ISSN 0965-5425, Computational Mathematics and Mathematical Physics, 2009, Vol. 49, No. 7, pp. 1264–1275. © Pleiades Publishing, Ltd., 2009.
Original Russian Text © Yu.V. Maksimov, 2009, published in Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 2009, Vol. 49, No. 7, pp. 1327–1339.

Correct Algebras over Estimation Algorithms in the Set of Regular Recognition Problems with Nonoverlapping Classes

Yu. V. Maksimov

Moscow Institute of Physics and Technology, Institutskii per. 9, Dolgoprudnyi, Moscow oblast, 141700 Russia
e-mail: yuramaksimov@gmail.com

Received May 23, 2008; in final form, December 16, 2008

Abstract—Algebras over estimation algorithms in the set of regular problems with nonoverlapping classes are considered. A correctness criterion for the arbitrary degree algebraic closure of the model of estimation algorithms in the classification problems of this type is proposed; this criterion can be efficiently verified. An estimate of the minimal degree of the algebraic closure that is sufficient for constructing a correct classifier in an arbitrary regular problem with nonoverlapping classes is found.

DOI: 10.1134/S0965542509070173

Key words: classification problems, algebraic approach, estimation algorithms, correct algorithms.

INTRODUCTION

In the framework of the algebraic approach to pattern recognition proposed by Zhuravlev, it was shown that simple algebraic operations over the algorithms based on incorrect (heuristic) algorithms make it possible to construct correct algorithmic compositions. The base algorithms can be described in the model of estimation algorithms proposed by Zhuravlev. For the case of the problems with nonoverlapping classes, we study the question concerning the sufficiency of the compositions that can be obtained by an arbitrary degree algebraic closure of the model of estimation algorithms (EAs) for the correct classification of a test sample for a given set of training objects. The case of the linear closure of the EA model is considered separately. A sharp estimate of the degree of the algebraic closure that is sufficient for the construction of a correct classifier in an arbitrary regular problem is obtained.

1. STATEMENT OF THE PROBLEM: BASIC DEFINITIONS AND REPRESENTATION OF RECOGNIZING OPERATORS

Consider the pattern recognition problem with two nonoverlapping classes in the classical statement [1, 2]. The set of admissible objects is a union of nonempty nonoverlapping classes

$$M = \bigcup_{i=1}^{l} K_i \subseteq M_1 \times \ldots \times M_n, \quad K_i \cap K_j = \varnothing, \tag{1}$$

where M_i is the set of values of feature i equipped with the semimetric ρ_i. Every admissible object is defined by its standard feature description $I(S) = (a_1(S), ..., a_n(S))$, where $a_i \in M_i$, $i \in \{1, 2, ..., n\}$. Each admissible object S is assigned a genuine Boolean information vector $\alpha(S) = (\alpha_1(S), ..., \alpha_l(S))$, where $\alpha_i(S)$ are the values of the predicate $S \in K_i$ for $i \in \{1, 2, ..., l\}$. The genuine standard information is defined as the collection of the sets $(I(S_1), ..., I(S_m))$ and $(\alpha(S_1), ..., \alpha(S_m))$ of the complete standard descriptions of the training objects and their genuine information vectors.

The pattern recognition problem is to construct an algorithm A such that

$$A(I_0(K_1, \ldots, K_l), I(S)) = \alpha^A(S) \equiv (\alpha_1^A(S), \ldots, \alpha_l^A(S)), \quad \text{where} \quad \alpha^A(S) = \alpha(S), \quad S \in M,$$

for the genuine standard information $I_0(K_1, ..., K_l)$.

The algorithms in the EA model we want to examine are a superposition of a recognizing operator and a decision rule. The recognizing operator calculates estimates of the distances from the objects to the classes, and the decision rule classifies the objects on the basis of these estimates.

The recognizing operator B in the EA model calculates an estimate of the distance from the object i to the class j using the following formula proposed in [2]:

$$\Gamma_{ij}[B] = \frac{x_1}{N_1(j)} \sum_{\Omega \in \Omega_A} \sum_{S^t \in \tilde{S}^m \cap K_j} \omega^t \omega(\Omega) B_\Omega^{\tilde{\varepsilon}}(S^t, S_i) + \frac{x_0}{N_0(j)} \sum_{\Omega \in \Omega_A} \sum_{r \neq j} \sum_{S^t \in \tilde{S}^m \cap K_r} \omega^t \omega(\Omega)(1 - B_\Omega^{\tilde{\varepsilon}}(S^t, S_i)),$$

here, $x_0, x_1 \in \{0, 1\}$; $N_0(j)$ and $N_1(j)$ are normalizing coefficients; $\{S^1, \ldots, S^m\} = \tilde{S}^m$ is the set of training objects; $\{S_1, \ldots, S_q\} = \tilde{S}_q$ is the test sample; $K_1, \ldots, K_l$ are the classes of objects defined in (1); Ω_A is the system of support sets (the set of subsets of $\{1, 2, \ldots, n\}$); $\omega(\Omega) \in \mathbb{Q}^+$ is the weight of the support set formed by the weights of the features belonging to Ω; ω^t are the weights of the admissible objects; and $B_\Omega^{\tilde{\varepsilon}}(S^t, S_i)$ is the binary proximity function of which the value is determined on the basis of the analysis of the similarity between the descriptions of the objects S^t and S_i.

