

PREDICTION OF THE INTEGRATED INDICATOR
OF QUALITY OF A NEW OBJECT UNDER THE CONDITIONS
OF MULTICOLLINEARITY OF REFERENCE DATA*S.B. Akhlyustin¹, A.V. Melnikov¹, R.A. Zhilin¹*¹Institute of the Ministry of Internal Affairs, Voronezh, Russian Federation

E-mails: cerg7676@yandex.ru, meln78@mail.ru, zhilin@yandex.ru

Prediction of a new object state at a lack of the known characteristics and estimates of quality indicators of a number of studied objects (a set of reference data) often leads to the problem of multicollinearity of basic data. We propose the following three ways to overcome this problem relating to the sphere of data mining: use a ridge regression, train with the teacher a two-layer neural network, consecutive adapt a single-layer neural network. Also, we compare characteristics of the proposed ways. In the ridge regression method, the introduction of a regularizing term into the LMS equation gives an approximate solution with a sufficient degree of accuracy. A disadvantage of use of the two-layer neural network “feed-forward backprop” and the procedure of training with the teacher “train” is that adjusted weights of the neural network take chaotic (and even negative) values that contradicts a common practice of examination. The following features are revealed: considerable dispersion of weights and shifts of a neural network, ambiguity of the solution due to the choice of random initial conditions, strong dependence on a training algorithm. In order to overcome this shortcoming, we propose a transition to consecutive adaptation of a single-layer neural network with fixing shifts of neurons at zero level.

Keywords: examination of objects; prediction; multiple regression; ridge regression; regularization; neural network; training of model; adaptation.

Introduction

The problem of overcoming multicollinearity of input data is one of the main problems of modern calculus mathematics. The multicollinearity is a consequence of existence of strong correlation, even functional, communications of vectors of signs (factors) of various compared objects. The simplest methods to eliminate the multicollinearity (reduction of amount of signs, addition of low significant signs, nonlinear transformation of the equations of regression) are artificial and, as a rule, do not yield positive results. In this work, we propose three regular methods of overcoming multicollinearity based on the theory of *data mining* and compare them with each other [1].

During examination of difficult technical objects (systems, devices, samples), need of approximate quality assessment of a new object often evolves from the analysis of reference data, i.e. a set of the known characteristics X and estimates of the quality indicator J of similar objects. This problem is complicated by the fact that the technique to obtain such estimates is, as a rule, unknown.

Usually, for joint examination of several technical objects, the theory of multiple linear regression is used [2,3]. Then for k objects (samples) having characteristics (signs), linear equations for each of them can be written in a matrix form.

The mathematical model corresponds to the model of a linear multiple regression, which we write in the compact matrix form

$$J = XV + E, \tag{1}$$

where J is a vector of indicators of the order $k \times 1$, X is a matrix of independent variables (predictors) of the order $k \times m$, V is a vector of regression coefficients of the order $k \times 1$, E is a matrix of the remains (regression noise) of the order $k \times 1$.

In our case, it is necessary to solve the following problem: for the given integrated indicators of the known objects (samples) J and the matrix X for several known samples, to define a set of weight coefficients to be regression coefficients, and then to use this set for assessment of the indicators of new objects $\{V_i\}$, i.e. to carry out their prediction. In econometrics, such a procedure is called *multidimensional calibration* [2, 4].

As a rule, in problems of examination, there are data (X, J) for a number of the known objects called *reference data*. It is required to use these data in order to make training of regression model, i.e. to find unknown weight coefficients \hat{V}_i by the method of the least mean squares (LMS) [2, 5]:

$$V = (X^T X)^{-1} X^T J. \tag{2}$$

However, actually often, we do not manage to calculate already the first factor of expression (2), i.e. to find the inverse matrix of $X^T X$, since the determinant of this matrix is equal to zero.

In this situation, we observe the phenomenon of multicollinearity of reference data which reflects strong correlation (or even functional) communication between the matrix X vectors (such a matrix is called badly caused). The strict multicollinearity breaks one of the main conditions of the Gauss–Markov theorem (about a matrix rank), and makes it impossible to construct a regression.

The following factors are among the main reasons for multicollinearity [2, 5].

1. The quantity of samples k (rows of the matrix X) insufficiently also approaches the amount of signs m (a saturation status).
2. Two or more linearly dependent variables are wrongly included in the equation.
3. Two or more explanatory variables, which are poorly correlated in a normal situation, become strongly correlated under specific conditions of selection.
4. The model includes the variable which is strongly correlating with a dependent variable (such independent variable is called *dominant*).

Further, to overcome the computing problems connected with multicollinearity, we use the data mining methods [6]: theory of multiple linear regression, method of the analysis of hierarchies, method of ridge regression, methods of the theory of neural networks.

1. Forming of Reference Data

For assessment of technical objects (in the problems of multidimensional calibration called *samples*), the works [1, 5, 7] propose the complex quality price indicator, which we rewrite in the modified form as follows:

$$J_{empl} = (V_{qw} \hat{J} + V_{pr} \hat{P}) / (V_{qw} + V_{pr}), \tag{3}$$

where J, P are the integrated indicators of quality and price of a sample; V_{qw}, V_{pr} are the weighting coefficients of quality and price that allow to make a choice about the

required ratio “quality-price”; the character “^” over designation of a variable means a normalization.

