criteria is generally determined by the PESS of the upper layer. We often use the average or minimum of PESS as the criterion of neuron selection. However, using the mean value of the upper layer PESS often makes the number of hidden neurons increased and causes the heavy computational burden, and using the minimum value of upper layer PESS often makes the number of hidden neurons decreases sharply resulting in network can not converge to the minimum.

The data used in this paper is from the IEEE 2008 PHM Conference challenge motor failure test data. There are great differences in the RUL of motors, and the failure process is a dynamic process. We assume that the initial RUL of the motor is an unpredictable constant value, in the subsequent operation due to the emergence of the damage, the remaining useful life of the motor is shortened and eventually damaged, so its life curve can be described as: In the initial stage of the motor, the motor is in good health, and the RUL is the maximum value of the factory $Y_{\max}$, during the subsequent operation, the remaining life of the motor decreases linearly with a slope of -1.

Using the mean value of the upper PESS as the neuron selection criteria, the network construction process and the data prediction results are shown in the Fig.1.

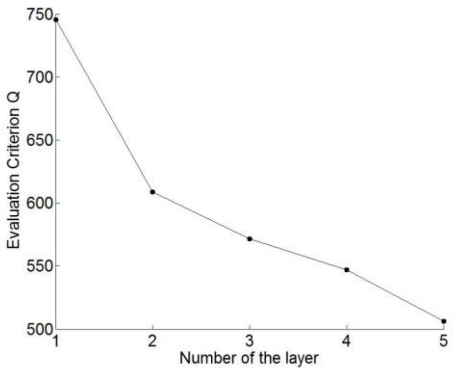

Fig. 1 Classical GMDH-type neural network construction process

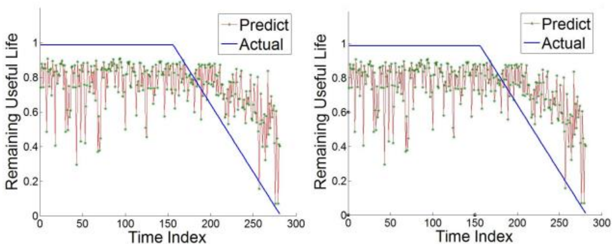

Fig. 2 The results of fifth and the twenty-fifth data set

As is show in Fig.1, the horizontal axis represents the network layers, and the vertical axis represents the smallest PESS of each layer. The number of hidden neurons in the process were 6, 12, 11, 32 and 152. Too many neurons in the fifth hidden layer lead to a huge computational burden. Fig.2 is the result of using the classical method to predict the RUL of motor with the second data set and the twenty-fifth data set of the validation data set. The horizontal axis represents time index, the vertical axis represents the remaining life in Fig.2. The initial life of the equipment is 1, with the use of the equipment life gradually reduced, the final life value of 0 indicates that the damage. As is show in Fig.2, the residual life prediction results using traditional methods have a large deviation from the actual results. In order to reduce the prediction error and avoid the excessive number of neurons in the hidden layer, we need to design a new method to select the hidden layer neurons.

3 GMDH-type neural network based on PCA method

PCA is a multivariate statistical analysis method, which is widely used in many fields, such as management, data statistics, process monitoring and so on. Using PCA method for the selection of hidden layer neurons can effectively avoid the excessive number of neurons in the hidden layer. The process of using this method to build neural network is as follows:

1) Preprocessing the input data

It is assumed that $\boldsymbol{X} \in R^{L \times N}$ is the input data, $\boldsymbol{Y} \in R^{L}$ is the output data, where L is the number of samples and N is the input data dimension. Firstly, the output data are normalized to get $\boldsymbol{Y}^*$. Then the PCA method is used to reduce the dimension of input data $\boldsymbol{X}$.

There are two kinds of PCA methods, which are based on covariance and correlation. The method based on covariance is each column of $\boldsymbol{X}$ minus the corresponding variable mean and the method of correlation is each column of $\boldsymbol{X}$ minus the corresponding variable mean and divided standard deviation of the corresponding variables. In this paper, using PCA method based on correlation to deal with the problem, each column of $\boldsymbol{X}$ is processed into a variable of unit variance and zero mean. The covariance matrix of the standardized sample input is defined as:

$$\boldsymbol{C}_{\mathrm{cov}} = \frac{1}{L-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} \tag{1}$$

The Singular Value Decomposition（SVD）is used to decompose the covariance matrix $\boldsymbol{C}_{\mathrm{cov}}$ and obtain the eigenvector matrix $\boldsymbol{U}$、$\boldsymbol{V}$ and the singular value diagonal matrix $\boldsymbol{S}$, and $\boldsymbol{C}_{\mathrm{cov}} = \boldsymbol{U} * \boldsymbol{S} * \boldsymbol{V}^{\mathrm{T}}$ .The diagonal matrix element $\lambda_1, \lambda_2, \cdots \lambda_N$ of singular value diagonal matrix $\boldsymbol{S}$ is the eigenvalue corresponding to the orthogonal eigenvector in U. In U, select the number of A eigenvalues corresponding to the combination of the column as the main feature vector $\boldsymbol{U}^* \in R^{A*N}$. Finally, the processed input data $\boldsymbol{X}^*$ can be obtained by $\boldsymbol{X}^* = \boldsymbol{U}^{*\mathrm{T}}\boldsymbol{X}$.

