4. Monteiro, P. Minimal and Maximal Fibonacci Representations: Boolean Generation [Text] / P. Monteiro, R. Newcomb // The Fibonacci Quarterly. – 1976. – Vol. 14, Issue 1. – P. 613–638.
5. Vajda, S. Fibonacci & Lucas Numbers, and the Golden Section [Text] / S. Vajda. – Ellis Horwood limited, 1989. – 192 p.
6. Staxov, A. P. Vvedenie v algoritmicheskuyu teoriyu izmerenij [Text] / A. P. Staxov. – Moscow: Sov. radio, 1977. – 288 p.
7. Staxov, A. P. Kody Fibonachchi i zolotoj proporcii kak alternativa dvoichnoj sistemy schisleniya. Chast 1 [Text] / A. P. Staxov. – Germany: Academic Publishing, 2012. – 305 p.
8. United States Patent No. US8983023 B2. Digital Self-Gated Binary Counter [Text] / Gupta N., Agarval A., Goyal G. – No. US 13/935,552; declared: 04.07.2013; published: 17.03.2015. – 23 p.
9. Yeh, C.-H. Designs of Counters with Near Minimal Counting/Sampling Period and Hardware Complexity [Text]: conference / C.-H. Yeh, B. Parhami, Y. Wang. – Pacific Grove, CA, USA, 2000. – P. 894–898. doi: 10.1109/acssc.2000.910642
10. Thamaraiselvan, K. A High Speed CMOS Parallel Counter Using Pipeline Partitioning [Text] / K. Thamaraiselvan, C. Gayathri, N. Divya // International Journal of Engineering Research. – 2013. – Vol. 2, Issue 8. – P. 491–495.
11. Bindu, S. A Review on High Speed CMOS Counter Using Altera MAX300A [Text] / S. Bindu, V. Vineeth, J. Paulin // International Journal of Advanced Research in Computer and Communication Engineering. – 2015. – Vol. 4, Issue 12. – P. 589–592. doi: 10.17148/ijarcce.2015.412140
12. Azarov, O. D. Shvidkodiyuchij reversivnij fibonachchievij lichilnik [Text] / O. D. Azarov, O. I. Chernyak, O. G. Murashhenko // Informacijni texnologii ta komp'yuterna inzheneriya. – 2015. – Issue 1. – P. 27–32.
13. Azarov, O. D. Metod shvidkodiyuchoi obernenoi lichbi z linijnim zrostannyam aparaturnix vitrat pri naroshhuvanni rozryadnosti [Text] / O. D. Azarov, O. I. Chernyak // Visnyk Vinnyc'kogo politehnichnogo instytutu. – 2015. – Issue 2 (119). – P. 57–61.
14. Azarov, O. D. Metod pobudovi shvidkodiyuchix fibonachchievix lichilnikiv [Text] / O. D. Azarov, O. I. Chernyak, O. G. Murashhenko // Problemy informatyzacii' ta upravlinnja. – 2014. – Issue 2 (46). – P. 5–8.
15. Pat. na vynahid № 104939 U Ukrajna. MPK H03K 23/00 (2014.01). Pereshkodostijkyj lichyl'nyk impul'siv Borysenko-Stahova [Text] / Borysenko O. A., Stahov O. P.; zajavnyk ta patentovlasnyk Sums'kyj derzh. un-t. – № 201210506; zajavl. 05.09.2012; opubl. 25.03.2014, Bjul. № 6.

UDC 004.93
DOI: 10.15587/1729-4061.2016.76171

LOCALIZATION OF VECTORS-PATTERNS IN THE PROBLEMS OF PARAMETRIC CLASSIFICATION WITH THE PURPOSE OF INCREASING ITS ACCURACY

Mourad Aouati
Chief Commissioner of Police
Central City Police Department of Constantine
Ali Mendjeli UV 01 Ilot 03 Bt H n°123
Constantine, Algeria

Розглянуто проблему точності класифікації, яка визначається часткою правильно класифікованих об'єктів при використанні параметричних методів. Показано, що нерівність коваріаційних матриць класів приводить до зміщення поверхні, яка їх розділяє, однак це може не суттєво впливати на точність класифікації. Показано, що локалізація векторів-образів на основі планів повного факторного експерименту забезпечує можливість підвищення точності класифікації

Ключові слова: параметрична класифікація, коваріаційна матриця, розпізнавання образів, точність класифікації об'єктів, вектор-образ, кластер

Рассмотрена проблема точности классификации, определяемой долей правильно классифицированных объектов при использовании параметрических методов. Показано, что неравенство ковариационных матриц классов приводит к смещению разделяющей поверхности, однако это может не существенно влиять на точность классификации. Показано, что локализация векторов-образов на основе планов полного факторного эксперимента обеспечивает возможность повышения точности классификации

Ключевые слова: параметрическая классификация, ковариационная матрица, распознавание образов, точность классификации объектов, вектор-образ, кластер

1. Introduction

Solving the problems of pattern recognition, which appear in many applications, requires, mainly, an increase in the accuracy of classification that is understood in the sense of minimization of the share of incorrect classifications of objects. When such problems are solved for face recognition [1, 2], identification of objects by their text, graphic or video images [3–8], for medical [9, 10] or technical diagnostics [11, 12], the named optimization criterion is a priority. Its achieving depends on the selected method of classification, caused often by specific subject area of the studies Thus, for instance, nonparametric methods are mainly used in pharmacological applications [13] while in the tasks of trouble-shooting in the series-connected units of chemical engineering systems they use parametric methods [14]. It is known that nonparametric classifiers are based on the construction of decision function through approximation of the decision surface, whereas parametric classifiers approximate the function of distribution of experimental data. However, in any case, the task comes down to obtaining a classifying or decisive rule, which makes it possible to relate the object to one of the classes. Certainly, selection of the most suitable method for this application and provision of all conditions, necessary for its correct application, must provide the best solution of the problem in the selected sense. Otherwise the situation is probable when the error of classification is obtained, which leads to unpleasant consequences. For example, a mistaken attribution of a man to the class of "criminal" or "terrorist" and his punishment as the consequence of using insufficiently informative features, which describe his appearance, or errors and incorrectness of the construction of decision function. Impossibility of determining the deviations from normal modes of operation of industrial equipment, caused by incorrect classification of the normal and the nonstandard modes of its operation, may lead to emergencies, etc. All these circumstances allow us to argue about the relevance of the studies, directed toward the search for the ways of increasing the accuracy of classification of objects, first of all, by improving the methods of pattern recognition. The latter implies mandatory consideration of specific features of the studied objects, created by specific applications.

2. Literature review and problem statement

Comparative analysis of parametric and nonparametric methods of classification can be found in the paper [15]. In particular, based on the established information connection between these methods, the algorithm of the combined method of statistical classification was developed, whose essence consists in the following. On the basis of the sample of experimental statistical data, they are divided into two classes and the parameters of discriminant function of the form (1) or (2) are determined:

$$d(x)=-\frac{1}{2}(x-m_1)' S_1^{-1}(x-m_1)+ +\frac{1}{2}(x-m_2)' S_2^{-1}(x-m_2)+\frac{1}{2}\ln\frac{|S_2|}{|S_1|}, \qquad (1)$$

$$d(x)=-\frac{1}{2}(x-m_1)' S_1^{-1}(x-m_1)+ +\frac{1}{2}(x-m_2)' S_2^{-1}(x-m_2)+\frac{1}{2}\ln\frac{|S_2|}{|S_1|}+\lambda. \qquad (2)$$

Here X is the matrix with the number of lines, equal to the number of measurements, and with the number of columns, equal to the number of features that characterize the object, S_i is the covariance matrix of measured parameters of state for the i-th class; $|S_i|$ is the determinant of the matrix S_i; m_1, m_2 are the mathematical expectations of the vector X_{ij} for classes 1 and 2, respectively, d(x) is the value of discriminant function, λ is the threshold of classification, which satisfies constraints on the probability of errors of the 1st and 2nd type, selected on this sequence of the values of the functions d(x).