The additive integrated indicator of quality has the form

$$\hat{J} = \frac{1}{\sum V_{gr}} \left[V_{gr,1} \frac{\sum_j V_{j,q} \hat{x}_j}{\sum_j V_{j,q}} + V_{gr,2} \frac{\sum_i V_{i,ex} \hat{x}_i}{\sum_j V_{j,ex}} + V_{gr,3} \frac{\sum_l V_{l,ql} \hat{x}_l}{\sum_l V_{l,ql}} \right], \quad (4)$$

where $\hat{x}_j, \hat{x}_i, \hat{x}_l$ are rated signs of quantity, existence and quality, respectively; $V_{j,q}, V_{i,q}, V_{l,q}$ are partial weight coefficients in each group of signs; $V_{gr,1}, V_{gr,2}, V_{gr,3}$ are weight coefficients of groups of signs; $\sum V_{gr}$ is a sum of group coefficients.

In order to provide a uniform contribution of various composed terms in weighed sum (4), it is necessary to reduce their values to uniform range [7]. Distinguish the positive effect signs (PES) and negative effect signs (NES) depending on the impact of growth of sign value on the size of the integrated indicator J . For these two groups of signs, the normalization is carried out differently.

For quantitative and qualitative signs of PES, we consider the following normalization:

$$\hat{x}_j = \frac{x_j}{x_{j,base}}, \quad j = 1, 2, \dots, m, \quad (5)$$

where $x_{j,base}$ is a basic (maximum) value of the sign for all k objects

$$x_{j,base} = \max_l x_j^{(l)}, \quad l = 1, 2, \dots, k, \quad \hat{x}_j \in [0; 1] \quad \forall j. \quad (6)$$

Second composed term of formula (4) gives only relative size of observation of *existence signs*.

For NES, as a basic object for comparison with other objects of examination, we choose an object with the minimum value of sign in the group of the compared objects

$$x_{j,base} = \min_l x_j^{(l)}, \quad l = 1, 2, \dots, k. \quad (7)$$

Then, excess of the j -th sign value of NES for the l -th object over basic (minimum) value is characterized by the rated sign

$$\hat{x}_j = \frac{x_{j,base}}{x_j}. \quad (8)$$

For the return price, the formula takes the following form:

$$\hat{P}^{(l)} = \frac{P_{base}}{P^{(l)}}, \quad l = 1, 2, \dots, k. \quad (9)$$

As the considered technical objects, we choose the integrated security systems (ISS), which technical characteristics are described in the works [8, 9] in detail.

Let us select signs of ISS in decreasing order of their importance: x_1 is the maximum length of the communication line, (m); x_2 is the number of the events which are stored in memory; x_3 is a response time on violations, (ms); x_4 is existence of the special controller for work with address annunciators; x_5 is integration with the system of security television; x_6 is the scale of an object to which ISS can be applied. Let us find the rated values of signs for 3 samples (Table 1).

Table 1

Rated values of signs of examination objects

ISS model	“Boundary-08”	“Kodos-A20”	“C2000M”	“AST-I”
ϵ	1	2	3	4
\hat{x}_1	0,400	0,533	1,000	0,900
\hat{x}_2	0,693	1,000	0,267	0,240
\hat{x}_3	1,000	0,571	0,800	0,800
$\sum V \hat{x}_i$	0,598	0,667	0,789	0,715
\hat{x}_4	1,000	0,000	1,000	1,000
$\sum V \hat{x}_j$	1,000	0,000	1,000	1,000
\hat{x}_5	1,000	1,000	0,800	0,400
\hat{x}_6	1,000	1,000	0,700	0,600
$\sum V \hat{x}_l$	1,000	1,000	0,700	0,460
\hat{J}	0,705	0,683	0,701	test
\hat{P}	0,545	0,960	1,000	0,750
\hat{x}_{empl}	0,651	0,775	0,801	test

For quantitative signs, we define the vector V_q using method of the analysis of hierarchies [10, 11]. We choose the sign x_1 as the basic (general). The matrix of pair comparisons takes the form

$$W = \begin{pmatrix} 1 & 3 & 5 \\ 0,33 & 1 & 3 \\ 0,2 & 0,33 & 1 \end{pmatrix}, \quad (10)$$

for which, taking into account the equation

$$WV = \lambda V, \quad (11)$$

we obtain the greatest eigenvalue $\lambda_{\max} = 0,033$ and the index of coherence $IS = 0,017$.

For the defined matrix of eigenvectors for matrix (10), we find the vector of priorities

$$V_q = (0,917 \quad 0,371 \quad 0,151), \quad (12)$$

and, after a normalization division by $\sum V_i = 1,439$, we obtain

$$\hat{V} = (0,637 \quad 0,258 \quad 0,105)^T. \quad (13)$$

Taking into account (13) and private values of signs (see Table 1), we determine the weighed sums of signs (see Table 1).

Let us calculate further integrated indicators J, J_{empl} , take

$$\hat{V} = (0,637 \quad 0,105 \quad 0,258)^T, \quad (14)$$

and construct the expanded vector of weights

$$\hat{V}_{empl} = (0,637 \quad 0,105 \quad 0,258 \quad 0,500)^T, \quad (15)$$

for the case $V_{qw}/V_{pr} = 2$, i.e. when choosing a ratio between quality and the cost to be 1 : 0,5.