The proximity function (see [2] for details) is determined by the parameters $\tilde{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$, where $\varepsilon_i \in \mathbb{R}^+$ for any $i \in \{1, 2, \ldots, n\}$, which indicate the accuracy of measuring the features; in the most general case, it also depends on the integer nonnegative parameters q_1 and q_2:

$$B_\Omega^{\tilde{\varepsilon}, q_1, q_2}(S^t, S) = 1 \Leftrightarrow \begin{cases} \left|\{i \in \Omega : \rho_i(a_i(S^t), a_i(S)) \leq \varepsilon_i\}\right| \geq q_1, \\ \left|\{i \in \Omega : \rho_i(a_i(S^t), a_i(S)) > \varepsilon_i\}\right| \leq q_2. \end{cases} \quad (2)$$

It was proved in [1] that an arbitrary proximity function defined by Eq. (2) is a linear combination of binary proximity functions

$$B_{\{1, 2, \ldots, n\}}^{\tilde{\eta}, n, 0} = 1 \Leftrightarrow \left|\{i \in \{1, \ldots, n\} : \rho_i(a_i, b_i) \leq \eta_i\}\right| = n, \quad (3)$$

when the parameters $\tilde{\eta} = \tilde{\eta}(\tilde{\varepsilon}, \Omega)$ are chosen appropriately. The functions defined by Eq. (3) are denoted by $B[\tilde{\eta}] : B[\tilde{\eta}] = B_{\{1, 2, \ldots, n\}}^{\tilde{\eta}, n, 0}$. Therefore, in the analysis of the closures of the EA model, one can consider only the proximity functions of this type.

Consider the function

$$\Delta_\delta[\tilde{\varepsilon}] = \sum_{(\sigma_1, \ldots, \sigma_n) \in \{-1, +1\}^n} B[(\varepsilon_1 + \sigma_1\delta_1, \ldots, \varepsilon_n + \sigma_n\delta_n)] \prod_{i=1}^n \sigma_i. \quad (4)$$

It was proved in [1] that, for an arbitrary $\tilde{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$ and an arbitrary finite set of admissible objects, a small vector $\tilde{\delta}'$ ' with positive components can be found such that, for every $\delta < \delta'$, every term on the right-hand side of (4) is independent of δ. Therefore, the function $\Delta_\delta[\tilde{\varepsilon}]$ is also independent of small values of δ. It is called a labeling function and is denoted by $\Delta[\tilde{\varepsilon}]$:

$$\Delta[\tilde{\varepsilon}] = \sum_{(\sigma_1, \ldots, \sigma_n) \in \{-1, +1\}^n} B[(\varepsilon_1 + \sigma_1\delta_1, \ldots, \varepsilon_n + \sigma_n\delta_n)] \prod_{i=1}^n \sigma_i.$$

Let $E_t(\tilde{\varepsilon}) = \{S \in M : \tilde{\rho}(S^t, S) = \tilde{\varepsilon}\}$. In the same book [1], the following propositions are proved.

Theorem 1. *For any finite set of admissible objects $\tilde{S}_q$, there exists a real (rational) vector $(\delta_1, \ldots, \delta_n)$ $(\delta_i > 0$, $i \in \{1, 2, \ldots, n\})$ such that, for any object, we have $S \in \tilde{S}_q : \Delta[\tilde{\varepsilon}](S^t, S) = 1 \Leftrightarrow S \in E_t(\tilde{\varepsilon})$.*

Theorem 2. *Let M be a finite set. Then,*

$$B[(\varepsilon_1, \ldots, \varepsilon_n)](S^t, \cdot) = \sum_{\tilde{\rho} \in P} \Delta[\tilde{\rho}](S^t, \cdot),$$

where $P = \{\tilde{\rho}(S^t, S) : S \in M, \rho_i(a_i(S^t), a_i(S)) \leq \varepsilon_i \ \forall i \in \{1, 2, \ldots, n\}\}$.

Define the set of labeling operators $D^* = \{D_{t,d}[\tilde{\varepsilon}]\}_{t,\tilde{\varepsilon},\tilde{N}_d,d}$ as follows (the correctness of this definition was proved in [1]):

$$\Gamma_{ij}(D_{t,d}[\tilde{\varepsilon}]) = \frac{1}{N_d(j)}\begin{cases}\Delta[\tilde{\varepsilon}](S^t, S_i), & S^t \in K_j^d \\ 0, & S^t \in K_j^{1-d},\end{cases} \tag{5}$$

$$\Gamma(D_{t,d}[\tilde{\varepsilon}]) = \Delta^t[\tilde{\varepsilon}] \otimes (\tilde{\alpha}(S^t)^d \circ \tilde{N}_d) \in \mathbb{Q}^{q\times l}. \tag{6}$$

By the product of two vectors $\alpha \otimes \beta = \alpha^{\mathrm{T}} \otimes \beta$, where $\alpha = (\alpha_1, \ldots, \alpha_l)^{\mathrm{T}}$, and $\beta = (\beta_1, \ldots, \beta_q)^{\mathrm{T}}$, we mean the matrix $A = \|a_{ij}\| \in \mathbb{R}^{ql}: a_{ij} = \alpha_i\beta_j$. By the product of two matrices $A \circ B$, where $A = \|a_{ij}\| \in \mathbb{R}^{ql}$ and $B = \|b_{ij}\| \in \mathbb{R}^{ql}$, we mean the matrix $C = \|c_{ij}\| \in \mathbb{R}^{ql}: c_{ij} = a_{ij}b_{ij}$. The properties of these operations are thoroughly considered in [1]. In this paper, the properties of the operations $\otimes$ and $\circ$ are used without explanations.