The parameters of distributions, determined at that, together with the values of the measurements of parameters of the states of the objects for classes 1 and 2 are used for obtaining the vector of initial values c[0]:

$$A=\frac{1}{2}\left(S_2^{-1}-S_1^{-1}\right), \qquad (3)$$

$$b=S_1^{-1}m_1-S_2^{-1}m_2, \qquad (4)$$

$$c=-\frac{1}{2}\left(m_1{}'S_1^{-1}m_1-m_2{}'S_2^{-1}m_2\right)+\frac{1}{2}\ln\frac{|S_2|}{|S_1|}, \qquad (5)$$

$$c=-\frac{1}{2}\left(m_1{}'S_1^{-1}m_1-m_2{}'S_2^{-1}m_2\right)+\frac{1}{2}\ln\frac{|S_2|}{|S_1|}+\lambda, \qquad (6)$$

where $A=\left\|a_{i_1 i_2}\right\|$ is the symmetrical matrix of dimensionality n×n, $b=(b_1,b_2,\ldots,b_n)'$ is the vector of coefficients, respectively, of the members of the second and first degree, relative to x; c is the free coefficient.

Then the procedure of additional training takes place with the use of the algorithm of stochastic approximation by applying the recurrent expression to the vector of initial values c[0]

$$c[k]=c[k-1]+ +\Gamma[k]F'\left(y[k]-c'[k-1]\varphi(x[k])\right)\varphi(x[k]), \qquad (7)$$

where c is the vector of coefficients of decision rule; k is the number of the step of the search for solution; Γ[k] is the diagonal matrix of the values of steps that determine the speed of motion to the optimum point of the vector of coefficients c by different coordinates at the k-th step of training; F is the measure of deviation of approximating function from the optimum; y is the reaction of "teacher", φ(x) is the vector of basic functions.

The procedure of training finishes as soon as the dividing function begins to ensure the assigned accuracy of classification, evaluated as the probability of correct recognition.

$$P=\frac{\sum_{i=1}^{m} l_i}{\sum_{i=1}^{m} L_i}\cdot 100, \qquad (8)$$

where m is the number of classes in the training sample; l_i is the quantity of correctly classified elements of the i-th class in the training sample; L_i is the number of elements of the i-th class in the training sample.

The authors [15] note that the described combined method, based on the combination of two algorithms, ensures high speed of convergence, independent of the laws of distribution of features in the diagnostic sample, and accuracy. In this case, it, however, remains unclear how the algorithms considered the form of the law of distribution of experimental data. In this connection, it was mentioned only that the proposed method has advantages over the method of statistical solutions concerning the accuracy of the objects recognition with the laws of distribution of parameters in the classes, different from the normal one. As the proofing base, they indicated only the percentage indicator of the results of testing on the sample with the volume of 400 measurements.

Researchers who solve specific problems in industrial manufacturing prefer to use parametric methods, based on the Bayes statistics [16, 17]. In this case, the problems of recognition are considered for the purpose of improving technological processes or diagnostics of localization of internal flaws in metal [17, 18]. However, construction of classifying rules is not illustrated by the accuracy of recognition, the authors probably refer to the point of view that the obtained result is sufficient for provision of the set requirements due to the specific conditions of manufacturing. In contrast to this approach, the authors of the paper [8] examine precisely the efficiency of performance of the system of pattern recognition. A characteristic feature of the approach is application of the principle of evaluation by using preliminarily prepared databases of patterns for classes and input patterns. The rule of decision making is based on the criterion of similarity of Dice, applied usually for classification of input patterns, represented by the two-dimensional spectrum of the video image [19]. Similarity coefficients in the form (9) are calculated with the use of eigenvectors, obtained after orthogonal conversion of the two-dimensional spectrum.

$$D(A,B)=\frac{2\times\sum_{i=1}^{n}a_i b_i}{\sum_{i=1}^{n}a_i^2+\sum_{i=1}^{n}b_i^2}. \quad (9)$$

Particularly interesting is the circumstance, which the authors [8] pay attention to: the use of a large number of vectors of features at the recognition of input patterns leads to the worsening in the quality of recognition. The following substantiation is referred to as an argument: the most informative features are in the first eigenvectors of a pattern while the remaining vectors have considerably smaller values, close to the threshold of noise value. And although this work addresses the issue of the influence of the sample size on the accuracy of classification, the algorithm of obtaining the decision rule and its particular form are not explored.

Interesting conclusions regarding the accuracy of classification are drawn in the paper [7], where it is noted, in particular, that the majority of the algorithms of pattern recognition are representatives specifically of the parametric families

$$A=\left\{\frac{(\alpha)}{\alpha}\in D\right\}.$$

The parameter α, used by them, determines the best of the algorithms

$$A^*(\alpha)=\underset{A^*\in A}{\arg\max}\,\Phi(A)$$

in the sense of maximization of the number of correctly identified objects of the training or control sample. The main disadvantage of this representation of optimization problem, in the opinion of the authors [7], is the fact that the functional of quality Φ(A) can take only integer values. The description of a regularizer is proposed as the alternative in the work, which allows, in the case of impossibility of linear separation of classes, building up the continuum of hyperareas, which correctly classifies all the objects of the sample. In this case, they also draw the conclusion that the obtained optimum hyperplane is the only one. However, the author confines himself to theoretical study of the problem, without providing detailed results of experimental or commercial tests, which would make it possible to estimate the adequacy of the algorithm, proposed in the article [7].

Further development of the accuracy of classification is connected to the use of neuron networks by a number of researchers [20–22]. In this case, the main attention of researchers is concentrated on the quality of training a neuron network. For example, the article [21] described results of applying one of the classical algorithms for developing the subsystem of decision making support in the system of neural network pattern recognition. As the criteria of efficiency, the authors examine the possibility of reduction in the subjectivity and an improvement in the quality of expert solutions when constructing training samples according to statistical data of observations. The method of support of making classification decisions, proposed by the authors, implies sequential execution of the stages of self-organizing the neurons of the Kohonen computational layer, calibration of the elements of the output vector of the training sample and their final marking. As a result of applying the algorithm, the following data are cited: the relative share of correct expert estimations increases on average by 20 % while the relative share of the false ones decreases by 50 %. In this case, however, this work does not present constraints, associated with the use of the proposed method.

In contrast to the results, described in [20], the paper [21] states the fact that the existing basic "classical" models of training, for example, based on the correction of errors, using the memory, competitive training and the Boltzmann's method, possess a number of shortcomings. The latter, according to the authors, include impossibility to apply only one of the models of training in the creation of universal systems of pattern recognition. This is especially obvious in specific subject areas, where the construction of universal quality system of recognition encounters a practically unsolvable problem of taking into account all specific features, inherent to these subject areas. As one of the variants of solving the problem, the authors [21] propose to use the Levenberg-Marquardt method [22], which, in contrast to the classical algorithms of training, uses Z-training by the epochs and an error of network is averaged for the entire epoch of training. The main indicator of the performance efficiency of the algorithm is considered to be the minimum time period of its operation; however, nothing is said about its other advantages, as well as about conditions of its application. In addition to this, the account of results is limited by theoretical description of the algorithm while the data that would make it possible to estimate the claimed possibilities are missing.

