

Modeling and Simulation of Soft measurement Based on Improved BP Neural Network

Xinchen Cui^{a*}, Zhenlin Chen^b

Department of Engineer Weapon Science and Technology, Naval Aeronautical and Astronautical University, Yantai 264001, P.R.China

^acuixinchen@sina.com, ^bczlzhzy@163.com

Keywords: Back Propagation (BP), Principal Component Analysis (PCA), Genetic Algorithms (GA), Soft measurement.

Abstract. To solve the problem of Back Propagation (BP) neural network easy to get in local least value and the initial weight is chosen randomly, Principal Component Analysis (PCA) and Genetic Algorithms (GA) were introduced to the BP Neural Network to achieve their complementary advantages. Based on the BP neural Network a GA-BP neural Network improved network based on PCA is built and practically applied. The simulation results show that, the improved network could improve the generalization ability of the model and the ability to predict dynamic measurement data, which make the BP neural Network can be used even more widely.

Introduction

Soft measurement is essentially a problem about modeling, which describe the relationship between process variables easy to measure (instrumental variables) and process variables difficult to measure (dominant variables) through mathematical modeling, apply process variables measurement by variable mathematical calculation and estimation methods.

Recently, neural network is used more and more in soft measurement modeling [1]. Compare to traditional method, information processing could reduce data analysis and model works through artificial neural network. Mathematical model based on neural network is built through large number of samples training, which is adapt to special problem but also promoted to a class of problems. At present, among so many neural network modeling, Back Propagation (BP) neural network is the best delegate[2], which is not only the core part of the forward type network, but also reflect the artificial neural networks essence.

In this article, BP neural network is improved for its defects. The main idea is that, to introduce Principal Component Analysis (PCA) and Genetic Algorithms (GA) to the BP neural network [2]. The improved BP neural network is used in soft measurement modeling works, and simulation realized in MATLAB.

GA-BP improved neural network based on PCA

GA (Genetic Algorithm) could optimize the network initial weights, but can not input the sample quality. However, the network convergence speed increscent and the generalization ability need input sample optimized urgently. Using least variables to reflect rich enough information is expected. So based on BA-BP improved network, PCA in introduced into the network.

GA-BP optimization algorithm. GA (Genetic Algorithm) have the characteristic of global search using GA optimize the connection weight and the network structure could overcome the network oscillation brought by the network structure determination [3], the problem of easy to get in local solution could improve the generalization ability. The algorithm is consists of six steps.

- (1) Determine the network structure.
- (2) Generate initial population.
- (3) Input training sample.
- (4) Get individual error, calculate the fitness value.

(5) Determine whether meet the requirements, yes, stop and go to (6); no, go on training optimize the network, generate new threshold value or weight.

(6) Get GA optimized network.

As can be seen from the algorithm process, the aim of the GA optimized BP neural network is to determine the threshold value or weight, while during training threshold value or weight are adjusted constantly, until total error function meet the conditions, the training stop.

Improved algorithm process. The GA-BP optimized algorithm based on PCA is consists of three steps, firstly, principal component analysis, secondly, genetic algorithms, lastly, using optimized input variable and initial weight to network training [4]. The specific process as follows.

(1) Do initial sample preprocessing and calculate contribution and total contribution.

(2) Select point which principal component total rate having certain percentage.

(3) Initialize population, select network individual.

(4) Copy and intersect, create new individual and calculate estimate new individual.

(5) Check the judgment meet the criterion or not.

(6) No, go to (4); yes, construct three layers BP network with inspect sample.

Soft measurement modeling and simulation of using GA-BP neural network based on PCA

Variable principal component analysis. Principle component analyzed 13 values among the data get by preprocessing. Use MATLAB to realize PCA programming. Firstly, calculate the correlation array R, secondly, calculate the Eigen value of R, lastly, and analyze the contribution respectively. The main codes were showed as follows.

Cwstd.m

```
function std=cwstd(vector)
cwsum=sum(vector,1);
[a,b]=size(vector);
for i=1:a
    for j=1:b
        std(i,j)=vector(i,j)/cwsum(j);
    end
end
```

end

Cwscore.m

```
function score=cwscore(vector1,vector2);
sco=vector1*vector2;
csum=sum(sco,2);
[newcsum,i]=sort(-1*csum);
[newi,j]=sort(i);
fprintf('Score:\n')
score=[sco,csum,j]
```

GA-BP processing. Network training is as usual by the batching way. All the test data were treated as training sample. The network structure is 13-13-25-1, the contain layer transmit function is “tensing” and the output layer transmit function is “purely”.

Set the GA code accuracy as $eps = 1e-6$. On the population problem and heredity generations selecting ,too small, could not reach the expected accuracy, too large, would lead to redundancy, which add the encumbrance network training. After repeat simulation experiment, finally, set the GA initial population number as $popu = 50$ and heredity generation as $GEN = 100$. In programming, Goat tool case was called and the objective function was defined as “gabpEval”. The codes were showed as follows.