In formulas (5) and (6), we use the following notation:

$$K_j^0 = S^m \backslash K_j, \quad K_j^1 = K_j, \quad \tilde{N}_d = \left(\frac{1}{N_d(1)}, \ldots, \frac{1}{N_d(l)}\right)$$

$$\text{for} \quad d \in \{0, 1\}, \quad \Delta^t[\tilde{\varepsilon}] = (\Delta[\tilde{\varepsilon}](S^t, S_1)\ldots\Delta[\tilde{\varepsilon}](S^t, S_q))^{\mathrm{T}}.$$

Let $P(t) = \{\tilde{\rho}(S^t, S_i)\}_{i=1}^q = \{\tilde{\rho}_1^t, \ldots, \tilde{\rho}_{p(t)}^t\}$ and $|P(t)| = p(t)$. Assume that $p = \sum_{t=1}^m p(t)$. From the binary vectors $\Delta^t[\tilde{\varepsilon}]$, we compose the matrices

$$\Theta^t = [\Delta^t[\tilde{\rho}_1^t]\ldots\Delta^t[\tilde{\rho}_{p(t)}^t]] \in \mathbb{Q}^{q\times p(t)}, \tag{7}$$

$$\Theta = [\Theta^1\ldots\Theta^m] \in \mathbb{Q}^{q\times p}. \tag{8}$$

Theorem 3. *The set of the estimation matrices for the operators in $L(B^*)$ coincides with the set of estimation matrices in $L(D^*)$; that is, it coincides with the set of linear combinations of the matrices $\tilde{\theta} \otimes (\tilde{\alpha}(S^t)^d \circ \tilde{N}_d)$ in which $\tilde{\theta}$ is a column in the matrix Θ^t defined by Eq.* (7); *here, $t \in \{1, 2, \ldots, m\}$ and $d \in \{0, 1\}$.*

Below, we only consider the operators with fixed weight parameters $\omega(\Omega)$, ω^t, and normalizations $\tilde{N}_d(j)$. To simplify certain implications, we assume that all these parameters are equal to unity. However, all the results proved below remain valid for arbitrary fixed weights and normalizations.

Definition 1. The model of recognizing operators R^* is said to be *correct* for the recognition problem $Z = (I_0, \tilde{S}^q)$ if

$$\forall\Gamma \in \mathbb{Q}^{q\times l}\exists B \in R^*: \Gamma[B] = \Gamma.$$

Definition 2. The recognition problem $\tilde{Z}(I_0, \tilde{S}_q)$ is said to be *regular* if the following conditions are satisfied:

$$\tilde{S}^m \cap \tilde{S}_q = \varnothing, \tag{9}$$

$$|\{K_1, \ldots, K_l\}| = l, \tag{10}$$

$$|\{(\tilde{\rho}(S^1, S_i), \ldots, \tilde{\rho}(S^m, S_i))\}_{i=1}^q| = q. \tag{11}$$

Below, we only consider regular recognition problems when the linear and algebraic closures of the EA model of an arbitrary degree are examined. Note that, due to (1), condition (10) is automatically fulfilled for the problems with nonempty nonoverlapping classes.

2. LINEAR CLOSURE OF THE MODEL

Consider the linear closure of the EA model. Because this is the simplest closure, it is interesting to find out if it is sufficient for constructing a correct algorithm for the given set of training and test objects. This question was considered in many studies including [1–5]. However, the correctness conditions obtained in those studies only reduce the problem to checking other, maybe more descriptive conditions. Unfortunately, such conditions cannot be checked efficiently in most practical problems.

The main result obtained in this paper is the correctness criterion for the linear closure of the EA model that can be checked in polynomial time.

Let us define some concepts that are convenient for the further presentation.

Definition 3. The set of columns $\Delta[\tilde{\varepsilon}](S^t, S_i)$ of the matrix Θ defined by Eq. (8) for which $S^t \in K_j$ is called the representation *with respect to class* K_j. The representation with respect to class K_j is denoted by Θ^{K_j}.

Where there is no source of confusion, we will interpret the representation as the matrix consisting of the corresponding columns. The linear hull of a representation is defined as the linear closure of the corresponding set of vectors.

Definition 4. A representation is said to be *complete* if the vectors included in the representation form a basis.

Definition 5. The union of the vectors included in the representations with respect to the classes $K_1, \ldots, K_t$ is called the representations with respect to the set of the classes $K_1, \ldots, K_t$. It is denoted by $\Theta^{K_1} \cup \ldots \cup \Theta^{K_t}$.

Theorem 4. *The linear closure of the EA model is correct if and only if the condition*

$$\forall i \neq j : rank(\Theta^{K_i} \cup \Theta^{K_j}) = q, \quad i, j = \overline{1, l},$$

is fulfilled in the regular classification problem with nonoverlapping classes.

