ARTIFICIAL NEURAL NETWORKS FOR SHAPE MODELING OF SEA BOTTOM

JERZY BALICKI

Naval University of Gdynia
Śmidowicza 69 81-103 Gdynia, Poland
J.Balicki@amw.gdynia.pl

Artificial neural networks are applied for approximation and interpolation of function with multiple variables. Because of concurrent processing of data by neurons, fhis approach can be seen as promising alternate for standard algorithms. From these reasons, the analysis of capabilities for some models of neural networks has been carried out in the purpose for modeling the shape of sea bottom. The feed-forward multi-layer networks with different transfer functions have been tested. These networks have been trained by backpropagation algorithm and its versions with some improvements. Moreover, the gradient optimization technique by Levenberg-Marquardt has been applied. Finally, for determination of the depth in a point of the water area the two-layer network with the hidden layer of the radial neurons has been proposed.

INTRODUCTION

Modeling the contour of the sea bed is related to construction an algorithm of finding the depth of the bottom according to the coordinates of the given point. However, there are two important questions: the first one is the technique of acquisition data and the second ne is the selection of the algorithm for sample interpolation.

Artificial neural networks (ANN) are applied for approximation and interpolation of function with multiple variables. Currently, this term tends to refer mostly to neural models employed in statistics, cognitive psychology and artificial intelligence. Neural models designed with emulation of the central nervous system in mind are a subject of theoretical neuroscience.

In modern software implementation of ANNs the approach inspired by biology has more or less been abandoned for a more practical approach based on statistics and signal processing. In some of these systems neural networks, or parts of neural networks are used as components in larger systems that combine both adaptive and non-adaptive elements. While the more general approach of such adaptive systems is more suitable for real-world problem solving, it has far less to do with the traditional artificial intelligence connectionist models.

What they do however have in common is the principle of non-linear, distributed, parallel and local processing and adaptation [3].

In this paper, the analysis of capabilities for some models of neural networks has been carried out in the purpose for modeling the shape of sea bottom. The feed-forward multi-layer networks with different transfer functions have been tested. These networks have been trained by back-propagation algorithm and its versions with some improvements.

Moreover, the gradient optimization algorithm by Levenberg-Marquardt (LMA) has been applied. It provides a numerical solution to the mathematical problem of minimizing a function, generally nonlinear, over a space of parameters of the function. This minimization problem arises especially in least squares curve fitting. LMA interpolates between the Gauss-Newton algorithm and the method of gradient descent. The LMA is more *robust* than the Gauss-Newton algorithm, which means that in many cases it finds a solution even if it starts very far off the final minimum [2]. On the other hand, for well-behaved functions and reasonable starting parameters, the LMA tends to be a bit slower than the Gauss-Newton algorithm.

Finally, for determination of the depth in a point of the water area the two-layer network with the hidden layer of the radial neurons has been proposed [4]. Radial basis functions (RBF) are powerful techniques for interpolation in multidimensional space. A RBF is a function which has built into a distance criterion with respect to a centre. Radial basis functions have been applied in the area of neural networks where they may be used as a replacement for the sigmoidal hidden layer transfer characteristic in multi-layer perceptrons. RBF networks have two layers of processing: In the first, input is mapped onto each RBF in the 'hidden' layer. The RBF chosen is usually a Gaussian. In regression problems, the output layer is then a linear combination of hidden layer values representing mean predicted output. The interpretation of this output layer value is the same as a regression model in statistics. In classification problems, the output layer is typically a sigmoid function of a linear combination of hidden layer values, representing a posterior probability. Performance in both cases is often improved by shrinkage techniques, known as ridge regression in classical statistics and known to correspond to a prior belief in small parameter values (and therefore smooth output functions) in a Bayesian framework.

1. MESEURMENTS OF LOCATIONS AND DEPTH

Let us assume that for the given water region the measurements have been performer and coordinates (ϕ,λ) from the Global Position Systems GPS as well and $d(\phi,\lambda)$ from the synchronized sonar. Echo sounding is the technique of using sound pulses directed from the surface or from a submarine vertically down to measure the distance to the bottom by means of sound waves. Distance is measured by multiplying half the time from the signal's outgoing pulse to its return by the speed of sound in the water. Echo sounding is effectively a special purpose application of sonar used to locate the bottom.