R=size(P,1);

S2=size(T,1);

S1=25;

```

S=R*S1+S1*S2+S1+S2;
aa=ones(S,1)*[-1,1];
popu=50;
initPpp=initializega(popu,aa,'gabpEval');
gen=100;
[x,endPop,bPop,trace]=ga(aa,'gabpEval',[], initPpp,[1e-6 1 1],'maxGenTerm',gen,
...'normGeomSelect',[0.09],[arithXover],[2],'nonUnifMutation',[2gen3]);

```

Simulation experiment and result

To test and verify the algorithm effectiveness, the improved BP neural network was applied to sewage treatment soft measurement modeling. The experiment result was analyzed. The simulation experiment was implemented under the conditions of 3.2GHz i3 CPU, 2GB memory and windows XP operating system, and programmed by MATLAB R2008a.

Variable principal component analysis. The PCA result get from MATLAB is showed in Table1.

Table 1 Correlation array Eigen value and contribution

SN	Eigen value	Contribution rate (%)	Accumulate contribution rate (%)
1	1.77079	13.62	13.62
2	1.6823	12.94	26.56
3	1.64171	12.63	39.19
4	1.49293	11.48	50.67
5	1.20496	9.27	59.94
6	1.16527	8.98	68.92
7	0.890511	6.85	75.77
8	0.829988	6.38	82.15
9	0.756538	5.82	87.97
10	0.68029	5.22	93.19
11	0.468085	3.60	96.79
12	0.41056	3.16	99.94
13	0.00832209	0.06	100

Generally, contribution rate is required to reach 80%~90%.As showed by Table1.1,accumulate contribution rate of Eigen value 1 to 9 reached 87.97%.So select the first 9 principle component and evaluate comprehensively. Calculate the Eigen vector of the first 9 Eigen value, get 9 principle components as follows.

$$y_1 = 0.2087x_1 + 0.2004x_2 - 0.1113x_3 - 0.3125x_4 + 0.2603x_5 + 0.0471x_6 + 0.4481x_7 - 0.7326x_8 + 0.2275x_9 + 0.0303x_{10} + 0.4477x_{11} - 0.7109x_{12} + 0.1051x_{13}$$

$$y_2 = 0.2591x_1 + 0.2637x_2 - 0.6644x_3 + 0.3230x_4 + 0.1351x_5 + 0.4379x_6 - 0.6103x_7 + 0.0785x_8 + 0.1121x_9 + 0.4159x_{10} - 0.2487x_{11} - 0.3467x_{12} + 0.2090x_{13}$$

$$y_3 = 0.7055x_1 - 0.3639x_2 + 0.2451x_3 - 0.4075x_4 + 0.4098x_5 - 0.1263x_6 - 0.1271x_7 - 0.1607x_8 + 0.0873x_9 + 0.4893x_{10} - 0.3656x_{11} + 0.1787x_{12} - 0.3833x_{13}$$

$$y_4 = 0.0621x_1 - 0.4886x_2 - 0.1198x_3 + 0.2324x_4 + 0.4127x_5 - 0.5479x_6 - 0.0710x_7 + 0.0629x_8 + 0.5921x_9 - 0.3213x_{10} - 0.1723x_{11} - 0.0618x_{12} + 0.4638x_{13}$$

$$y_5 = -0.3104x_1 + 0.2687x_2 + 0.2584x_3 - 0.4664x_4 + 0.0825x_5 + 0.3429x_6 - 0.3173x_7 - 0.2850x_8 + 0.4694x_9 - 0.1916x_{10} - 0.3409x_{11} + 0.2409x_{12} + 0.1204x_{13}$$

$$y_6 = 0.0684x_1 + 0.3226x_2 - 0.3580x_3 - 0.1946x_4 + 0.5288x_5 - 0.0780x_6 - 0.1121x_7 + 0.2968x_8 + 0.0163x_9 - 0.4091x_{10} + 0.3223x_{11} + 0.1896x_{12} - 0.4435x_{13}$$

$$y_7 = 0.1089x_1 + 0.2840x_2 + 0.3563x_3 + 0.3752x_4 + 0.3065x_5 - 0.0160x_6 - 0.0646x_7 - 0.1877x_8 - 0.3975x_9 - 0.3463x_{10} - 0.2977x_{11} - 0.1672x_{12} + 0.0473x_{13}$$

$$y_8 = 0.2408x_1 + 0.2858x_2 - 0.0326x_3 - 0.2072x_4 - 0.3888x_5 - 0.2860x_6 - 0.1294x_7 - 0.2163x_8 + 0.0295x_9 + 0.0295x_{10} + 0.2294x_{11} + 0.2538x_{12} + 0.4490x_{13}$$

$$y_9 = 0.0677x_1 - 0.2614x_2 - 0.2476x_3 - 0.2298x_4 + 0.1563x_5 + 0.3790x_6 + 0.0459x_7 + 0.2004x_8 - 0.3360x_9 + 0.2041x_{10} - 0.0974x_{11} + 0.1601x_{12} + 0.3808x_{13}$$

Therefore, take place the previous 13 variables with the 9 principles component as the input variables of the network.