Proof. *Necessity.* Consider the representations with respect to the arbitrary classes K_{i_0}, K_{j_0}. For definiteness, assume that $j_0 > i_0$. Assume that the linear closure of the EA model is correct and the condition of the theorem is not satisfied; that is, assume that

$$rank\left(\Theta^{K_{i_0}} \cup \Theta^{K_{j_0}}\right) < q \Rightarrow \exists \hat{\mathbf{l}} = (\hat{l}_1, \ldots, \hat{l}_q) \in \mathbb{Q}^q \forall \theta \in \Theta^{K_i} \cup \Theta^{K_j} : (\hat{\mathbf{l}}, \theta) = 0, \quad \hat{\mathbf{l}} \neq 0. \tag{12}$$

Consider the following function of the vector $\hat{l}$ and of the columns i_0 and j_0 of the matrix Γ:

$$f(i_0, j_0, (\hat{l}_1, \ldots, \hat{l}_q), \Gamma) = \sum_{k=1}^{q} (\Gamma_{ki_0} - \Gamma_{kj_0}, \hat{\mathbf{l}}_k).$$

Under the assumptions of the theorem, any estimation matrix can be represented as a linear combination of the matrices $\theta_{jt} \otimes e_j^d$:

$$\forall \Gamma \exists \{c_{jtd}\} : \Gamma = \sum_{j=0}^{l} \sum_{\theta_{jt} \in \Theta^{K_j}} \sum_{d \in \{0,1\}} c_{jtd} (\theta_{jt} \otimes e_j^d)$$

$$= \sum_{j \in \{i_0, j_0\}} \sum_{\theta_{jt} \in \Theta^{K_j}} \sum_{d \in \{0,1\}} c_{jtd} (\theta_{jt} \otimes e_j^d) + \sum_{j \in \{1, 2, \ldots, l\} \setminus \{i_0, j_0\}} \sum_{\theta_{jt} \in \Theta^{K_j}} \sum_{d \in \{0,1\}} c_{jtd} (\theta_{jt} \otimes e_j^d),$$

$$\Gamma_1 = \sum_{j \in \{i_0, j_0\}} \sum_{\theta_{jt} \in \Theta^{K_j}} \sum_{d \in \{0,1\}} c_{jtd} (\theta_{jt} \otimes e_j^d), \quad \Gamma_2 = \sum_{j \in \{1, 2, \ldots, l\} \setminus \{i_0, j_0\}} \sum_{\theta_{jt} \in \Theta^{K_j}} \sum_{d \in \{0,1\}} c_{jtd} (\theta_{jt} \otimes e_j^d).$$

Note that the columns i_0 and j_0 in the matrix Γ_2 are identical. Therefore, $f(i_0, j_0, (\hat{l}_1, \ldots, \hat{l}_q), \Gamma_2) = 0$. Now, consider the terms of the sum constituting the matrix Γ_1. Every column of the matrix $\theta \otimes \alpha$, where α is an arbitrary binary vector, either coincides with θ or is zero; therefore, condition (12) implies that $\forall \theta' \in \mathrm{col}(\theta' \otimes \alpha) : (\hat{\mathbf{l}}, \theta') = 0$, where $\theta \in \Theta^{K_i} \cup \Theta^{K_j}$ and $\alpha \in \{e_{i_0}, e_{j_0}, \bar{e}_{i_0}, \bar{e}_{j_0}\}$. (For an arbitrary matrix M, $\mathrm{col}(M)$ denotes the set of its columns.) Therefore, the function $f(i_0, j_0, (\hat{l}_1, \ldots, \hat{l}_q), \Gamma_1) = 0$ due to the linearity in the last argument. Hence, we have the following equality for the matrix Γ:

$$\forall \Gamma : f(i_0, j_0, (\hat{l}_1, \ldots, \hat{l}_q), \Gamma) = f(i_0, j_0, (\hat{l}_1, \ldots, \hat{l}_q), \Gamma_1) + f(i_0, j_0, (\hat{l}_1, \ldots, \hat{l}_q), \Gamma_2) = 0.$$

For Γ, we take an arbitrary matrix that contains the only unity in the column i_0 in the first row i for which $\hat{l}_i \neq 0$ (the vector $\hat{\mathbf{l}}$ is nonzero by condition). Then, we arrive at the contradiction.

Sufficiency. **Case 1.** Assume that all the representations Θ^{K_i}, $i \in \{1, 2, ..., l\}$ have a non-full rank. Denote by $\{l^{K_i}\}$ the basis of the orthogonal complement of the subspace spanned by the vectors included in Θ^{K_i}:

$$\forall \mathbf{l} \in \{l^{K_i}\} \quad \forall \theta \in \Theta^{K_i} : (\mathbf{l}, \theta) = 0, \quad \left|\{l^{K_i}\}\right| = \operatorname{rank}(\{l^{K_i}\}) = q - \operatorname{rank}(\Theta^{K_i}). \tag{13}$$