Then

$$\hat{J} = (0,705 \quad 0,683 \quad 0,707)^T, \quad (16)$$

the vector of cost sign

$$\hat{P} = (0,545 \quad 0,960 \quad 1,000)^T. \quad (17)$$

The vector of the complex indicator “quality-price” is calculated by formula (4) and is equal

$$J_{empl} = (0,651 \quad 0,775 \quad 0,871)^T. \quad (18)$$

As before, we tabulate the obtained values of integrated indicators $\hat{J}, \hat{P}, \hat{J}_{empl}$.

Let us construct finally *reference data* for three ISS samples in the form

$$\hat{X} = \begin{pmatrix} 0,538 & 1,000 & 1,000 & 0,545 \\ 0,667 & 0,000 & 1,000 & 0,960 \\ 0,789 & 1,000 & 0,770 & 1,000 \end{pmatrix}; \quad J_{empl} = \begin{pmatrix} 0,652 \\ 0,775 \\ 0,871 \end{pmatrix}, \quad (19)$$

where rows and columns of the matrix X correspond to samples and signs, respectively.

2. Ridge Regression Method

As we see from (18), the quantity of samples $k = 3$ are less than the amount of signs $m = 4$ and it is not possible to use formula (2) of the LMS method for assessment of the weights vector V .

Owing to a lack of samples for training model, there exists a problem of multicollinearity for which overcoming we suggest to use one of the data mining methods that is a method of ridge regression. The method of regularization proposed by the academician A.N. Tikhonov and applied to the solution of incorrect problems [12] is the cornerstone of the procedure of ridge regression.

The theory of regularization proposed by A.N. Tikhonov assumes representation of mistakes functional in the form of a sum of two composed terms [13]:

$$E(F) = E_S(F) + E_\alpha(F).$$

In this formula, the first term $E_s(F)$ describes a mistake that is a distance between the desirable response of the model J_i and the actual output signal y_i for the chosen example of training. For instance, it is possible to define

$$E_S = \frac{1}{2} \sum_{i=1}^k (J_i - y_i)^2 = \frac{1}{2} \sum_{i=1}^k [J_i - F(x_i)]^2.$$

The second composed term (regularization augend) $E_c(F)$ depends on “geometrical” properties of the approximation function $F(x)$. As a rule, some square functional can be used in the form of this stabilizing functional $\Omega(F)$. In the last expression, existence of the second composed term stabilizes the solution in the sense of smoothness and, therefore, guaranteeing the continuity of display.

In the theory of regularization, it is required to minimize the size

$$E(F) = E_s(F) = \frac{1}{2} \sum_{i=1}^N [J_i - F(x_i)]^2 + \frac{1}{2} \alpha \Omega(F),$$

where α is a positive real number called regularization parameter.

Therefore, ridge regression assumes assessment of unknown vector of weight coefficients V_c in the calibration model by the modified formula of the LMS:

$$V_c = (X^T X + \alpha I_m)^{-1} X^T J, \tag{20}$$

where I_m is a diagonal single matrix of the order m .

Adding of the regularization parameter α solves the problem of bad conditionality of a matrix. Nevertheless, there exist no accurate rules of the choice of this parameter. Therefore, in calculations, it is required to reduce consistently α , since the value $\alpha = 0, 1$, before achievement of the best accuracy of estimates.

In order to assess an error of the proposed calibration model in the course of training, the “exact” vector of weight coefficients (16) (in the sum, the first three elements equal 1,0, and the fourth one is 0,5 that reflects the choice of the ratio “quality price” to be 1:0,5)

$$V = (0,637 \quad 0,105 \quad 0,258 \quad 0,500)^T \tag{21}$$

is compared with the vector $\hat{V}_c = (\hat{V}_{c,1}, \dots, \hat{V}_{c,m})$ estimated by formula (20), consistently reducing size. At the same time, the relative error of calibration is

$$\delta_V = \frac{\sqrt{\sum_{i=1}^m (V_{c,i} - V_i)^2}}{\|V\|_E^m}, \tag{22}$$

where $\|V\|_E^m$ is the Euclidean norm of the vector V .

Let us calculate the vector of weight coefficients V_c and the relative error of calibration δ_V for various values of the regularization multiplier α . As we see, if α takes values from 0,1 to 10^{-8} , then the relative error of calibration decreases to very small values.

At the same time, after calibration, the best approach to the vector of weight coefficients takes the form

$$V_c = (0,279 \quad 0,088 \quad 0,185 \quad 0,421)^T. \tag{23}$$

Let us go to the second stage of multidimensional calibration, i.e. assessment of the integrated indicator J of a new sample. Consider a new ISS “AST-I” (see Table 1, Column 5) as the characteristic test and choose its signs x_i to be the same as for three known samples. Find the rated values of its signs x_i according to the normalization technique given above:

$$X_{Test} = (0,715 \quad 1,000 \quad 0,460 \quad 0,750)^T. \tag{24}$$

Further, we use weight coefficients (16) – (18) and calculate “true” value of the integrated indicator $J_{empl} = 0,703$.