2) Separate original data into training and test set

It is necessary to divide the input data $\boldsymbol{X}^*$ into two parts: training and test set, which are used to estimate the parameters of neurons. Finally, the input data is divided into the form like $\left(x_{k,i}, x_{k,j}\right)$, where $x_{k,i}, x_{k,j}$ represents the *i*-th and *j*-th input variable in the *k*-th sample and $i \neq j$.

3) Neuron parameter estimation

The above data are used as inputs to each neuron:

$$\hat{y}_{k,m}^{(l)} = w_{0,m}^{(l)} + w_{1,m}^{(l)} x_{k,i} + w_{2,m}^{(l)} x_{k,i} + w_{3,m}^{(l)} x_{k,j} x_{k,j} + w_{4,m}^{(l)} x_{k,i}^2 + w_{5,m}^{(l)} x_{k,j}^2 \quad (2)$$

Where $\hat{y}_{k,m}^{(l)}$ is the predicted output of the m-th neuron of the l-th layer in the k-th sample, $w_{p,m}^{(l)}\left(p=0,1,\cdots,5\right)$ is the neuron parameters obtained by regression. The output sample $\boldsymbol{Y}^*$ is used as the output of y_k, and the least square method is used to estimate the parameters of the polynomial. After obtaining the parameter $w_{p,m}^{(l)}$, we need to use the test data set to calculate the PESS of the neurons:

$$PESS_m^{(l)} = \sum_{k=1}^{L}\left(y_k - y_{k,m}^{(l)}\right)^2 \quad (3)$$

Where $PESS_m^{(l)}$ is the PESS of the m-th neuron in the l-th layer, and the $\hat{y}_{k,m}^{(l)}$ is the predicted output of the m-th neuron in the l-th layer in the k-th sample, and the y_k is the output of the k-th sample. The mean value $\overline{PESS}^{(l)}$ of $PESS_m^{(l)}$ is used as the selection criteria for the next layer:

$$\overline{PESS}^{(l)} = \frac{1}{M}\sum_{m=1}^{M} PESS_m^{(l)} \quad (4)$$

M is the number of neurons in the l-th layer.

4) Select neurons

First of all, we use the selection criterion $\overline{PESS}^{(l)}$ of the current layer to filter the neuron for the first time, the neuron which satisfies the condition $PESS_m^{(l)} < \overline{PESS}^{(l-1)}$ is selected, and m is the index number of the selected neuron. Then we get the first selection of the network output result $Y_1^{(l)}$. After the first selection, the PCA method is used to select the neurons secondly. Its process is as follows:

a. Calculate the covariance of the $Y_1^{(l)}$ after the first selection:

$$C_Y = \frac{1}{H}\left(Y_1^{(l)}\right)^{\mathrm{T}} Y_1^{(l)} \quad (5)$$

b. Singular value decomposition of covariance matrix:

$$[\boldsymbol{U_Y} \quad \boldsymbol{S_Y} \quad \boldsymbol{V_Y}] = \mathrm{svd}(\boldsymbol{C_Y}) \quad (6)$$

As show in formula (6), $\boldsymbol{U_Y}$ 、 $\boldsymbol{V_Y}$ 、 $\boldsymbol{S_Y}$ are the eigenvector matrix and eigenvalue diagonal matrix respectively , and $\boldsymbol{C}_{\mathrm{cov}} = \boldsymbol{U_Y} * \boldsymbol{S_Y} * \boldsymbol{V_Y}^T$, $svd()$ is a singular value decomposition function.

c. Select main element：

The eigenvalues of diagonal matrix $\boldsymbol{S_Y}$ are arranged from large to small to get $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_H$, and the number of principal components q is satisfied:

$$1 - \frac{\sum_{h=1}^{q}\lambda_h}{\sum_{h=1}^{H}\lambda_h} \leq 0.01 \quad (7)$$

Finally, select the corresponding eigenvector matrix $\boldsymbol{U_Y^*}$ from the eigenvector matrix $\boldsymbol{U_Y}$.

d. Calculate the input of the next layer：

$$\boldsymbol{Y}_2^{(l)} = \left(\boldsymbol{U_Y^*}\right)^T \boldsymbol{Y}_1^{(l)} \quad (8)$$

Where $\boldsymbol{Y}_2^{(l)}$ is the neuron output which is selected in the l-th layer and is used as the input of the next layer.

5) Generate next layer

The upper neuron output $\boldsymbol{Y}_2^{(l)}$ is used as the input of the next layer, and the 2) ~5) process is repeated until the termination condition is satisfied.

6) Stop calculation to get the final network structure

After generating a new network layer, comparing the current layer with the previous layer neuron minimum process error square $PESS_{min}^{(l+1)}$ and $PESS_{min}^{(l)}$, if the condition of $PESS_{min}^{(l+1)} > PESS_{min}^{(l)}$ is satisfied, the calculation is terminated. Take the corresponding neurons in the previous layer as the final output of the network. Finally, the final network structure and neuron parameters are determined.