The representation Θ^{K_i} is of no full rank; therefore, $\operatorname{rank}(\{l^{K_i}\}) \geq 1$ and, in addition, for the linear hulls spanned by the corresponding sets of vectors, we have $L(\Theta^{K_i}) \oplus L(\{l^{K_i}\}) = \mathbb{Q}^q$ for $(i \in \{1, 2, ..., l\}$. We show that any vector that is orthogonal to all the vectors in the representation Θ^{K_i} is a linear combination of the vectors of the representation with respect to the class Θ^{K_j} $\forall j \neq i$. By the assumptions of the theorem, we have

$$\forall i = 1, 2, \ldots, l : \operatorname{rank}(\Theta^{K_1} \cup \Theta^{K_i}) = \ldots = \operatorname{rank}(\Theta^{K_l} \cup \Theta^{K_i}) = q;$$

that is,

$$L(\Theta^{K_1}) + L(\Theta^{K_i}) = \ldots = L(\Theta^{K_l}) + L(\Theta^{K_i}) = L(\Theta^{K_i}) \oplus L(\{l^{K_i}\}) = \mathbb{Q}^q,$$

From this, we obtain

$$\forall j \neq i : L(\{l^{K_i}\}) \subseteq L(\Theta^{K_j}) \Rightarrow L(\{l^{K_i}\}) \subseteq \bigcap_{j \neq i} L(\Theta^{K_j}) \Rightarrow \forall \mathbf{v} \in L(\{l^{K_i}\}) : \mathbf{v} \in \bigcap_{j \neq i} L(\Theta^{K_j}). \tag{14}$$

Now, we describe a method for constructing an algorithm producing an arbitrary matrix of estimates $\Gamma = \|\gamma_{ij}\| \in \mathbb{Q}^{q \times l}$. The proof consists of four steps:

Step 1. Consider the basis $\{l^{K_1}\}$ defined by Eq. (13). For the matrix Γ, calculate the products of the vector $\tilde{\gamma}_1 = \{\tilde{\Gamma}_{11}, ..., \tilde{\Gamma}_{q1}$ by each of the vectors of the basis $\{l^{K_1}\}$:

$$(\tilde{\gamma}_1, l_1^{K_1}) = r_1^{K_1},$$
$$(\tilde{\gamma}_1, l_2^{K_1}) = r_2^{K_1},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\left(\tilde{\gamma}_1, l^{K_1}_{\operatorname{rank}\left(\left\{l^{K_1}\right\}\right)}\right) = r^{K_1}_{\operatorname{rank}\left(\left\{l^{K_1}\right\}\right)}.$$

Assume that at least one of these products is distinct from zero. Otherwise, if all these products are zero, the vector $\tilde{\gamma}_1$ belongs to the space $L(\Theta^{K_1})$, and steps (2) and (3) of the proof may be omitted.

Step 2. It follows from (14) that

$$\exists \tilde{\mathbf{v}} \in \bigcap_{j \neq 1} L(\Theta^{K_j}) : \forall i \in \{1, \ldots, \operatorname{rank}(\{l^{K_1}\})\} : (\tilde{\mathbf{v}}, l_i^{K_1}) = r_i^{K_1};$$

moreover,

$$\forall j \neq 1 \exists \{\mathbf{c}_{jt}\} : \tilde{\mathbf{v}} = \sum_t c_{jt}\theta_{jt}, \quad \text{where} \quad \theta_{jt} \in \Theta^{K_j};$$

Step 3. Using the same notation as in the preceding step, we consider the matrix

$$\Gamma^1 = \sum_t c_{kt}(\theta_{kt} \otimes \mathbf{e}_k) + \sum_t c_{kt}(\theta_{kt} \otimes \bar{\mathbf{e}}_k) - \sum_{j \neq 1} \sum_t c_{jt}(\theta_{jt} \otimes \mathbf{e}_j),$$

where k is an arbitrary class index distinct from unity. Note that all the columns in Γ^1, except for the first one, consist of zeros, and the first column equals the vector $\tilde{\mathbf{v}} = (\tilde{v}_1, \ldots, \tilde{v}_q)$ determined at the preceding step.

Step 4. Repeating the reasoning used at steps 1–3 for the other classes, we obtain the set of matrices $\Gamma^1, \ldots, \Gamma^l$ such that only the ith column in Γ^i, $i \in \{1, 2, \ldots, l\}$ is distinct from zero, and it holds that

$$\forall i \in \{1, \ldots, l\} \forall j \in \{1, \ldots, \operatorname{rank}(\{l^{K_i}\})\} : (\tilde{\gamma}_i^i, \mathbf{l}_j^{K_i}) = (\tilde{\gamma}_i, \mathbf{l}_j^{K_i}). \tag{15}$$

Consider the matrix $\Gamma^* = \Gamma - \sum_{i=1}^{l} \Gamma^i$, and denote by $\tilde{\gamma}_i^*$ the ith column of Γ^*. Due to (15), we have

$$\forall i \in \{1, \ldots, l\} \forall j \in \{1, \ldots, \operatorname{rank}(\{l^{K_i}\})\} : (\tilde{\gamma}_i^*, \mathbf{l}_j^{K_i}) = 0.$$

Therefore, $\forall i \in \{1, 2, \ldots, l\}$, we have $\tilde{\gamma}_i^* \in L(\Theta^{K_i})$. Hence,

$$\forall i \in \{1, 2, \ldots, l\} \exists \{b_{it}\} : \tilde{\gamma}_i^* = \sum_{t=0}^{t_i} b_{it}\theta_{it}, \quad \text{where} \quad \theta_{it} \in \Theta^{K_i}. \tag{16}$$