Table 2

Relative error of prediction of the integrated quality indicator

α	0,1	0,01	0,001	10^{-4}	10^{-5}	10^{-6}
$\delta_J, \%$	2,94	1,36	1,09	1,07	1,06	1,06

Consistently reducing the regularization parameter from 0,1 to 10^{-6} , we estimate the relative error of prediction of the integrated quality indicator (Table 2)

$$\delta_J = \frac{|J_c - J|}{J}. \quad (25)$$

Therefore, the relative error of the prediction is less than 1,07% that is quite satisfactory result. Note that this small error of prediction is reached at much bigger relative error of assessment of weight coefficients. This is due to the fact that, in the summation composed term $V_{c,i} x_i$, the integrated indicator part J_c has big value, than corresponding composed indicator J and other part have smaller value. Therefore, the effect of partial mutual compensation of composed terms is observed. The last statement can be proved using Cauchy's inequality known from the functional analysis:

$$|(V_c - V, X)| \leq \|V_c - V\| \cdot \|X\|. \quad (26)$$

3. Neural Network: Procedure of Training

For overcoming the problem of multicollinearity of reference data, the neural networks methods are perspective [14–17]. When training (calibration) of a model by a neural network, formula of LMS (3) is not used at all and, therefore, equality to zero of the determinant $|X^T X|$ does not matter. In neural networks, as a rule, control of the weights w_i by criterion of mistakes functional minimum is carried out by gradient methods. However, these methods rather slowly work and demand a large number of iterations. Therefore, further, for control of a model, we propose an alternative approach.

Newton's method [15,18] can be considered as an alternative to the method of gradients for the accelerated training of neural networks. The main step of this method is defined by the expression

$$\omega_{j+1} = \omega_j - H_J^{-1} g_j,$$

where w_j is a vector of the adjusted neuron weights $w_j = (w_{1j}, w_{2j}, \dots, w_{mj})^T$; H_j is a Hesse's matrix of the second private derivatives of the mistakes functional in the configured settings; g_j is a gradient of the mistakes functional. Procedures of minimization on the basis of Newton's method, as a rule, work quicker than the same procedures on the basis of the interfaced gradients method. The best one is the class of quasi Newtonian algorithms, which uses some approximate assessment of a Hesse's matrix at each iteration.

In particular, the Levenberg–Markuardt's algorithm (LM) realizes the following strategy for assesment of a Hesse's matrix. Under the assumption that the functional is defined as the sum of squares of mistakes that is a characteristic when training neural networks with direct transfer, the Hessian can be approximately calculated as

$$H \cong Y^T Y$$

and the gradient is calculated by the formula

$$g = Y^T e,$$

where $Y = \partial e / \partial W$ is the Jacoby's matrix of mistake functional derivatives in the configured settings; e is the vector of mistakes of a neural network.

The LM algorithm uses approximation of a Hessian of the form

$$w_{j+1} = w_j - (Y^T Y + \mu I)^{-1} Y^T e_j.$$

If coefficient $\mu = 0$, then we have Newton's method, while if μ is bigger, then we obtain the method of gradient descent with a small step.

Let us make the following hypothesis: as at the solution of a direct examination problem (assessment of the integrated indicator J on the set of signs X and the vector of weight coefficients V), two-stage division of weights on group and partial, natural way to solve the inverse problem (assessment of weights on the basis of the sets X, J) is also to use application of a two-layer linear network.

Let us construct a two-layer neural network of direct signal transmission with the return distribution of a mistake (feed-forward backprop) in the beginning. To this end, we use means of GUI (Graphical User Interface) of the *nntool* package of the MATLAB environment (Fig. 1) [19]. Here the first layer is hidden, the second layer is output; $W^{1,2}$ are matrixes of weights of the first and second layers; $b^{1,2}$ are constants of shifts signals of the first and second layers. Numbers given at the bottom of the corresponding blocks show the corresponding parameters of the network: 4 is the dimension of a vector of an entrance, 1 is existence of one neuron at the hidden and output layers. As shown, linear functions of activation purelin are chosen in both layers.

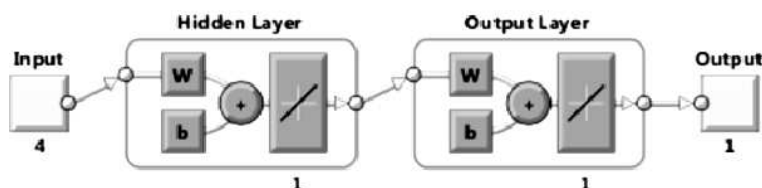


Fig. 1. Block diagram of two-layer neural network

Two approaches can be applied to the calibration of a neural model: training with the teacher (train) and consecutive adaptation (adapt) [20]. In both cases, at an entrance of network move data (19), the vector of indicators J_{empl} is chosen as a target vector.

For training of network, we use the Levenberg–Marquardt's algorithm. After 5 steps of training, square of the mean square error of calibration of the model

$$CKO = \frac{1}{n} \sum_{i=1}^n (\delta J_1)^2 \quad (27)$$

decreases to negligible value (Fig. 2).