One additional aspect should also be noted, not explored in the works described above, the possibility of constructing efficient systems of recognition under conditions when the components of vectors-patterns cannot be measured with a sufficient degree of accuracy. This aspect is examined, for example, in the papers [23–25], devoted to attempts at construction of the efficient systems of recognition, based on clustering, taking into account fuzzy input data. The article [23] demonstrated that the use of the methods of self-organizing maps (SOM) and the principal components analysis (PCA) for the problem of forecasting, in particular, for impending volcanic eruptions, does not provide possibility of direct restoration of the spectra of spectrograms for the qualitative recognition without the a priori knowledge of the set of templates. It is shown that PCA in combination with hierarchical clustering is the more powerful practical tool for the automated identification of characteristic models in the seismic spectra. In this case, in contrast to PCA, the algorithms of general clustering cannot ensure ideal grouping of essential features on the constructed spectrograms. Thus, preference is given specifically to the method of principal components, which makes it possible, at the minimum losses of informativeness, to ensure reduction in dimensionality of the space of features.

Interesting is the approach, described in the paper [24] and based on the point estimation for the conditions of fuzzy pattern recognition. In this approach, the space is considered to consist of two fuzzy sets ("True" and "Deceiver"). In the first stage of the procedure each separate matcher is simulated as fuzzy set on the basis of the method of generation, using the function of automatic membership. The purpose of this procedure is elimination of uncertainty and imperfection in the evaluation of the corresponding points. Then new fuzzy estimations of compliance combine with the fuzzy operator of aggregation, after which the final decision about the recognition is made. The authors note that the proposed method possesses, in particular, high stability. At the same time, the articles [23, 24] should have mentioned the constraints, imposed on the given methods by specific conditions of the subject areas.

The approach to data processing for classification with the use of the methods of computational intellect is described in the paper [25]. The authors proposed the adaptive algorithm of fuzzy clustering with the use of objective function of the form

$$E\left(w_j(k),c_j\right)=\sum_{k=1}^{N}\sum_{j=1}^{m} w_j^{\beta}(k)\,d^2\left(x(k),c_j\right)\to \min, \tag{10}$$

with constraints:

$$\sum_{j=1}^{m} w_j(k)=1,\ k=1,\dots,N,\ 0<\sum_{k=1}^{N} w_j(k)<N,\ j=1,\dots,m. \tag{11}$$

Here $w_j(k)\in[0,1]$ is the level of belonging of the vector x(k) to the j-th cluster, c_j is the centroid of the j-th cluster, $d^2\left(x(k),c_j\right)$ is the distance between x(k) and c_j in the accepted metric, β is the non-negative parameter, named "fuzzifier".

The result of the work of this algorithm is formation of the matrix of fuzzy partition, in which the objects are divided into clusters (diagnoses). The characteristic feature, noted by the authors, is the fact that the form of clusters may vary from hypersphere to hyper-ellipsoid depending on the form of initial data. In other words, decisive is the selection of distance between x(k) and c_j, described in the form

$$d\left(x(k),c_j\right)=\sqrt{\left(x(k)-c_j\right)^T A_j\left(x(k)-c_j\right)}, \tag{12}$$

where A_j is the reverse fuzzy covariance matrix of each cluster.

As noted by the authors [25], a peculiarity of this approach is insensitivity to the ratio of the number of objects to the quantity of indicators that characterize these objects, and also to the law of distribution of data. But it is not defined what requirements the area of location of initial data in the space of features must satisfy.

Such a brief survey allows us to state that the number of questions, which are not paid sufficient attention to, should include the stage of data preparation, which precedes the procedure of constructing classifying rules. In particular, the discussion may tackle the conversions of clustering even when there is no need for decreasing the dimensionality of the space of features with the minimum loss of informativeness. The discussion rather deals with the selection of the area of the space of features and maintaining the conditions on the compliance of the law of their distribution to the normal one and to the equality of covariance matrices of the divided classes.

3. The purpose and objectives of the study

The purpose of this work is to examine the influence of localization of vectors- patterns in the space of features on the position of the dividing surface, whose results may be used for the subsequent purposeful selection of the optimum algorithm of parametric classification.

To achieve the set goal, the following tasks are to be solved:

– the study of the nature of influence of the inequality of covariance matrices of classes on the position of the dividing surface and the accuracy of parametric classification;

– formation of directions of selecting the area of the space of features as the factor of improvement of the accuracy of parametric classification.

4. Numerical experiments on construction classifying rules for arbitrary area of the space of features of dimensionality (N×2)

Let, as a result of normalization, the values of the variables x_i that characterize position of objects in the space of features of dimensionality (N×2) entered the range [–1; +1]. Let it also be known that N_1 objects belong to class A and N_2 objects belong to class B, moreover, $N_1=N_2$. The area of values for both classes is arbitrary (Fig. 1). The problem is to construct a straight line with the aid of the methods of parametric classification, which divides both classes in the space of features. In this case, taking into account that localization of both areas is arbitrary, the components of vectors– patterns are selected arbitrarily and there are no constraints from the subject area, let us name this task as a test.

It is natural to assume that the first stage, which precedes the procedure of construction of classification rule, must be checking the compliance of the laws of distribution of variables x_1 and x_2 with the normal one. The parameters of distributions, in this case, may be assessed as follows:

$$m\left(x_i\right)=\frac{1}{N}\sum_{j=1}^{N} x_i^j, \tag{13}$$

$$S^2\left(x_i\right)=\frac{1}{N-1}\sum_{j=1}^{N}\left[x_i^j-m\left(x_i\right)\right]^2, \tag{14}$$

$$S\left(x_i\right)=\sqrt{S^2\left(x_i\right)}, \tag{15}$$

where x_i^j is the value of the i-th variable in the j-th experiment, $m(x_i)$ is the mathematical expectation of the i-th variable,

$S^2(x_i)$ is the estimation of dispersion of the i-th variable, mean-square deviation, N is the number of objects in the class.

The theoretical probability of distribution is calculated based on the analytical description of the density function of the probability of the normal law of distribution:

$$p=\varphi(x_i)\Delta x_i=\frac{\Delta x_i}{S\sqrt{2\pi}}e^{-\frac{\left(x_i^j-m(x_i)\right)^2}{2S^2}}, \quad (16)$$

where Δx_i is the width of the interval of values of the i-th random magnitude, equal to $0{,}5S_i$ and plotted on the x-axis of the histogram of distribution.

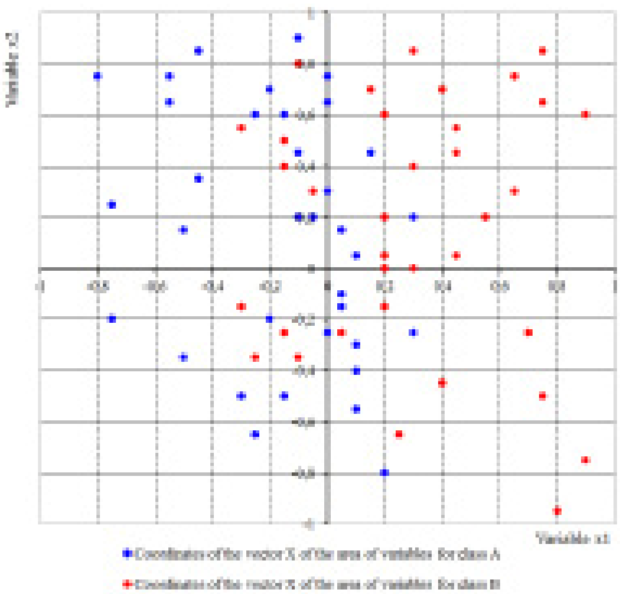

Fig. 1. Space of features for the test problem

The results of data processing with the use of (13)–(16) for classes A and B are represented in Fig. 2–5.