Consider the matrices Γ^{*i} defined by Eq. (16):

$$\Gamma^{*i} = \tilde{\gamma}_i^* \otimes \mathbf{e}_i = \sum_{t=0}^{t_i} b_{it}(\theta_{it} \otimes \mathbf{e}_i), \quad i \in \{1, 2, \ldots, l\}.$$

Note that each of the matrices $\{\Gamma^{*i}\}_{i=0}^{l}$ and $\{\Gamma^i\}_{i=0}^{l}$ can be constructed in the linear closure of the EA model; therefore, the matrices

$$\Gamma^* = \sum_{i=0}^{l} \Gamma^{*i}, \quad \Gamma = \Gamma^* + \sum_{i=0}^{l} \Gamma^i = \sum_{i=0}^{l} \Gamma^i + \sum_{i=0}^{l} \Gamma^{*i},$$

being the sums of matrices belonging to the linear closure, can be constructed in this closure. Because Γ was chosen arbitrarily, case 1 is completely proved.

Case 2. Some of the representations are of full rank. Let the representations with respect to the classes $K_{i_0}, \ldots, K_{i_t}$ have the full rank. It is clear that, in the linear closure of the EA model, an arbitrary classification that does not change the entries of the estimation matrix in the other columns can be realized in the columns $i_0, \ldots, i_t$. Therefore, case 2 is naturally reduced to case 1 for the problem with a smaller number of classes. It remains to note that the validity of the theorem is obvious for the case of the problem with two nonoverlapping classes and for the case when all the representations have the full rank. This completes the proof of Theorem 4.

Remark 1. Note that a necessary condition for correctness is the full rank of the matrix Θ defined by (8). In the case of two classes, this condition is also sufficient.

Remark 2. The proposed criterion can be efficiently checked in a time not exceeding $O(mq^2l^2)$ in the worst case. Here, m is the length of the training sample, q is the amount of test data, and l is the number of classes. Indeed, the time needed to construct Θ defined by (8) is of order $O(qm)$, and the check of the criterion for the given pair of classes (i, j) does not exceed $O(mq^2)$. Since the number of pairs (i, j) is $O(l^2)$, the total time does not exceed $O(qm) + O(l^2)O(mq^2)$.

We illustrate the use of the proposed criterion by the following simple example.

Example 1. Consider the following problem defined by the set of training objects $\tilde{S}_m$ and the set of test objects $\tilde{S}_q$

$\tilde{S}_m$:

	ω_1	ω_2	ω_3
S^1	0	1	1
S^2	2	2	1
S^3	0	0	0

$\tilde{S}^q$:

	ω_1	ω_2	ω_3
S_1	1	2	3
S_2	3	0	3
S_3	3	2	3
S_4	0	1	2
S_5	0	1	0

Furthermore, let the object S_1 belong to the class K_1, S_2 to the class K_2, and S_3 to the class K_3. In addition, we assume that the metric in each feature space (the set of values of each feature is the set of integers) be defined by $\rho(x, y) = |x - y|$. Let us calculate the sets of pairwise distances $P(t)$, which define the matrix Θ by formulas (7) and (8): $P(1) = \{(1, 1, 2), (3, 1, 2), (0, 0, 1)\}$, $P(2) = \{(1, 0, 2), (1, 2, 2), (2, 1, 1)\}$ and $P(3) = (1, 2, 3), (3, 0, 3), (3, 2, 3), (0, 1, 2), (0, 1, 0)\}$.

Construct the matrix Θ, and find the representations with respect to each of the classes:

$$\Theta = \left[\begin{array}{ccc|ccc|ccccc} 1&0&0&1&0&0&1&0&0&0&0\\ 0&1&0&0&1&0&0&1&0&0&0\\ 0&1&0&1&0&0&0&0&1&0&0\\ 0&0&1&0&0&1&0&0&0&1&0\\ 0&0&1&0&0&1&0&0&0&0&1 \end{array}\right],$$

$$\Theta^{K_1} = \{(1, 0, 0, 0, 0)^{\mathrm{T}}, (0, 1, 1, 0, 0)^{\mathrm{T}}, (0, 0, 0, 1, 1)^{\mathrm{T}}\},$$

$$\Theta^{K_2} = \{(1, 0, 1, 0, 0)^{\mathrm{T}}, (0, 1, 0, 0, 0)^{\mathrm{T}}, (0, 0, 0, 1, 1)^{\mathrm{T}}\},$$

$$\Theta^{K_3} = \{(1, 0, 0, 0, 0)^{\mathrm{T}}, (0, 1, 0, 0, 0)^{\mathrm{T}}, (0, 0, 1, 0, 0)^{\mathrm{T}}, (0, 0, 0, 1, 0)^{\mathrm{T}}, (0, 0, 0, 0, 1)^{\mathrm{T}}\}.$$

Consider the representation with respect to the classes K_1 and K_2. Note that the vector $l = (0, 0, 0, 1, -1)$ is orthogonal to all the vectors both in the representation with respect to the class K_1 and with respect to K_2. Therefore, $(\mathrm{rank}(\Theta^{K_1} \cup \Theta^{K_2}) < q$; by the correctness criterion proved above, the linear closure of the EA model is incorrect. Note that the second-degree closure for this problem is correct as will be proved in the next section.