After the end of a procedure of training, the weights and shifts of the network neurons in terms of the procedure *nntool* take the form

$$\begin{aligned} iw\{1, 1\} &= [0, 096 \quad 0, 912 \quad 0, 926 \quad -0, 687]; \\ lw\{2, 1\} &= [0, 904]; \quad b\{1\} = [0, 939]; \quad b\{2\} = [-0, 03]. \end{aligned} \quad (28)$$

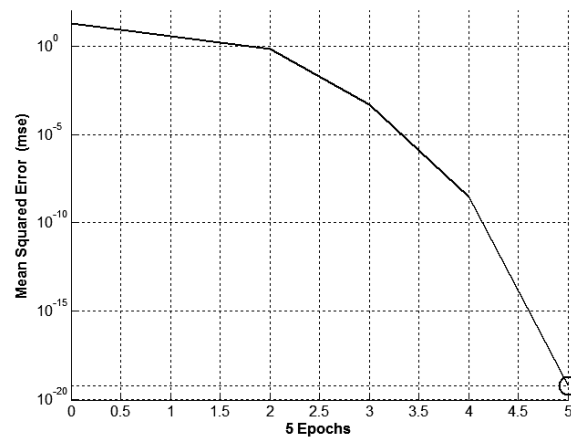


Fig. 2. Training quality of two-layer neural network

Let us calculate the integrated “quality-price” indicator of the new (test) ISS “AST-I” using assessment (28) of weights and shifts of a two-layer neural network received after calibration. The characteristic vector X_{Test} of this sample is calculated by the same technique which was applied when determining matrix (19), and has form (24).

Comparing the received value $J_{Test} = 0,681$ and the exact value $J = 0,703$, we are convinced that the relative error of prediction of the integrated indicator of the new ISS “AST-I” is $\delta_J = 0,031$, or 3,1%. Note that the accuracy of the neural modelling prediction is slightly lower than when using the method of ridge regression (the error is 1,06%)

The obtained results allow to confirm the hypothesis given above: application of a two-layer neural network is enough for the solution of the inverse problem of examination. Besides, at numerical modelling, the method of training revealed the following features: considerable dispersion of weights and shifts of a neural network, ambiguity of the solution due to the choice of accidental entry conditions, strong dependence on a training algorithm.

4. Consecutive Adaptation of Neural Network

At high accuracy of estimates of the generalized indicators, a two-layer neural network has the noticeable shortcoming: training is made in a multidimensional space (weights of the hidden neuron, weights of an output neuron, signals of the general shift) and therefore the received estimates of neurons weights considerably differ from initial vector of scales (15) used for synthesis of reference data in the course of solution of the direct problem of the carried-out examination.

To eliminate such discrepancy, unlike the method of training of a neural network considered earlier with use of the automated editor of GUI, we apply the procedure of adaptation on the basis of our program in the MATLAB environment (the so-called M-file) and using specific designations of neural networks.

Let us choose a 4-dimensional space of scales corresponding to number of signs, and consider a single-layer network. Namely, construct a single-layer neural network, for the chosen linear function of activation as it corresponds to the rules of construction of the integrated indicator described above. At the same time, there is an elimination of the second layer (see Fig. 1). Besides, we consider the reference values of the weights w_i^1 and

shift b^1 to be zero:

$$net.IW\{1\} = [0 \ 0 \ 0 \ 0]; \quad net.b\{1\} = 0. \quad (29)$$

For the considered example with a four-dimensional vector of an entrance, the block diagram of network takes the form given in Fig. 3.

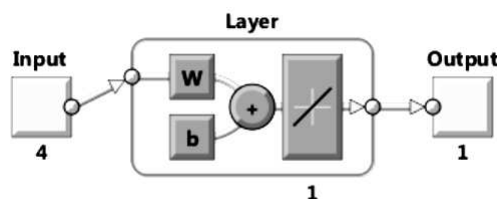


Fig. 3. The block diagram of a single-layer network with a linear function of activation

In the course of adaptation, we apply the procedure nullifying the control shift speed lr to an exception of changes of zero shift

$$net.biases\{1, 1\}.learnParam.lr = 0. \quad (30)$$

For rather effective control of scales, we choose the speed of adaptation of a compromise between the speed and stability of the process:

$$net.inputWeights\{1, 1\}.learnParam.lr = 0, 01. \quad (31)$$

This parameter is very important to assess the degree of stability of an algorithm [19]. However, it is necessary to determine lr in the course of calculations by the trial and error method.

Vectors of rated signs of various objects (19) are consistently analyzed. For their description, we choose the specific structure of the MATLAB environment, i.e. the array of elements (cell)

$$P = \begin{bmatrix} 0,538 & 1,000 & 1,000 & 0,545 \\ 0,667 & 0,000 & 1,000 & 0,960 \\ 0,789 & 1,000 & 0,770 & 1,000 \end{bmatrix}. \quad (32)$$

As a target vector to which exits of neural network will aspire, the vector of integrated indicators is considered (also the array of elements)

$$T = (0,652 \ 0,775 \ 0,871). \quad (33)$$

Fig. 4 presents results of adaptation by consecutive consideration of vectors of an entrance for a single-layer neural network.