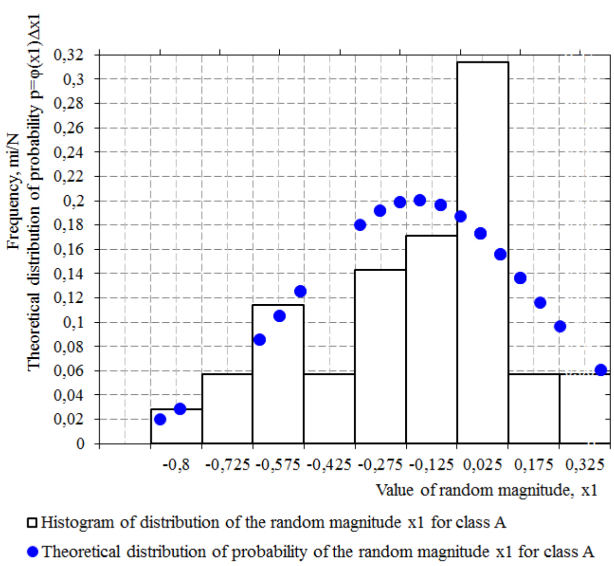

Fig. 2. Histogram of distribution and the envelope curve of theoretical distribution of probability of the random magnitude x_1 for class A. Parameters of distribution: m=–0,161429, S=0,298301

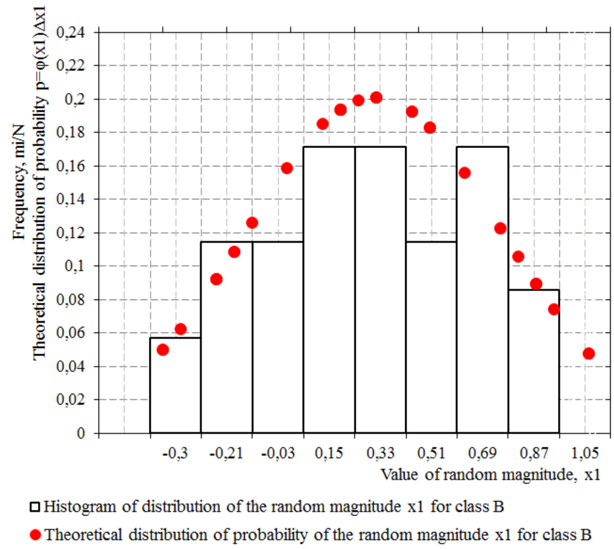

Fig. 3. Histogram of distribution and the envelope curve of theoretical distribution of probability of the random magnitude x_1 for class B. Parameters of distribution: m=0,295714, S=0,357354

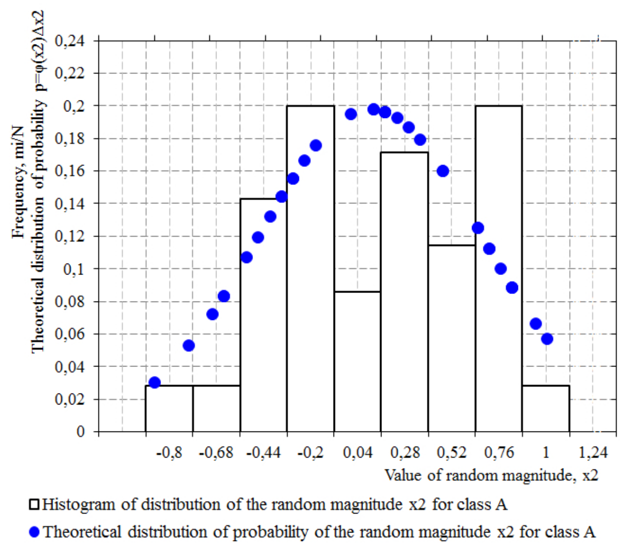

Fig. 4. Histogram of distribution and the envelope curve of theoretical distribution of probability of the random magnitude x_2 for class A. Parameters of distribution: m=0,135714, S=0,48364

In a general case it is necessary to verify the hypothesis about the compliance of the distribution law to the normal one, as one of the main prerequisites of the correctness of applying the Bayes statistics for the construction of a classifying rule. However, taking into account that the initial data were selected arbitrarily, it is sufficient to confine ourselves to the visual analysis of matching experimental and theoretical description. This analysis shows that the hypothesis about the normal distribution will be rejected in all four variants. Nevertheless, if there is reason to believe that in the specific case the distribution law nevertheless corresponds to normal, further solution of the problem of construction of the dividing straight line is possible. In favor of this assumption one may put forward the following substantiation. When looking at Fig. 2–5, it is possible to see that in each of the cases the data can be burdened by systematic errors, which is observed in the form of several peaks on the histograms of distribution. This can lead to asymmetry on the density curves of the probability of distribution. Moreover, exactly this situation is often revealed at the attempts to evaluate distribution for the actually functioning objects, for example, in the industry. If such systematic errors are removed, then, probably, verification of the hypothesis, which postulates normal law of distribution, will yield a positive result. Taking into account that the data sample for both classes is formed in an arbitrary manner, let us suppose that this is exactly the case.

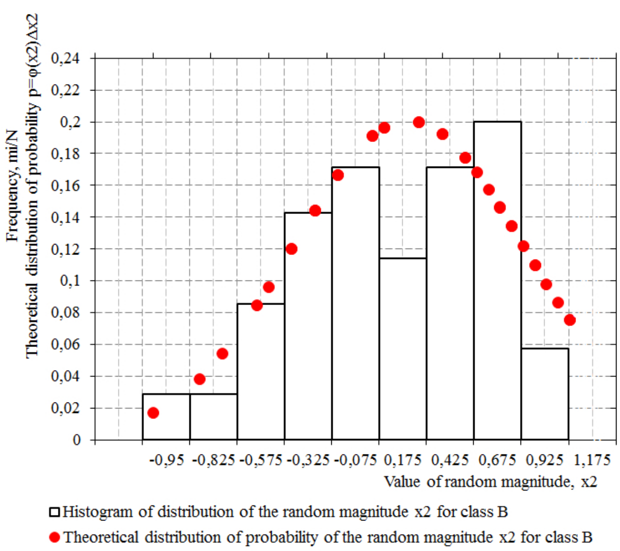

Fig. 5. Histogram of distribution and the envelope curve of theoretical distribution of probability of the random magnitude x_2 for class B. Parameters of distribution: m=0,154286, S=0,496957

The classifying rule with the use of the likelihood ratio $l(x^j)$ can be represented in the form [14]

$$L(x^j)=\frac{P(A)p_A(x^j)}{P(B)p_B(x^j)}=\frac{P(A)}{P(B)}l(x^j), \tag{17}$$

where P(A), P(B) are the priori probabilities of classes, and the densities of distribution of the probabilities $p_A(X)$ and $p_B(X)$ take the form:

$$p_A(X)=d\exp\left(-\frac{1}{2}(x-m_A)^T\operatorname{cov}^{-1}(x)(x-m_A)\right),$$

$$p_B(X)=d\exp\left(-\frac{1}{2}(x-m_B)^T\operatorname{cov}^{-1}(x)(x-m_B)\right). \tag{18}$$

Here d is the constant multiplier, m_A, m_B are the mathematical expectations of classes A and B, respectively,

$$\operatorname{cov}(x)=\frac{1}{N}X^TX-m^Tm$$

is the covariance matrix, $\operatorname{cov}^{-1}(x)$ is the matrix, reverse of covariance.