An echo-sounder sends an acoustic pulse directly downwards to the seabed and records the returned echo. The sound pulse is generated by a transducer that emits an acoustic pulse and then “listens” for the return signal. The time for the signal to return is recorded and converted to a depth measurement by calculating the speed of sound in water. As the speed of sound in water is around 1 500 meters/second, the time interval, measured in milliseconds, between the pulse being transmitted and the echo being received, allows bottom depth and targets to be measured. The value of underwater acoustics to the fishing industry has led to the development of other acoustic instruments that operate in a similar fashion to echo-

sounders but, because their function is slightly different from the initial model of the echo-sounder, have been given different terms.

Measurements are supposed to be carried out by uniform way for the sea area. The density of sampling should be greater for the sub-regions outstandingly important or for areas that are explored the first time. Data of depth are gathering along profiles because of possibilities of the rejection of fault data [1]. We reject data that exceed beyond admissible values and also data with too large changes of depth. Measurements are averaged because of avoiding the sequences of errors. Interferences can be filtered by artificial neural networks, too [5]. An effectiveness of multi-layer neural models is very high according to errors filtering and may be even equal to 40% [7]. The problem is how to interpolate given points of depths into the proper function of depth for variables ϕ and λ.

In the studied multi-layer neural networks, there are two input neurons and one output neuron with linear transfer function. Moreover, one or two hidden layer of neurons can be used for interpolation. Hidden neurons have sigmoid transfer functions or radial transfer functions. The training techniques for hidden neurons with sigmoid activation functions based on several back-propagation algorithms have been tested. We studied standard back-propagation algorithm, back-propagation algorithm with momentum, and back-propagation algorithm with adaptive rate of learning. Furthermore, the Levenberg-Marquardt algorithm has been applied for training the "depth map" network. The training techniques for hidden neurons with radial activation functions based on the orthogonal least square algorithms [8]. The training set of patterns, preparing input data, training of network, and evaluation of the quality of networks are crucial steps according to methodological approach. Software has been implemented in Matlab language, and numerical experiments have been developed by PC with processor Core2Duo/3 GHz.

If the training set has too many patterns, then the learning time can be too long, as well. On the other hand, sampling of patterns is supposed to be representative for all sub-regions. If there is a rapid whole in the seabed and there in no pattern of it, then model of network is not capable to recognize this whole. It is important to carry out the measurement for different wheatear conditions as well as different perturbations. So, we suggest to consider as much as possible patterns in the learning set for the given time limit of training.

2. TRAINING OF NEURAL MODEL FOR THE MAP

Capability of network over-fitting to the training set is related to the size of that network. The over-fitted network has difficulties with generalization of the knowledge for untrained positions.

Figure 1 shows the process of training the multilayer network by the standard back-propagation algorithm. We can observe the minimization of the root from the sum of squares for errors calculated at the outputs of ANN. There were 100 elements in the training set. Each triple from a set was randomly chosen and new values of synaptic weights and biases were calculated for the epoch. The global error was reduced to 0.471 after 1000 epochs. However, this results for 100 elements is not very promising because it should be less than 0.02 for the relative error 1%. There were two hidden layers with 5 and 10 neurons, respectively. Although, the numbers of neurons were increased, no improvement can be done. Similar outcomes have been obtained by the model with the hidden layer.