Let us consider the number of steps of adaptation to be equal 100. As a result, we obtain estimates of the separate indicators $J_1 - J_3$:

$$y_1 = 0,7216, \quad y_2 = 0,7254, \quad y_3 = 0,8493, \quad (34)$$

and the average square of LMS of calibration calculated by formula (27) is equal to 0,0027. As we see, after the procedure of adaptation, the constructed vector of weight of a neuron

$$i\omega\{1\} = (0,2247 \ 0,1503 \ 0,2863 \ 0,3012)^T \quad (35)$$

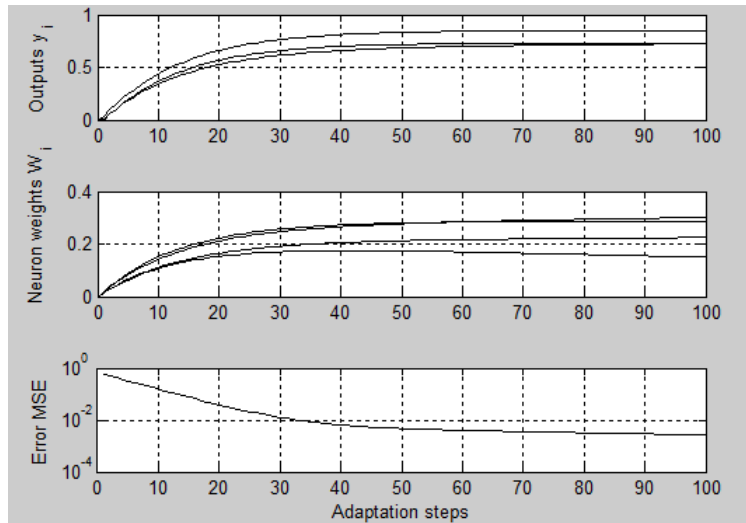


Fig. 4. Values of exits, neuron scales, and mistakes at consecutive consideration of vectors of an entrance

contains only values of weights from the interval $[0; 1]$ that corresponds to a common practice of expert systems [5, 7], unlike data (28) when the adjusted weights can take negative values. This is an important advantage of the consecutive adaptation method over the previously discussed method using the GUI graphical editor procedure.

Let us apply to modelling a characteristic vector of a new sample (in the form of the array of elements)

$$P_{Test} = [0, 715 \quad 1, 000 \quad 0, 460 \quad 0, 750]. \quad (36)$$

For assessment of an indicator of a new object, we use the procedure of modelling (sim) of the MATLAB language

$$[Y_{Test}, pf] = sim(net, P_{Test}) \quad (37)$$

and, for a test object, obtain the value of the indicator of quality with the relative error 4,8% (in comparison with the exact value 0,703). The value of the found error is quite admissible for prediction of the integrated indicator of quality of new objects, especially under the conditions of multicollinearity of basic data.

Therefore, all three considered above methods allow to overcome the multicollinearity problem when forecasting the integrated indicator of a new object on the basis of the previously known reference data on similar objects.

Conclusion

1. In the course of examination of a model, when using multiple linear regression for high-quality training of a model, it is required that the number of objects (samples) k considerably exceeds the amount of signs of an object m : $k > (2 \div 3) m$. However, in practice, this condition often is not satisfied and there exists the multicollinearity problem in such cases.

2. A reliable way of overcoming this problem is to use the method of ridge regression based on the theory of solution of incorrect problems. It is shown that introduction of the

regular adding LMS composed terms in the equation gives an approximate solution with the sufficient degree of accuracy.

3. Another way is to use the two-layer neural network *feed-forward backprop* and the procedure of supervised learning train on the basis of the package *nntool* of the graphic user interface GUI of the MATLAB environment. An advantage of such approach is the lack of need to develop a special program (M-file).

4. A lack of the previous approach is that configured weights of a neural network take chaotic (and even negative) values that contradicts customary practice of examination. For overcoming this shortcoming, we propose a transition to consecutive adaptation of a single-layer neural network with fixing shifts of neurons at zero level.

References

1. Bukharin S.V., Melnikov A.V., Menshih V.V. Method for Expanding the Object Base of Examination by Stitching Solutions of Hierarchy Analysis Method. *Bulletin of the South Ural State University. Series: Mathematical Modelling, Programming and Computer Software*, 2017, vol. 10, no. 2, pp. 74–82. DOI: 10.14529/mmp170206
2. Esbensen K.H. *Multivariable Data Analysis in Practice*. Oslo, CAMO Software AS, 2019.
3. Dreyper N., Smit G. *Applied Regression Analysis*. N.Y., John Wiley and Sons, 2007.
4. Ayvazyan S.A. *Prikladnaya statistika. Klassifikatsiya i snizhenie razmernosti* [Application Statistics. Classification and Reduction of Dimension]. Moscow, Finance and Statistics, 1989.
5. Magnus Ya.R., Katyshev P.K., Peresetsky A.A. *Econometrics*. Moscow, Delo, 1997.
6. Witten I.H., Frank E., Hall M.A. *Data Mining: Practical Machine Learning Tools and Techniques*. N.Y., Morgan Kaufmann, 2011.
7. Bukharin S. *Metody teorii neyronnykh setey v ehkspertize tekhnicheskikh i ehkonomicheskikh ob'ektov* [Methods of the Theory of Neural Networks in the Examination of Technical and Economic Objects]. Voronezh, Scientific Book, 2015. (in Russian)
8. Akhlyustin S. Fuzzy-Multiple Modeling of Generalized Hazard Indicators of Fuel and Energy Complex Facilities. *Bulletin of the Voronezh Institute of the Ministry of Internal Affairs of Russia*, 2018, no. 3, pp. 44–55. (in Russian)
9. Melnikov A., Akhlyustin S. Determination of the Particular Criterion of Integrated Security Systems by the Symmetric Method of Hierarchy Analysis. *Bulletin of the Voronezh Institute of the Ministry of Internal Affairs of Russia*, 2018, no. 2, pp. 37–44. (in Russian)
10. Melnikov A. Neural Modeling of the Generalized Quality Indicator of the Repeaters of the Notification Transmission Systems. *Bulletin of the Voronezh Institute of the Ministry of Internal Affairs of Russia*, 2015, no. 1, pp. 17–26. (in Russian)
11. Saaty T.L. Relative Measurement and Its Generalization in Decision Making Why Pairwise Comparisons Are Central in Mathematics for the Measurement of Intangible Factors the Analytic Hierarchy/Network Process. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, 2008, vol. 102, no. 2, pp. 251–318. DOI: 10.1007/BF03191825
12. Akhlyustin S. [Formation of a Complex Indicator of the Quality of Objects with Poorly Formalized Signs]. *Digitization of the Agro-Industrial Complex*, Tambov, Tambov State Technical University, 2018, pp. 185–187. (in Russian)
13. Tikhonov A. *Metody resheniya nekorrektnykh zadach* [Methods for Solving Ill-Posed Problems]. Moscow, Nauka, 1986. (in Russian)