Discriminant function in this case takes the form

$$y=f(x)=x'\operatorname{cov}^{-1}(x)(m_A-m_B). \tag{19}$$

The threshold value of discriminant function, which makes it possible to make a comparison for making a a decision about belonging of the object to the specific class, is calculated based on the equation

$$y_0=\frac{1}{2}(m_A+m_B)'\operatorname{cov}^{-1}(x)(m_A-m_B)-\ln\frac{P(A)}{P(B)}. \tag{20}$$

Let us assume that the priori probabilities of classes are identical: P(A)=P(B)=0,5. Parameters of distributions $p_A(x)$ и $p_B(x)$, calculated based on initial data of the test problem (Fig. 1), comprise:

$$m_A=\begin{pmatrix}-0,16143\\0,135714\end{pmatrix},\ m_B=\begin{pmatrix}0,29571\\0,15429\end{pmatrix},$$

$$\operatorname{cov}_A(X)=\begin{pmatrix}0,08644 & -0,0409\\-0,0409 & 0,22722\end{pmatrix},$$

$$\operatorname{cov}_B(X)=\begin{pmatrix}0,12405 & -0,0085\\-0,0085 & 0,23991\end{pmatrix}.$$

It is evident that the covariance matrices of classes A and B are not equal, i. e., one of the conditions for correct application of the Bayes statistics is not fulfilled – it is known that in this case the bordering surface ceases to be linear. Nevertheless, it is of interest to examine this situation in order to understand how the failure to meet this condition will influence the position of the dividing straight line in the space (N×2) and the accuracy of classification connected with it. We, in other words, remove the constraint to the equality of covariance matrices and build a classifying rule, moreover, for each class separately. In this case, the classifying rule, coefficients in the analytical description of which are calculated for class A, takes the form:

$$x^j\in A \text{ if } -5,8225x_1-1,1292x_2\geq -0,5547,$$

$$x^j\in B \text{ if } -5,8225x_1-1,1292x_2< -0,5547. \tag{21}$$

The classifying rule, coefficients in the analytical description of which are calculated for class B, takes the form:

$$x^j\in A \text{ if } -3,6993x_1-0,2082x_2\geq -0,2786,$$

$$x^j\in B \text{ if } -3,6993x_1-0,2082x_2< -0,2786. \tag{22}$$

Fig. 6 presents the results of construction of the dividing straight line, the position of which in the coordinates x_1–x_2 is readily available, by its description from the forms (21) and (22) into the description of straight line “in sections”:

$$\frac{x_1}{a}+\frac{x_2}{b}=1,$$

where a is the section, intercepted on the axis x_1, b is the section, intercepted on the axis x_2.

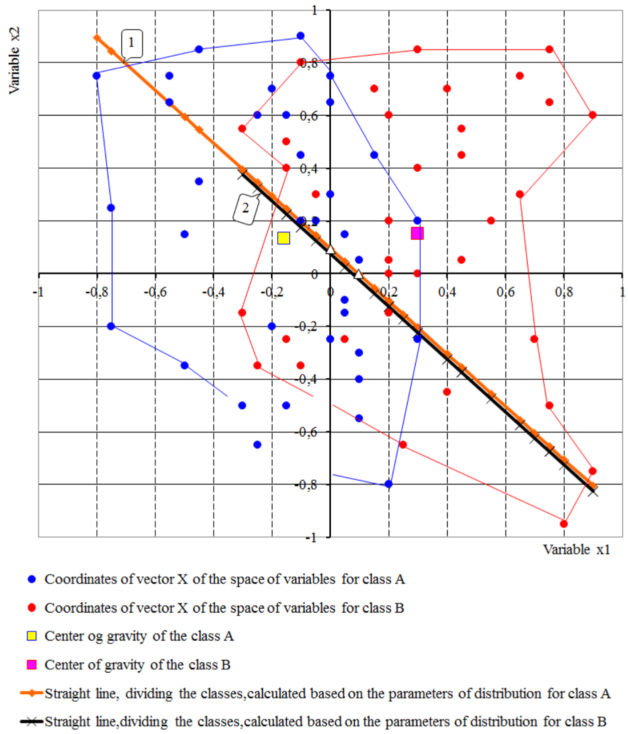

Fig. 6. Results of constructing the dividing straight lines in the space of dimensionality (N×2) for the test problem. Markers Δ designate the sections, intercepted on the axes at the transformation of the equations of straight lines to the form "in sections" (a=0,095, b=0,095 for the straight line 1, a=0,075, b=0,075 for the straight line 2)

It is clear from Fig. 6 that both dividing straight lines are located close to each other (Δa=0,02, Δb=0,02). In this case we observe the overlap of classes A and B, which includes a rather large quantity of objects of both classes. The accuracy of classification, assessed by the probability of inclusion of the objects into an appropriate class with the aid of the expressions (23), (24), amounts to 51,4 % for class A and 74,3 % for class B.

$$P_A = \frac{n_A}{N_A} \cdot 100, \qquad (23)$$

$$P_B = \frac{n_B}{N_B} \cdot 100, \qquad (24)$$

where n_A, n_B is the quantity of correctly classified objects for classes A and B, respectively, N_A, N_B is the total quantity of objects of classes A and B, respectively.

The displacement of the dividing straight lines relative to each other is so small that it does not influence the accuracy of classification in any way. The densities of probabilities of distribution of the value of the discriminant function $y=f(x)$ $p_A(y)$ and $p_B(y)$ for classes A and B, respectively, are given in Fig. 7.

As can be seen from Fig. 7, a classification error of the dividing straight line in the position $y=y_0$ is proportional to the area of the shaded figure ABCD. The manifestation of this fact is connected with the values of covariance matrices of the classes; therefore there is a question if it is possible, by changing the shape of the area of the space of features, to attain an increase in the accuracy of classification. In other words, if it is possible by some purposeful way to localize the area of vectors- patterns in the space of features in order to maximize the share of the correctly classified objects.

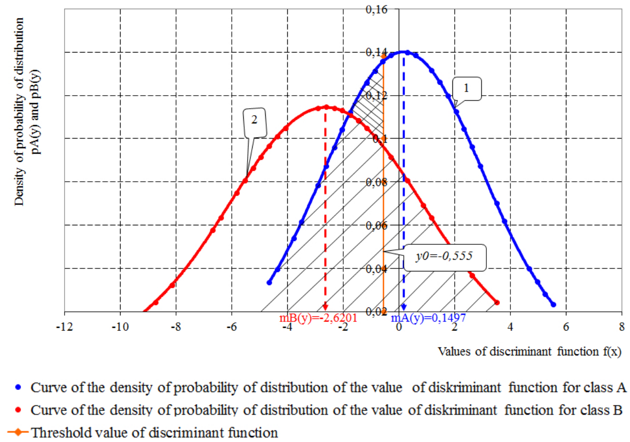

Fig. 7. Densities of probabilities of distribution $p_A(y)$ and $p_B(y)$ of the value of discriminant function for classes A and B: 1 is the curve $p_A(y)$, 2 is the curve $p_B(y)$

5. Localization of vectors-patterns based on the construction of the plans of full factorial experiment

If we refer to the known positions about the fact that the form of the plan of experiment determines the accuracy of estimations of the coefficients of mathematical models, one can assume the following. Assume that the linear dividing surface is built with the aid of the parametric method, moreover, under conditions when the covariance matrices of the divided classes are not equal to each other. Then for each of the classes one can select the area of values of factors– features, for which the accuracy of classification is absolute. In other words, to construct a classification rule, the area of the space of features is selected, not overlapped by the area of another class. For example, this area under the dividing straight line for class A, limited on the right by the left red broken line, i. e., by the left boundary of class B (Fig. 6). By analogy, the area for class B can be selected. Any point, which belongs to this area, according to the said above, can be related to the appropriate class. Any point, which belongs to this region, on the strength of what has been said, can be related to the appropriate class. Such points, therefore, may be the coordinates of the apexes of the plan of full factorial experiment (FFE). If the components of vectors- patterns from classes A and B are selected precisely according to this principle, then two problems are practically solved. The first – the vectors of factors-features will be orthogonal to each other, i. e., the elements of covariance matrix, which are not on the principal diagonal, are zero. The second – covariance matrices for classes A and B will be equal to each other. The latter will be executed under additional condition – if the intervals of variation of variables for classes A and B in terms of the appropriate variables will be equal to each other. In this case, there is a question in what particular part of the area of the space of features for each of the classes the vectors-patterns must be localized. To answer this question, one must perform the calculations of parameters of distributions with the subsequent classification of the objects for different combinations of the plans, placed in the different quadrants of plane in the coordinates x_1–x_2. The results of such calculations are presented in Fig. 8.