GA-BP simulation result. Fitness and sum of square error curve during the training are showed in Fig.1 and Fig.2 After 70 times iteration, the average adaption degree of chromosome get in stabilization, after 80 generations iteration get an optimum initial weight.

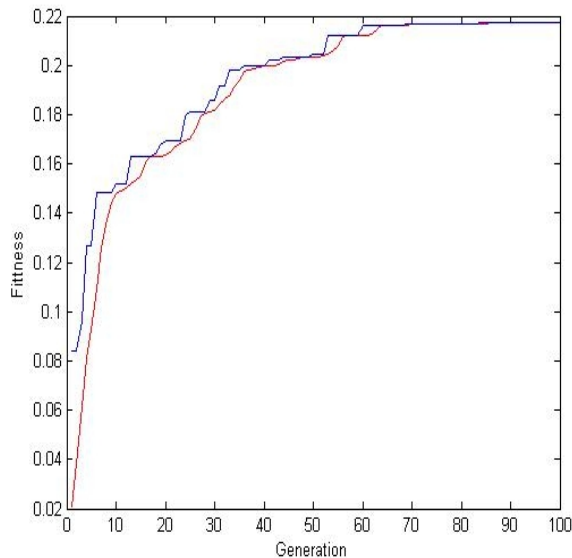


Fig.1 Fitness curve during training

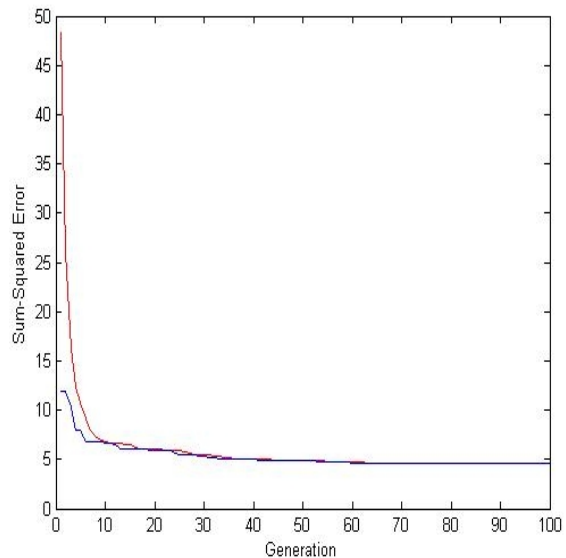


Fig.2 Sum of squared error cure during training

Observations and discussion. As showed in Fig.1, the network overcome the problem of local minimum, the curve of error has no horizontal part, change quickly.

As showed in Fig.2, the forecast output was close to expected output, the combine result was preferably, the network have better data combine ability. The GA-BP network forecast value relative error between the interval of $-0.015 \sim 0.02$, which improve the accuracy of forecast effectively.

The GA-BP neural forecast result is better than normal BP neural network, which prove that the GA-BP neural network has advantageous in forecast dynamic measurement data.

Conclusions

In this paper we undertook a systematic study of BP neural network, improved the network with PCA-GA, designed the algorithm process and applied the improved network to waste water conduction soft measurement modeling and simulation. The simulation experiment results show that the improved network has advantages in forecast of dynamic measurement data, has better generalization capability and could be applied to other soft measurement modeling.

References

- [1] Cho, Sung-bare, Fusion of neural networks with fuzzy logic and genetic algorithm, Integrated Computer-Aided Engineering, Volume 9, Number 4, 2002:363-372
- [2] Merigeault S, Betrayer M, Papillion J, Data fusion based on neural network for the mobile subscriber location, Vehicular Technology Conference Fall 2010. IEEE VTS Fall VTC2010. 52nd Vehicular Technology Conference, Volume 9, 2010:536-541
- [3] Muhammad J, Husain A, nektonic A, Magill E, New neural network based mobile location estimation in a metropolitan area, Artificial Neural Networks-ICANN 2005, PART 2, Proceedings Lecture Notes IN Computer Science 3697:935-941
- [4] X. Z. Gao, S. J. Osaka and A. V. Vasilakos, A modified Elman neural network-based power controller in mobile communications system, Soft Computing-A Fusion of Foundations, Volume 9, Number 2, February 2005:88-93