The construction of GMDH-type neural network is completed, and the PCA method is used to assist in the selection of hidden layer neurons, which avoids the problem of too many hidden neurons in the process of self-organization construction.

4 Simulation Results

Using the IEEE 2008 PHM Conference challenge motor failure data to validate the algorithm. This data set has been divided into three parts: training set, test set and validation set. Firstly, we use the training test set to train the network.

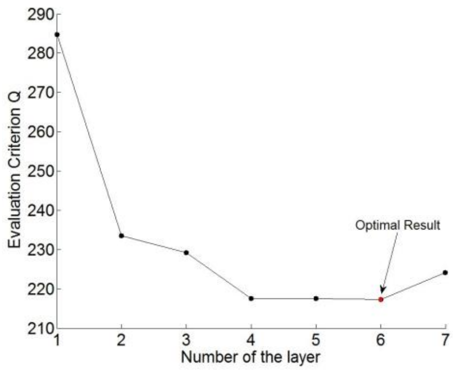

Fig. 3 GMDH-Type neural network based on PCA construction process

The GMDH-type neural network based on PCA method construction process is show in Fig.3, the optimal result is 217.33 which is obtained in the sixth layer. The number of neurons in each layer is: 28, 20, 10, 8, 4, 2, respectively. Finally selecting the neuron which have the smallest error in the network as the network output neuron.

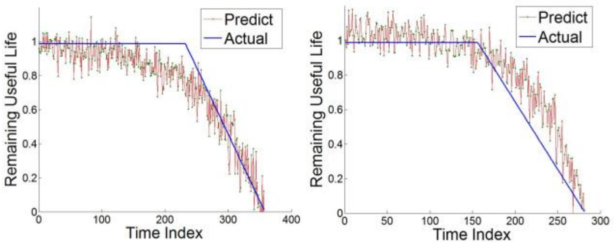

Fig.4. The results of the fifth and twenty-five data sets

The network is used to predict the validation data set , the predicted results of the fifth set and the twenty-fifth set of the data are shown in the Fig.4. It can be seen that the predicted results are significantly enhanced by comparing Fig.2 with Fig.4.

Table 1 Prediction and estimation accuracy

number of layers	Classic GMDH NN		GMDH-type network based on PCA	
	Number of neurons	Minimum value of process error	Number of neurons	Minimum value of process error
1	6	745.41	28	284.76
2	12	608.73	20	233.57
3	11	571.64	10	229.17
4	32	546.87	8	217.49
5	154	506.09	4	217.6
6	-	-	2	217.33

As is shown in table 1, the method proposed in this paper by using the PCA method to preprocess the input data, which makes the prediction error of the first layer is far less than the classical method. The final prediction error of this method is less than that of the classical method. Furthermore, the number of neurons in each layer has reduced by using the algorithm proposed in this paper.

5 Conclusion

Using the PCA method to preprocessing of input data can reduce the prediction error and reduce the computational burden. Compared with the classical method, the prediction error of the first layer of the proposed method is reduced by 61% and the final prediction error is reduced by 57%. Using PCA method to select the hidden layer neurons of GMDH-type neural network can effectively avoid the excessive number of hidden neurons.

References

[1] S. B. Johnson , T. J. Gormley , System Health Management: With Aerospace Applications , *John Wiley & Sons*,2011, chapter 1

[2] G. Kai, B. Saha, A. Saxena. A comparison of three data-driven techniques for prognostics, *Failure prevention for system availability ,62th meeting of the MFPT Society*, 2008, 119-131

[3] H. Bagan, J. Ma, Z. Zhou. Study of artificial neural network method for weather and AVHRR thermal data classification, I*nternational Symposium on Multispectral Image Processing and Pattern Recognition. International Society for Optics and Photonics*, 2003:179-182.

[4] Kondo T. GMDH neural network algorithm using the heuristic self-organization method and its application to the pattern identification problem Sice '98. *Proceedings of the, Sice Conference. International Session Papers.* IEEE Xplore, 1998:1143-1148.

[5] S J. Farlow. Self-organizing Method of Modeling, GMDH type Algorithms, *New York: Marcel Dekker Inc*1984.

[6] M. Mrugalski Confidence estimation of GMDH neural networks and its application in fault detection systems. *International Journal of Systems Science,* 2008, volume 39(8):783-800.

[7] M. Mrugalski, Advanced Neural Network-Based Computational Schemes for Robust Fault Diagnosis. *Studies in Computational Intelligence*, 2014, 510(1).

[8] Kondo T, Ueno J, Takao S. Feedback GMDH-Type Neural Network Algorithm Using Prediction Error Criterion Defined as AIC. *Intelligent Decision Technologies*. 2012:49-54.

[9] T. Kondo, J. Ueno. Nonlinear system identification by feedback GMDH-type neural network with architecture self-selecting function. *IEEE International Symposium on Intelligent Control* 2010:1521-1526.