Assume that a new training object $S^4 = (2, 1, 0)$ such that $S^4 \in K_2$ is added. The set of the distances $P(4)$ defining the matrix Θ^4 by formula (7) is $P(4) = \{(1, 1, 3), (2, 0, 2), (2, 0, 0)\}$. The matrix Θ and the representation with respect to K_2 (the representations with respect to the other classes do not change) take the form

$$\Theta = \left[\begin{array}{ccc|ccc|ccccc|ccc} 1&0&0&1&0&0&1&0&0&0&0&1&0&0\\ 0&1&0&0&1&0&0&1&0&0&0&1&0&0\\ 0&1&0&1&0&0&0&0&1&0&0&1&0&0\\ 0&0&1&0&0&1&0&0&0&1&0&0&1&0\\ 0&0&1&0&0&1&0&0&0&0&1&0&0&1 \end{array}\right],$$

$$\Theta^{K_2} = \{(1, 0, 1, 0, 0)^{\mathrm{T}}, (0, 1, 0, 0, 0)^{\mathrm{T}}, (0, 0, 0, 1, 1)^{\mathrm{T}}, (1, 1, 1, 0, 0)^{\mathrm{T}}, (0, 0, 0, 1, 0)^{\mathrm{T}}, (0, 0, 0, 0, 1)^{\mathrm{T}}\}.$$

The equalities $\mathrm{rank}(\Theta^{K_1} \cup \Theta^{K_3}) = \mathrm{rank}(\Theta^{K_2} \cup \Theta^{K_3}) = 5$ are obvious. It remains to verify the equality $\mathrm{rank}(\Theta^{K_1} \cup \Theta^{K_2}) = 5$. The vectors $\{\mathbf{e}_2, \mathbf{e}_4, \mathbf{e}_5\}$ belong to Θ^{K_2}, and the vector $\{\mathbf{e}_1\}$ belongs to Θ^{K_1}. The linear hull of each of the representations contains the vector $\tilde{1} = (1, 1, 1, 1, 1)$; therefore, $\mathrm{rank}(\Theta^{K_1} \cup \Theta^{K_2}) = q = 5$, and the linear closure of the EA model is correct by Theorem 4.

3. ALGEBRAIC CLOSURE OF THE EA MODEL

The operation of the elementwise multiplication defined in Section 1 and applied to the matrices of the operator estimates in the EA model naturally induces operations on recognizing operators. If $B^* = (B_1, \ldots, B_r)$ is the set of recognizing operators in the EA model, then the set of operators $U^k(B^*)$ of the algebraic closure of the EA model of degree k is defined by

$$U^k(B^*) = L(\{B_1 \cdot \ldots \cdot B_t \,|\, B_1, \ldots, B_t \in B^*, 1 \le t \le k\}). \tag{17}$$

By the set of recognizing operators of the algebraic closure $U(B^*)$ of the EA model, we mean the union of the operators belonging to the algebraic closure of the EA model of an arbitrary degree n, where $n \in \mathbb{N}$:

$$U(B^*) = \bigcup_{n \in \mathbb{N}} U^n(B^*).$$

As we established in the preceding section, the linear closure of the EA model is incorrect for certain training and test samples, which opens the question of the possibility to construct a classification in the algebraic closures of the EA model. For the problems with nonoverlapping classes, we propose and prove a criterion for the correctness of the algebraic closure of the EA model that completely solves this problem.

Note the following consequence of the correctness criterion of the EA model obtained in [2].

Theorem 5. *The algebraic closure of the EA model in the regular problem* $Z = (I_0, \tilde{S}_q)$ *with nonoverlapping classes is correct.*

As was mentioned in Section 1, only regular problems are considered when the EA model and its closures are examined. Because the linear closures of the labeling operators $D^* = \{D_{t,d}\}$ and the operators B^* of the EA model are identical, formula (17) can be rewritten in the form

$$U^k(B^*) = U^k(D^*) = L(\{D_1 \cdot \ldots \cdot D_t \mid D_1, \ldots, D_t \in D^*, 1 \le t \le k\}).$$

In [1], the matrix $A = \begin{bmatrix} A^1 \\ \ldots \\ A^m \end{bmatrix}$ was constructed. Here, $A^t = [1...1]_{1 \times l}$, for $\alpha(S^t) \in \{(0, \ldots, 0)_{1 \times l}, (1, \ldots, 1)_{1 \times l}\}$; otherwise, $A^t = \begin{bmatrix} \tilde{\alpha}(S^t) \\ \overline{\tilde{\alpha}(S^t)} \end{bmatrix}_{2 \times l}$, where $\tilde{\alpha}(S^t) = (\alpha_1, \ldots, \alpha_l)$, and $\overline{\tilde{\alpha}(S^t)} = (1 - \alpha_1, \ldots, 1 - \alpha_l)$, $t \in \{1, 2, \ldots, m\}$).