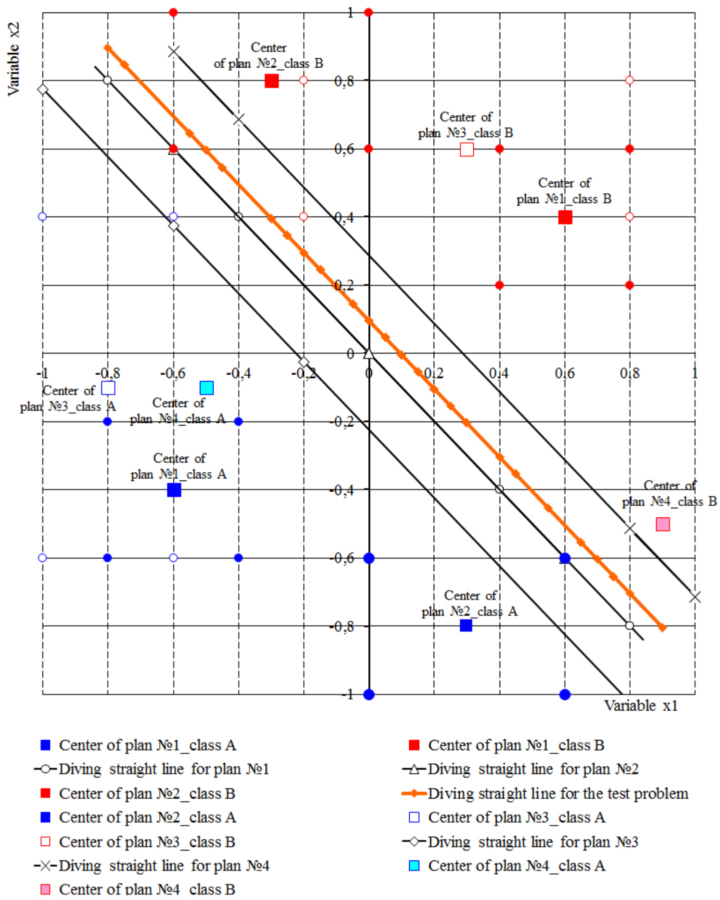

Fig. 8. Position of the dividing straight line for different localizations of the vectors-patterns of classes A and B, selected in accordance with the plans of the full factorial experiment

As can be seen in Fig. 8, the selection of different plans of FFE for calculating the parameters of distributions of classes determines the position of dividing straight line in the space of coordinates x_1–x_2. This means that the selection of the corresponding area of planning can provide, in a general case, for a change in the position of the hyperplane, which divides classes A and B.

6. Discussion of the possibility of increasing the accuracy of classification by localization of vectors-patterns with the aid of the FFE plans

As follows from the obtained results, the localization of vectors-patterns in the selected space of features based on the FFE plans may become the factor that contributes to the increase in the accuracy of classification of objects. The latter is connected with the position of the dividing straight line or, in a general case of multidimensional space of features, the hyperplane. Important in this case is the mutual arrangement of plans for classes A and B relative to each other. For example, as can be seen from Fig. 8, if the plans for both classes are selected at the identical distance from the dividing straight line, then it crosses the origin of coordinates. In this case, it does not matter if the identical intervals of variation for the objects inside one class (plans No. 1 for classes A and B) are selected, or the intervals of variation for x_1 and x_2 inside one class are not equal to each other (plans No. 2 for classes A and B). The main thing is that the intervals of variation of the corresponding variables for classes A and B are equal to each other. In this case, we see (Fig. 8) that the dividing straight lines coincide. By possessing the plans with identical covariance matrices for both classes and arbitrary areas of the space of features of the classes with unequal covariance matrices, it is possible to perform the estimation of the importance of inequality of covariance matrices relative to its influence on the accuracy of classification. It so happens that in this case the displacement of the dividing straight line is insignificant, and one object of class A and one object of class B are placed in the section between two dividing straight lines. This can be seen if one compares results of Fig. 6, 8. In this case, the accuracy of classification of the objects of class B increases while the accuracy of classification of the objects of class A is reduced. Nevertheless, such a change in the accuracy should be considered non-essential.

If we change the geometry of plans in such a way that the interval of variation by the variable x_1 for class A is equal to the interval of variation by the variable x_2 for class B, and the interval of variation by the variable x_2 for class A is equal to the interval of variation by the variable x_1 for class B, then the dividing straight line is displaced downwards (plan No. 3, Fig. 8). In this case, the accuracy of classification of the objects of class B (P_B=0,857) increases considerably, but the accuracy of classification of the objects of class A is lowered (P_A=0,343). Reverse situation is observed when the plans have identical values of the intervals of variation inside each class and for classes A and B by the corresponding variables; moreover, the plan for class A is located in the third quadrant, and for class B – in the fourth. The dividing straight line is displaced upwards in this case, increasing the accuracy of classification for class A (P_A=0,686) and decreasing for class B (P_B=0,514). This can be seen by comparing the results given in Fig. 6 and in Fig. 8.

Obviously, selecting by appropriate way the area of the space of features, in accordance with the FFE plans, there is a possibility of purposeful influence on the accuracy of classification. If one considers that the values of the components of vectors-patterns for the test problem were selected arbitrarily, there is a question to what degree the obtained results can be used for examining the real objects. Since the values of variables were assigned in a standardized form, it can be assumed that, independent of the nature of the variables of the space of features, the obtained results may be used for a wide range of applied problems.

Let us examine as an example the real chemical engineering system, which includes two consecutively operating industrial units – a shaft-type furnace, which works by the principle of countercurrent, and electrical furnace. The task of functioning of this system is obtaining alloy with the required properties, formed owing to the physical-chemical processes that occur in both units. Principal scheme of this chemical engineering system is represented in Fig. 9.

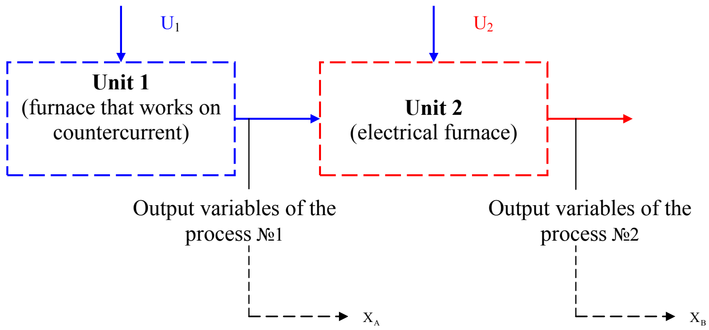

Fig. 9. Simplified schematic of chemical engineering system U_1, U_2 is the control, X_A, X_B are the matrices that describe the space of features for classes A and B, respectively

As the space of features, in accordance with the expert data, concentrations of chemical elements in the alloy, on which the formation of the properties depend [26, 27] can be selected. In the simplified variant, when only the main factors are examined – carbon concentration (designated below as x_1) and the carbonic equivalent of the alloy (designated below as x_2) – the data sample for classes A and B is formed. In this case, the main factors in this type of process are not the particular concentrations, but deviations in their values from the nominal ones, set by the modes of industrial functioning of the equipment. The sample out of 70 objects – 35 for each of the classes – includes standardized values of deviations by both variables.