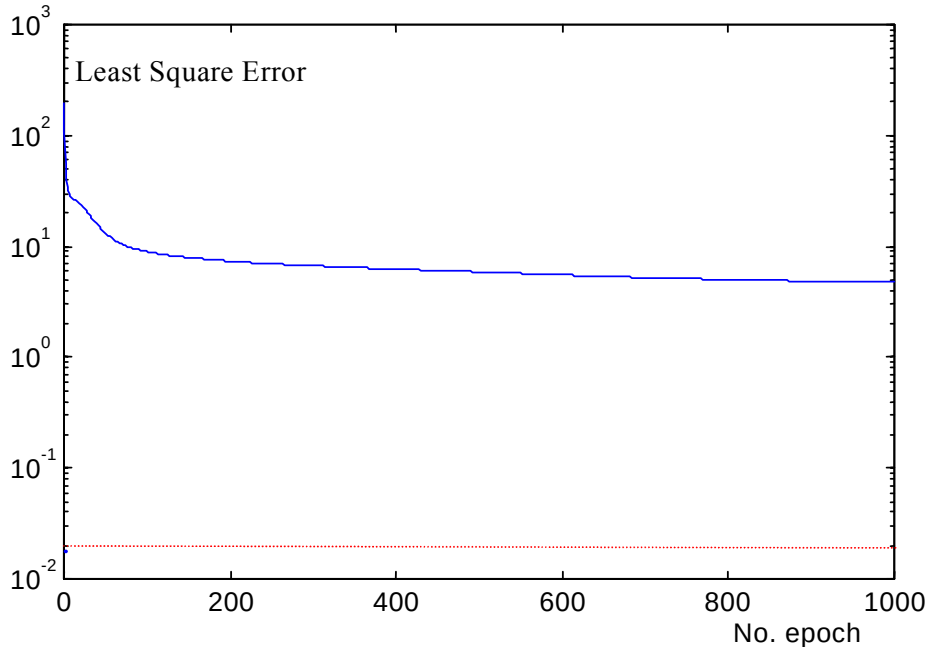

Fig.1 Minimization of the least square terror by back-propagation algorithm

In that case, arithmetic error was equal to 8.05% and maximal error - 41.2%. An assigning more time did not cause the improvement In the reasonable rate. So, we studied back-propagation algorithm with momentum and back-propagation algorithm with adaptive rate of learning [10].

Biases and weights from two inputs to *m* hidden neurons are decision variables. So, there are 3*m*+2 variables. To avoid the network over-fitting, we assume the number of depth patterns should be greater than 3*m*+2. The network can avoid random errors, but it has difficulties with systematic errors fro echo-sounder and GPS.

Preparing input data is important for the quality of finding the depth of seabed by neural model. Moreover, the rate of learning can be accelerated. Networks that were trained with the normalized data to the period (-1, 1) learned faster and gave better outcomes. Input data can be limited, as follows:

$$\phi^{min} \le \phi \le \phi^{max}, \tag{1}$$

$$\lambda^{min} \le \phi \le \lambda^{max}, \tag{2}$$

$$d^{max} \le d \le 0, \tag{3}$$

where
ϕ^{min}, ϕ^{max}, λ^{min}, λ^{max} – constraints for coordinates of position,
d^{max} – maximal depth of water region.

Value of the coordinate ϕ is normalized to the value $\bar{\phi}$ that is input data to the network, as follows:

$$\bar{\phi} = \bar{\phi}_{\min} + \frac{\bar{\phi}_{\max} - \bar{\phi}_{\min}}{\phi_{\max} - \phi_{\min}} (\phi_{\max} - \phi_{\min}), \tag{4}$$

where

$\bar{\phi}_{\min}, \bar{\phi}_{\max}$ - limits of normalization.

Similarly, we can detrmined $\bar{\lambda}$ and $\bar{d}$. Limits of normalization–1 and 1 are commonly used [9]. However, they can developed for network to values –0,9 and 0,9 because for that period of values sigmoid transfer function there is intense increase functional value with the relatively small value of variable increment.

3. ALGORITHMS FOR TRAINING OF NEURAL NETWORKS

The training set can include 100, 400 or 900 elements (ϕ,λ,d). Back-propagation algorithm with momentum, back-propagation algorithm with adaptive rate of learning and Levenberg-Marquardt are improved method for training artificial neural network [12].