14. Haykin S. *Neural Networks: a Comparison Foundation* New Jersey, Prentice Hall, 2001.
15. Medvedev V. *Neyronnye seti. Matlab 6* [Neural Networks. Matlab 6]. Moscow, Dialog-MYTHI, 2002. (in Russian)
16. Yakhyayeva G. *Nechetkie mnozhestva i neyronnye seti* [Fuzzy Sets and Neural Networks]. Moscow, Internet University of Information Technologies, 2011. (in Russian)
17. Diligensky N., Dymova L., Sevastyanov P. *Nechetkoe modelirovanie i mnogokriterial'naya optimizatsiya proizvodstvennykh sistem v usloviyakh neopredelennosti: tekhnologiya, ekonomika, ekologiya* [Fuzzy Modeling and Multi-Criteria Optimization of Production Systems in Conditions of Uncertainty: Technology, Economics, Ecology]. Moscow, Mashinostroenie, 2004. (in Russian)
18. Bukharin S. Prediction of a Generalized Indicator of the Quality of Protection Devices Based on Ridge Regression and Neural Modeling. *Bulletin of the Voronezh Institute of the Ministry of Internal Affairs of Russia*, 2014, no. 1, pp. 57–65. (in Russian)
19. Avsentiev O.S., Meshcheryakova T.V., Navoev V.V. Sequential Application of the Hierarchy Analysis Method and Associative Training of Neural Network in Examination Problems. *Bulletin of the South Ural State University. Series: Mathematical Modelling, Programming and Computer Software*, 2017, vol. 10, no. 3, pp. 142–147. DOI: 10.14529/mmp170312
20. Boyarkin D.A., Krupenev D.S., Yakubovsky D.V. Using Machine Training Methods for Assessing the Reliability of Electric Power Systems by the Monte–Carlo Method. *Bulletin of the South Ural State University. Series: Mathematical Modelling, Programming and Computer Software*, 2018, vol. 11, no. 4, pp. 146–153. (in Russian) DOI: 10.14529/mmp180411

Received July 23, 2019

УДК 004.891

DOI: 10.14529/mmp200406

ПРЕДСКАЗАНИЕ ИНТЕГРАЛЬНОГО ИНДИКАТОРА КАЧЕСТВА НОВОГО ОБЪЕКТА В УСЛОВИЯХ МУЛЬТИКОЛЛИНЕАРНОСТИ РЕФЕРЕНТНЫХ ДАННЫХ

С.Б. Ахлюстин¹, А.В. Мельников¹, Р.А. Жилин¹

¹Воронежский институт МВД России, г. Воронеж, Российская Федерация

Предсказание состояния нового объекта при недостатке известных характеристик и оценок показателей качества ряда изученных объектов (множества референтных данных) зачастую приводит к проблеме мультиколлинеарности исходных данных.

Предлагаются три способа преодоления этой проблемы, относящиеся к сфере data mining: использование гребневой (*ridge*) регрессии, обучение с учителем двухслойной нейронной сети, последовательная адаптация однослойной нейронной сети и сравниваются их характеристики. В методе гребневой регрессии введение регуляризирующего слагаемого в уравнение МНК дает приближенное решение с достаточной степенью точности. Недостатком использования двухслойной нейронной сети «feed-forward backprop» и процедуры обучения с учителем «train» является то, что настроенные веса нейронной сети принимают хаотичные (и даже отрицательные) значения, что противоречит обычной практике экспертизы. Выявлены следующие особенности: значительный разброс весов и смещений нейронной сети, неоднозначность решения за счет выбора случайных начальных условий, сильная зависимость от алгоритма обучения. Для преодоления этого недостатка предложен переход к последовательной адаптации однослойной нейронной сети с фиксацией смещений нейронов на нулевом уровне.

Ключевые слова: экспертиза объектов; прогнозирование; множественная регрессия; гребневая регрессия; регуляризация; нейронная сеть; обучение модели; адаптация.