The task of classification is to determine the belonging of object to class A (deviations from the nominal are caused by the modes of operation of the unit AHTS No. 1) or to class B (deviations from the nominal are caused by the modes of operation of the unit AHTS No. 2). Since the selection of the optimum modes of operation of the unit AHTS No. 2 depends on the values of the space of features, formed at the output from the unit AHTS No. 1 (in Fig. 9, they correspond to the output variables of process No. 1), it is very important to know if it is worthwhile passing from one operating mode to another one.

Fig. 10 demonstrates the results, obtained by the selection of the area of localization of vectors-patterns for classes A and B based on the plans of FFE.

As can be seen from Fig. 10, construction of the classifying rule by the initial sample provides a low indicator of the accuracy of classification for class B (P_B=0,514), while application of the localization, selected in accordance with the plan of FFE, ensures the increase in this indicator to P_B=0,714. In this case, the accuracy of classification of the objects that fall into class A, amounts to P_A=0,886. This testifies to the efficiency of the proposed variant of localization of vectors-patterns based on the selection of plans of FFE for the subsequent construction of classification rules. However, we should note an obvious shortcoming of this study – the lack of functional dependencies, which make it possible to quantitatively evaluate the accuracy of classification at various variants of arrangement of FFE plans. Nevertheless, the obtained results make it possible to see the field of further research, in particular, in the direction of the above-indicated shortcoming and subsequent experimental optimization. The purpose of such optimization mignt be the selection of such coordinates of the centers of plans and their geometric characteristics for classes A and B, which ensure the maximum share of correctly classified objects. Nevertheless, this approach has explicit constraints, connected with the fact that the dividing, in a general case, hyperplane, due to the displacement to the optimum position, will not ensure the accuracy of classification close to 100 %. And this may be linked not only to the linear nature of the dividing surface, but also to the informativeness of the selected variables. Obviously, the latter should include not only the variables that characterize object, but also the parameters, which describe the geometry of clusters in the space of features.

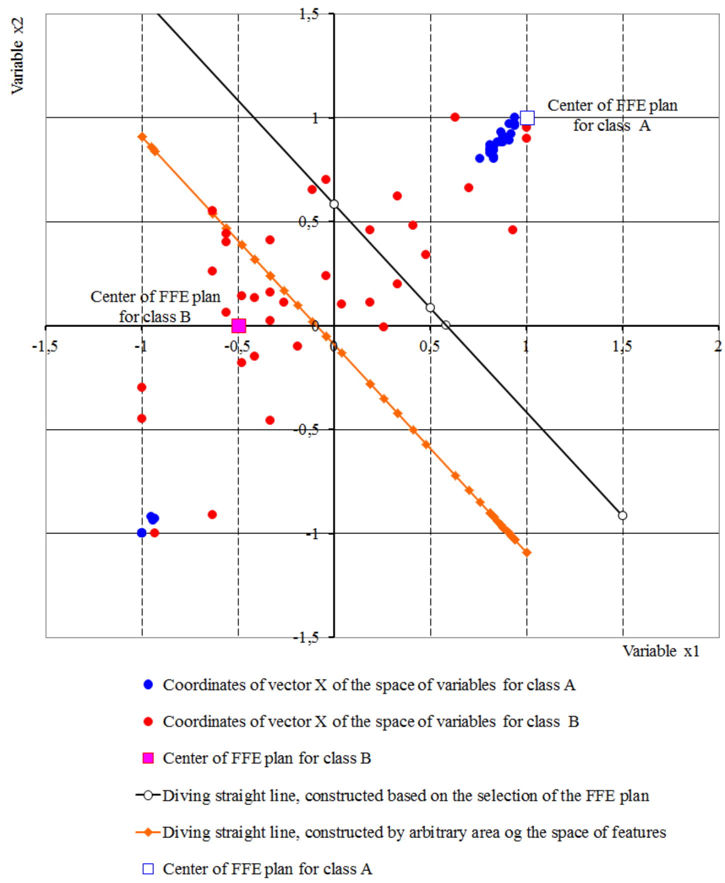

Fig. 10. Position of the dividing straight line for the initial localizations of the vectors-patterns of classes A and B and for the localizations, selected in accordance with the plans of full factorial experiment for the examined chemical engineering system

7. Conclusions

1. It was found that the inequality of covariance matrices of the divided classes with the use of parametric methods of classification leads to the shift of the dividing surface. The magnitude of this displacement may not become the essential factor, which influences the accuracy of classification, which is evaluated by the share of the correctly classified objects. For the space of variables with dimensionality (N×2), it is shown that for the randomly selected classes inside the

square with the length of the edge, equal to two, the accuracy of classification for classes A and B proves to be different. The latter depends on the position of the straight line, which divides classes in the space of factors-features.

2. It is shown that localization of vectors-patterns in the space of features may be selected in accordance with the plans of full factorial experiment. The selection of the appropriate area of positioning the plans for classes A and B is connected with the position of the dividing surface, in this case, as the apexes of the plan, the coordinates of vector-pattern are selected, which belong to the corresponding class with the probability of 100 %. Due to this localization, the equality is ensured of covariance matrices of the classes and the share of the correctly classified objects increases. The reserve for the increase in the accuracy may be in the realization of the purposeful procedures to change the coordinates of plans, for example, with the aid of the methods of experimental optimization.