In the standard back-propagation algorithm, the rate of training is constant and the best results have been obtained for 0.0001. On the other hand, a learning rate is increased after an epoch for the decrease of the last square error. If the error is not decreased or the decreasing is very small, then the rate factor is decreased because value of error can be in the neighborhood of the local minimum of error function. Figure 2 shows the process of training network by back-propagation algorithm with adaptive rate of learning.

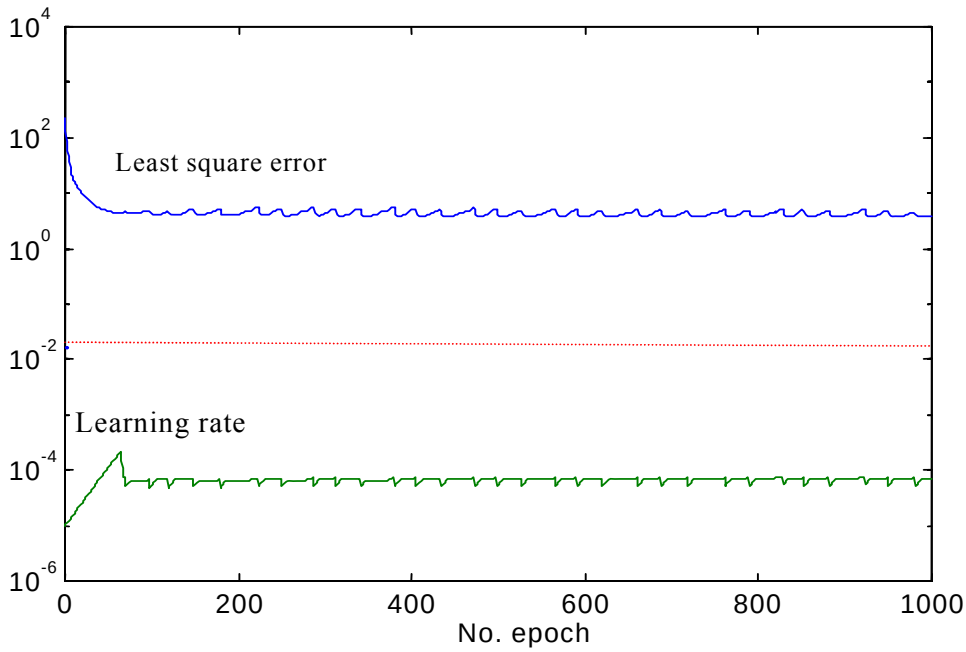

Fig.2 Training by back-propagation algorithm with adaptive learning rate

The least square error was decreased from 4.71% to 3.85% and the average error from 8.05% to 6.08%. However the maximal error increased from 41.2% to 42.8%. Back-propagation algorithm with momentum, that can omit local minimum of error function, decreased the maximal error to 39.4%.

The Levenberg-Marquardt algorithm is from 10 till 100 times faster than back-propagation ones. Figure 3 shows outcomes obtained during training by this algorithm. After 1000 epochs the least square error was 0.15% and the average error 1.18%. However the maximal error was equal to 10.2% that was unacceptable.