Литература

1. Bukharin, S.V. Method for Expanding the Object Base of Examination by Stitching Solutions of Hierarchy Analysis Method / S.V. Bukharin, A.V. Melnikov, V.V. Menshih // Вестник ЮУрГУ. Серия: Математическое моделирование и программирование. – 2017. – Т. 10, № 2. – С. 74–82.
2. Esbensen, K. Multivariable Data Analysis in Practice / K.H. Esbensen. – Oslo: CAMO Software AS, 2019.
3. Dreyper, N. Applied Regression Analysis / N. Dreyper, G. Smit. – N.Y.: John Wiley and Sons, 2007.
4. Айвазян, С.А. Прикладная статистика. Классификация и снижение размерности / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1989.
5. Магнус, Я.Р. Эконометрика / Я.Р. Магнус, П.К. Катыхев, А.А. Пересецкий. – М.: Дело, 1997.
6. Witten, I.H. Data Mining: Practical Machine Learning Tools and Techniques / I.H. Witten, E. Frank, M.A. Hall. – N.Y.: Morgan Kaufmann, 2011.
7. Бухарин, С.В. Методы теории нейронных сетей в экспертизе технических и экономических объектов / С.В. Бухарин, В.В. Навоев. – Воронеж: Научная книга, 2015.
8. Ахлюстин, С.Б. Нечетко-множественное моделирование обобщенных показателей опасности объектов топливно-энергетического комплекса / С.Б. Ахлюстин, А.В. Мельников, И.В. Щербакова // Вестник Воронежского института МВД России. – 2018. – № 3. – С. 44–55.
9. Мельников, А.В. Определение частного критерия интегрированных систем безопасности симметричным методом анализа иерархий / А.В. Мельников, С.Б. Ахлюстин // Вестник Воронежского института МВД России. – 2018. – № 2. – С. 37–44.
10. Мельников, А.В. Нейронное моделирование обобщенного показателя качества ретрансляторов систем передачи извещений / С.В. Бухарин, А.В. Мельников, В.В. Навоев // Вестник Воронежского института МВД России. – 2015. – № 1. – С. 17–26.
11. Saaty, T.L. Relative Measurement and Its Generalization in Decision Making Why Pairwise Comparisons Are Central in Mathematics for the Measurement of Intangible Factors the Analytic Hierarchy/Network Process / T.L. Saaty // Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas. – 2008. – V. 102, № 2. – P. 251–318.

12. Ахлюстин, С.Б. Формирование комплексного показателя качества объектов со слабо-формализуемыми признаками / С.Б. Ахлюстин, И.Р. Нарушев, С.А. Мальцев // Цифровизация агропромышленного комплекс. – Тамбов: Тамбовский государственный технический университет, 2018. – С. 185–187.
13. Тихонов, А.Н. Методы решения некорректных задач / А.Н. Тихонов, В.Я. Арсенин. – М.: Наука, 1986.
14. Haykin, S. Neural Networks: a Comparison Foundation / S.Haykin. – New Jersey: Prentice Hall, 2001.
15. Медведев, В.С. Нейронные сети. Matlab 6 / В.С. Медведев, В.Г. Потемкин. – М.: Диалог-МИФИ, 2002.
16. Яхъяева, Г.Э. Нечеткие множества и нейронные сети / Г.Э. Яхъяева. – М.: Интернет-университет информационных технологий, 2011.
17. Дилигенский, Н.В. Нечеткое моделирование и многокритериальная оптимизация производственных систем в условиях неопределенности: технология, экономика, экология / Н.В. Дилигенский, Л.Г. Дымова, П.В. Севастьянов. – М.: Машиностроение, 2004.
18. Бухарин, С.В. Прогнозирование обобщенного показателя качества приборов охраны на основе гребневой регрессии и нейронного моделирования / С.В. Бухарин, А.В. Мельников, В.В. Навоев // Вестник Воронежского института МВД России. – 2014. – № 1. – С. 57–65.
19. Avsentiev, O.S. Sequential Application of the Hierarchy Analysis Method and Associative Training of Neural Network in Examination Problems / O.S. Avsentiev, T.V. Meshcheryakova, V.V. Navoev // Вестник ЮУрГУ. Серия: Математическое моделирование и программирование. – 2017. – Т. 10, № 3. – С. 142–147.
20. Бояркин, Д.А. Использование методов машинного обучения при оценке надежности электроэнергетических систем методом Монте-Карло / Д.А. Бояркин, Д.С. Крупнев, Д.В. Якубовский // Вестник ЮУрГУ. Серия: Математическое моделирование и программирование. – 2018. – Т. 11, № 4. – С. 146–153.

Сергей Борисович Ахлюстин, старший преподаватель, кафедра «Радиотехнические системы и комплексы охранного мониторинга», Воронежский институт МВД России (г. Воронеж, Российская Федерация), serg7676@yandex.ru.

Александр Викторович Мельников, профессор, доктор технических наук, доцент, кафедра «Математика и моделирование систем», Воронежский институт МВД России (г. Воронеж, Российская Федерация), meln78@mail.ru.

Роман Андреевич Жилин, адъюнкт, кафедра «Математика и моделирование систем», Воронежский институт МВД России (г. Воронеж, Российская Федерация), zhilin99.zhilin@yandex.ru.

Поступила в редакцию 23 июля 2019 г.