References

1. Nemrodov, D. The time course of individual face recognition: A pattern analysis of ERP signals [Text] / D. Nemrodov, M. Niemeier, J. N. Y. Mok, A. Nestor // NeuroImage. – 2016. – Vol. 132. – P. 469–476. doi: 10.1016/j.neuroimage.2016.03.006
2. Liu, L. Extended local binary patterns for face recognition [Text] / L. Liu, P. Fieguth, G. Zhao, M. Pietikäinen, D. Hu // Information Sciences. – 2016. – Vol. 358-359. – P. 56–72. doi: 10.1016/j.ins.2016.04.021
3. Florindo, J. B. Local fractal dimension and binary patterns in texture recognition [Text] / J. B. Florindo, O. M. Bruno // Pattern Recognition Letters. – 2016. – Vol. 78. – P. 22–27. doi: 10.1016/j.patrec.2016.03.025
4. Barajas-García, C. Classification of fragments of objects by the Fourier masks pattern recognition system [Text] / C. Barajas-García, S. Solorza-Calderón, J. Álvarez-Borrego // Optics Communications. – 2016. – Vol. 367. – P. 335–345. doi: 10.1016/j.optcom.2016.01.059
5. Kushnirenko, O. A. Correlation algorithm of recognition of character information with low-complexity of technical implementation [Text] / O. A. Kushnirenko, A. V. Cadchenko, A. V. Troyanskiy // Eastern-European Journal of Enterprise Technologies. – 2013. – Vol. 3, Issue 9 (63). – P. 42–45. Available at: http://journals.uran.ua/eejet/article/view/14848/12650
6. Kasian, K. Development of modified method for text recognition in standardized picture [Text] / K. Kasian, V. Bratchykov, V. Shkarupylo // Eastern-European Journal of Enterprise Technologies. – 2015. – Vol. 3, Issue 2 (75). – P. 11–17. doi: 10.15587/1729-4061.2015.43047
7. Fraze-Frazenko, A. Method for the regularization of ill-posed problems of recognition in CCTV [Text] / A. Fraze-Frazenko // Technology audit and production reserves. – 2012. – Vol. 6, Issue 4 (8). – P. 19–20. Available at: http://journals.uran.ua/tarp/article/view/5638/5080
8. Anokhin, M. Moving objects recognition method by their video images [Text] / M. Anokhin // Eastern-European Journal of Enterprise Technologies. – 2014. – Vol. 4, Issue 9 (70). – P. 33–37. doi: 10.15587/1729-4061.2014.26275
9. Voloshyn, N. Identification of iris sign in automated systems of iridology based on Petri Nets [Text] / N. Voloshyn // Technology audit and production reserves. – 2012. – Vol. 4, Issue 1 (6). – P. 5–6. Available at: http://journals.uran.ua/tarp/article/view/4767/4418
10. Alfaro-Ponce, M. Pattern recognition for electroencephalographic signals based on continuous neural networks [Text] / M. Alfaro-Ponce, A. Argüelles, I. Chairez // Neural Networks. – 2016. – Vol. 79. – P. 88–96. doi: 10.1016/j.neunet.2016.03.004
11. Nesterenko, S. A. The wireless computer networks state recognition over the three-dimensional field of directions [Text] / S. A. Nesterenko, A. O. Stanovskiy, O. O. Oborotova // Technology audit and production reserves. – 2015. – Vol. 6, Issue 2 (26). – P. 28–35. doi: 10.15587/2312-8372.2015.56825
12. Unglert, K. Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra [Text] / K. Unglert, V. Radić, A. M. Jellinek // Journal of Volcanology and Geothermal Research. – 2016. – Vol. 320. – P. 58–74. doi: 10.1016/j.jvolgeores.2016.04.014
13. Stuper, A. J. Computer Assisted Studies of Chemical Structure and Biological Function [Text] / A. J. Stuper, W. E. Brugger, P. C. Jurs. – Published by John Wiley & Sons, 1979.
14. Hartmann, K. Statistische Versuchsplanung und auswertung in der Stoffwirtschaft [Text] / K. Hartmann, E. Lezki, V. Schafer. – Leipzig, 1974.
15. Pavlenko, V. D. Povyishenie tochnosti postroeniya reshayuschego pravila v metodah statisticheskoy klassifikatsii [Text] / V. D. Pavlenko, A. A. Fomin // Elektronnoe modelirovanie. – 2001. – Vol. 23, Issue 4. – P. 61–68.
16. Vasenko, Y. Technology for improved wear iron [Text] / Y. Vasenko // Technology audit and production reserves. – 2012. – Vol. 1, Issue 1 (3). – P. 17–21. Available at: http://journals.uran.ua/tarp/article/view/4870/4521
17. Ponomarenko, O. I. Computer modeling of crystallization processes as a reserve of improving the quality of pistons of ICE [Text] / O. I. Ponomarenko, N. S. Trenev // Technology audit and production reserves. – 2013. – Vol. 6, Issue 2 (14). – P. 36–40. Available at: http://journals.uran.ua/tarp/article/view/19529/17205
18. Ignaszaka, Z. Contribute to quantitative identification of casting defects based on computer analysis of X-ray images [Text] / Z. Ignaszaka, P. Popielarskia, K. Krawiecb // Archives of foundry engineering. – 2007. – Vol. 7, Issue 4 (18/4). – P. 89–94.
19. Lin, Z. MatchSim: a novel similarity measure based on maximum neighborhood matching [Text] / Z. Lin, M. R. Lyu, I. King // Knowledge and Information Systems. – 2011. – Vol. 32, Issue 1. – P. 141–166. doi: 10.1007/s10115-011-0427-z
20. Arsirii, E. Development of the decision support subsystem in the systems of neural network pattern recognition by statistical information [Text] / E. Arsirii, O. Manikaeva, O. Vasilevskaja // Eastern-European Journal of Enterprise Technologies. – 2015. – Vol. 6, Issue 4 (78). – P. 4–12. doi: 10.15587/1729-4061.2015.56429

21. Fraze-Frazenko, A. Algorithm of study neural network for image recognition [Text] / A. Fraze-Frazenko // Technology audit and production reserves. – 2012. – Vol. 4, Issue 1 (6). – P. 33–34. Available at: http://journals.uran.ua/tarp/article/view/4781/4432
22. Fertsev, A. A. Uskorenie obucheniya neyronnoy seti dlya raspoznavaniya izobrazheniy s pomoschyu tehnologii NVidia Cuda [Text] / A. A. Fertsev // Vestnik Samarskogo gosudarstvennogo tehnicheskogo universiteta. – 2012. – Vol. 1, Issue 26. – P. 183–191.
23. Unglert, K. Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra [Text] / K. Unglert, V. Radić, A. M. Jellinek // Journal of Volcanology and Geothermal Research. – 2016. – Vol. 320. – P. 58–74. doi: 10.1016/j.jvolgeores.2016.04.014
24. Fakhar, K. Fuzzy pattern recognition-based approach to biometric score fusion problem [Text] / K. Fakhar, M. El Aroussi, M. N. Saidi, D. Aboutajdine // Fuzzy Sets and Systems. – 2016. doi: 10.1016/j.fss.2016.05.005
25. Perova, I. G. (Adaptive treatment of these medico-biological researches by methods of computational intelligence [Text] / I. G. Perova // Eastern-European Journal of Enterprise Technologies. – 2014. – Vol. 1, Issue 4 (67). – P. 24–28. doi: 10.15587/1729-4061.2014.21202
26. Demin, D. A., Pelikh, V. F., Ponomarenko, O. I. (1998). Complex alloying of grey cast iron [Text] / D. A. Demin, V. F. Pelikh, O. I. Ponomarenko // Litejnoe Proizvodstvo, 10, 18–19.
27. Demin, D. A. Change in cast iron's chemical composition in inoculation with a Si-V-Mn master alloy [Text] / D. A. Demin // Litejnoe Proizvodstvo. – 1998. – Vol. 6. – P. 35.

UDC 004.891.2:550.8.052
DOI: 10.15587/1729-4061.2016.75679

INTERVAL FUZZY MODELING OF COMPLEX SYSTEMS UNDER CONDITIONS OF INPUT DATA UNCERTAINTY

N. Kondratenko
PhD, Associate Professor*
E-mail: kondrn@gmail.com
O. Snihur
Postgraduate student*
E-mail: olha.snihur@gmail.com
*Department of information security
Vinnytsia National Technical University
Khmelnytske highway, 95, Vinnytsia, Ukraine, 21021

Запропоновано систему моделей на основі інтервальної нечіткої логічної системи класифікації, що дозволяє отримати вихід за умов відсутності частини вхідних значень. Система будується виходячи з експериментальних даних, допускає залучення одного або кількох експертів, а також інтеграцію сторонніх підмоделей на основі інших методів та технологій прийняття рішень

Ключові слова: нечітка логічна система, інтервальні нечіткі множини, недовизначеність, кластерний аналіз

Предложена система моделей на основе интервальной нечеткой логической системы классификации, позволяющая получить выход при условии частичного отсутствия входных значений. Система строится исходя из экспериментальных данных, допускает привлечение одного или более экспертов, а также интеграцию сторонних подмоделей на основе других методов и технологий принятия решений

Ключевые слова: нечеткая логическая система, интервальные нечеткие множества, недоопределенность, кластерный анализ

1. Introduction

Input data uncertainty is one of the key factors in complex natural systems modeling. These include ecological, social, economic, technical systems of various nature. Constructing a single analytic expression that would mathematically describe such a system is a highly complicated task, and it is only possible to make assumptions about the way the system operates based on an experimental data set.

In [1] a number of UN-factors are described, that have a defining impact on the experimental data set quality, including measurements imprecision, lack of conditions for direct observations of the object, incompleteness and ambiguity of the knowledge related to the subject area and the task at hand, unaccounted for (hidden) parameters impact, lack of expert knowledge about the subject area or inability to formalize them, as well as uncertainty caused by input feature space dimensionality (redundancy and noise) [1, 2].

All these factors are inherent in natural systems and processes in one way or another. As an example of modeling a system of this class, later in this paper we show how the condition of an artesian well can be evaluated at any given time from the beginning of hydrogeological exploration up to its full completion. This task is characterized by difficulties in accessing experimental data, since obtaining input data necessary for operation of any given model requires significant effort. It is therefore to be expected, that geological exploration which precedes putting an artesian well into operation lasts ranging from 6 months and up to several years [3].