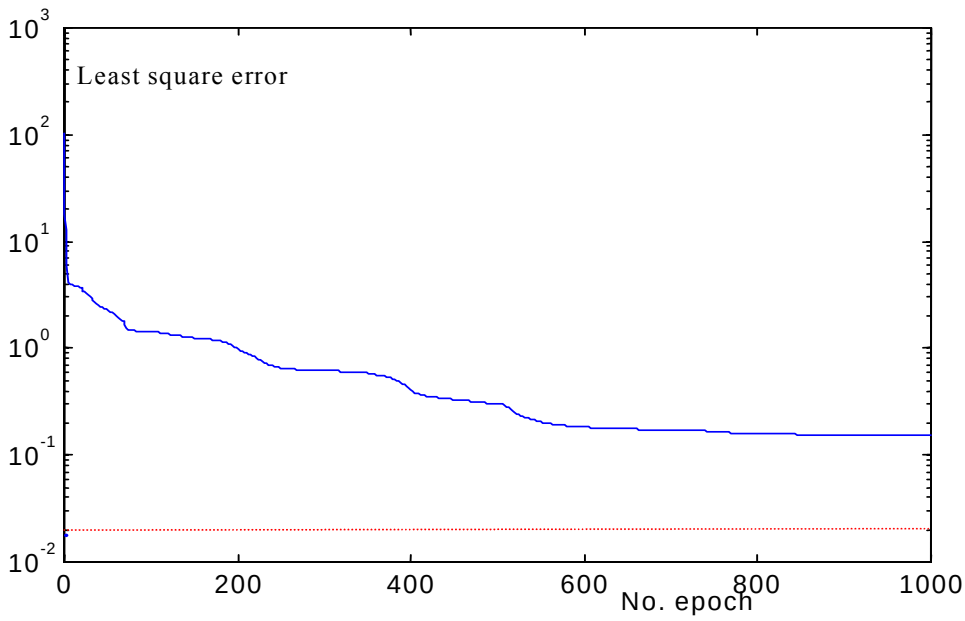

Fig.3 Minimization of the least square terror by LMA

4. ALGORITHMS FOR TRAINING OF NEURAL NETWORKS

A RBF is a function which has built into a distance criterion with respect to a centre. RBF networks have two layers of processing: In the first, input is mapped onto each RBF in the 'hidden' layer. The RBF chosen is usually a Gaussian (Figure 4). In classification problems, the output layer is typically a sigmoid function of a linear combination of hidden layer values, representing a posterior probability. Performance in both cases is often improved by shrinkage techniques, known as ridge regression in classical statistics and known to correspond to a prior belief in small parameter values (and therefore smooth output functions) in a Bayesian framework [11].

Standard multilayer networks develop the stochastic approximation function of two variables transferring the set of input variables $X=\{(\varphi,\lambda)\in R^2 \mid \varphi_{\min}\le\varphi\le\varphi_{\max}, \lambda_{\min}\le\lambda\le\lambda_{\max}\}$ into the set of depth. In the linear or sigmoid neurons, the level of activation u_m is calculated, as follows:

$$u_m = b_m + \sum_{r=1}^{R} p_r w_{rm}, \quad m=\overline{1,M}, \tag{6}$$

where

b_m – value of the mth bias,

p_r – value of the rth input,

w_{rm} – value of the synaptic weight from rth input to the mth neuron of the first layer,

R – number of inputs,

M – number of neurons in the first layer.

In the radial neurons of hidden layer, u_m is multiplication bias by the distance between the vector of inputs p and vector of weights w calculated, as follows:

$$u_m = b_m \sqrt{\sum_{r=1}^{R} (p_r - w_{rm})^2}, \quad m=\overline{1,M}. \tag{7}$$

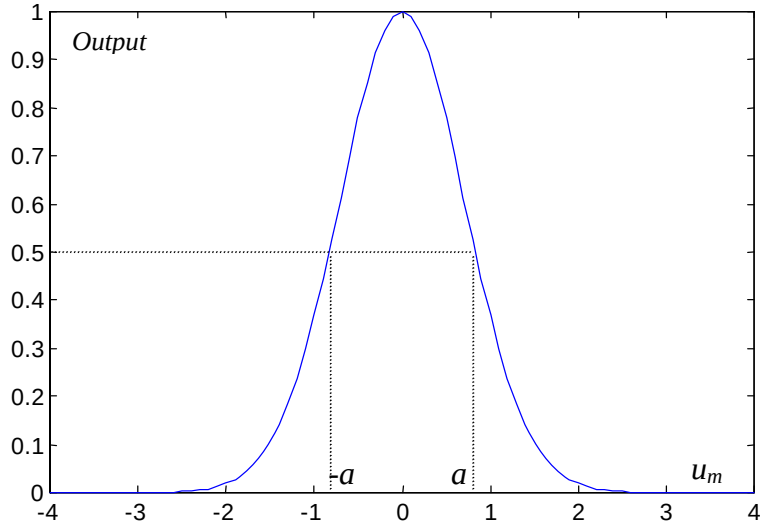

Fig.4 A radial function

If the distance between *p* and *w* decreases, then the output from this neuron increases until it obtains value 1 for the *p*=*w*. To some extend, a radial neuron is the similarity measure between *p* and *w*. A bias scales that distance. If this distance is equal to a/b_m, then output is 0.5. Radial neurons from the hidden layer are connected with the linear output neuron. That approach was the most effective and after 93 epochs the least square error was 10^{-6} %.

5. CONCLUDING REMARKS

Radial networks of networks support modeling of the shape of the seabed with the reasonable precision. Inputs of network are coordinates of position and an output is the depth of water. Time of the depth determination is equal time of few instruction run. There is possibility to present different cut of the depth map.

Patterns to the training set are selected randomly from the measured data carried out during preferred weather conditions. The set of training should not exceed 1000 elements for a sub-region. The plane of water region is supposed to be divided on set of rectangles with the similar size. Common areas with the size 10% of sub-region are required between two neighbors sub-regions.

The analysis of capabilities for some models of neural networks has been carried out in the purpose for modeling the shape of sea bottom. The feed-forward multi-layer networks with different transfer functions have been tested. These networks have been trained by back-propagation algorithm and its versions with some improvements. The Levenberg-Marquardt algorithm is from 10 till 100 times faster than back-propagation ones.

If the distance between input vector and weights decreases in the radial network, then the output from this neuron increases until it obtains value 1. A radial neuron is the similarity measure between these vectors. A bias scales that distance. Radial neurons from the hidden layer are connected with the linear output neuron.

REFERENCES

[1] J.Balicki, Z.Kitowski, A.Stateczny, Neural techniques for linear multiobjective optimization problems with continuous variables, Proceedings of the Third Conference „Neural Networks and Their Applications", Kule, 1997, pp. 505-510.

[2] S. Chen, C.F. Cowman, P.M. Grant, Orthogonal least squares learning algorithm for radial basis function learning, IEEE Trans. Neural Networks, Vol. 2, 1991, pp. 302-309.

[3] E.K.P. Chong, S.H. Zak, An introduction to optimization, John Wiley&Sons, Inc., New York, 1996.

[4] S. Coombes, S. H. Doole, Neuronal Population Dynamics with Post Inhibitory Rebound: A Reduction to Piecewise Linear Discontinuous Circle Maps, Dynamics and Stability of Systems, Vol. 11, 1996, pp. 193-217.

[5] H. Demuth, M. Beale, Matlab Neural Network Toolbox. User's Guide, The MathWorks, Inc., 2003.

[6] M. Frize, C. Ennett, M. Stevenson, H. Trigg, Clinical decision-support systems for intensive care units using artificial neural networks, Medical Engineering and Physics, 2001, Vol. 23, pp. 217-225

[7] J. Hertz, A. Krogh, R. Palmer, Introduction to the Theory of Neural Computation, Massachusetts: Addison-Wesley Publishing Company, Inc. Reading, 1991.

[8] T. Kujanpää, J. Roos, Efficient Initialization of Artificial Neural Network Weights for Electrical Component Models, Book of Abstracts of SCEE 2006, Sinaia, Romania, 2006, pp. 47-48.

[9] M. Mokhari, Classification by the Random Neural Network, Tenth International Conference on Mathematical and Computer Modelling and Scientific Computing, Boston, USA, 1995.

[10] S. Priddy, P. E. Keller, Artificial Neural Networks: An Introduction, SPIE Press, New York 2005.

[11] A.J.C. Sharkey, G.O. Chandroth, N.E. Sharkey, Acoustic Emisssion, Cylinder Pressure and Vibration: A Multisensor Approach to Robust Fault Diagnosis. In Proceedings of IJCNN2000, Como, Italy, 2000.

[12] L. M. Silva, L. A. Alexandre, J. Marques de Sá, New developments of the z-edm algorithm, In 6th International Conference on Intelligent Systems Design and Applications (ISDA'06), volume 1, pages 1067-1072, Jinan, China, October 2006