Set

$$V = \bigcup_{t=1}^{m} \bigcup_{\theta \in \Theta^t} \bigcup_{\alpha \in A^t} (\theta \otimes \alpha), \tag{18}$$

where the matrix Θ^t is defined by (7). Note that, for the types of problems under examination, equality (18) can be rewritten in terms of the representations as

$$V = \bigcup_{t=1}^{l} \bigcup_{\theta \in \Theta^{K_t}} \bigcup_{\sigma \in \{0,1\}} (\theta \otimes \mathbf{e}_t^{\sigma}). \tag{19}$$

In [1], the following theorem was proved.

Theorem 6. *The set of matrices of operator estimates belonging to* $U^k(B^*)$ *is the set of linear combinations of the matrices* $(\tilde{\theta}_1 \circ \ldots \circ \tilde{\theta}_k) \otimes (\tilde{\alpha}_1 \circ \ldots \circ \tilde{\alpha}_k)$, *where* $(\tilde{\theta}_1^{\mathrm{T}}, \tilde{\alpha}_1), \ldots, (\tilde{\theta}_k^{\mathrm{T}}, \tilde{\alpha}_k) \in V$.

Note that $(\tilde{\theta}_1 \circ \ldots \circ \tilde{\theta}_k) \otimes (\tilde{\alpha}_1 \circ \ldots \circ \tilde{\alpha}_k) = \Gamma[D_1 \cdot \ldots \cdot D_k]$, where $\Gamma[D_i] = \theta_i \otimes \alpha_i \; \forall i \in \{1, 2, \ldots, k\}$.

Now, we give some definitions, and then formulate the correctness criterion for the algebraic closure of degree k in terms of these definitions.

Definition 6. 1-closure $[\Theta]$ of the representation Θ is defined as the union of the set of vectors in Θ and the vector of an appropriate length consisting of unities. The operation $[\cdot]$ of the 1-closure is the completion of a representation to its 1-closure.

Definition 7. A representation is said to be 1-closed (1-representation) if its 1-closure coincides with the representation itself.

Definition 8. The product of the representations (1-representations) $\Theta_1 \circ \Theta_2$ is defined if the lengths of the vectors constituting the representations are identical; in this case, the product is the set of vectors U defined as

$$U = \{\theta : \theta = \theta_1 \circ \theta_2, \theta_1 \in \Theta_1, \theta_2 \in \Theta_2\}. \tag{20}$$

Definition 9. The degree k ($k > 0$) of the representation (1-representation) Θ is the set of vectors U defined as

$$U = \{\theta : \theta = \theta_1 \circ \ldots \circ \theta_k, \theta_1 \in \Theta, \ldots, \theta_k \in \Theta\}. \tag{21}$$

It is denoted by Θ^k.

The following theorems immediately follows from Definitions 6–8.

Theorem 7. *If the representations* Θ_1 *and* Θ_2 *are* 1*-representations*, *then the set U defined by* (20) *is a* 1*-representation*.

Theorem 8. *If* Θ *is a 1-representation*, *then the set* Θ^k *defined by* (21) *also is a 1-representation*.

The union of the representation $\Theta_1 \cup \Theta_2$ is the representation with the set of vectors formed by the union of the sets of vectors Θ_1 and Θ_2 without taking into account the multiplicities of their occurrences.

Theorem 9. *For an arbitrary regular problem with nonoverlapping classes, the set of estimation matrices of the operators from* $U^k(B^*)$ *is the set of linear combinations of the matrices of the form* $(\tilde{\theta}_1 \circ \ldots \circ \tilde{\theta}_k) \otimes \tilde{\alpha}_t^{\sigma_t}$, *where* $(\tilde{\theta}_1^{\mathrm{T}}, \tilde{\alpha}_1), \ldots, (\tilde{\theta}_k^{\mathrm{T}}, \tilde{\alpha}_k) \in V$, $\sigma_t \in \{0, 1\}$, and $t \in \{1, 2, \ldots, k\}$.

Proof. It follows from the equality $\bar{\mathbf{e}}_{i_1} \circ \ldots \circ \bar{\mathbf{e}}_{i_k} = \bar{\mathbf{e}}_{i_1} \circ \ldots \circ \bar{\mathbf{e}}_{i_{k-1}} - \bar{\mathbf{e}}_{i_k} \circ \bar{\mathbf{e}}_{i_1} \circ \ldots \circ \bar{\mathbf{e}}_{i_{k-1}}$ that any product of the form $\mathbf{e}_{i_1}^{\sigma_{i_1}} \circ \ldots \circ \mathbf{e}_{i_k}^{\sigma_{i_k}}$ can be represented as a linear combination of the vectors $\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_k}, \bar{\mathbf{e}}_{i_1}, \ldots, \bar{\mathbf{e}}_{i_k}$. The application of Theorem 6 with regard for Eq. (19) completes the proof.

Theorem 10. *For an arbitrary regular classification problem* $Z = (I_0, \tilde{S}_q)$ *with nonoverlapping classes*, *the algebraic closure of the EA model of degree k is correct if and only if*

$$\forall i \neq j \quad \operatorname{rank}([\Theta^{K_i} \cup \Theta^{K_j}] \circ [\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{k-1}) = q.$$

Proof. Consider the set of vectors of the matrices defined in Theorem 9 of the form $\theta \times \mathbf{e}_i^{\sigma_i}$, where $\sigma_i \in \{0, 1\}$. Due to Theorem 9 and Eq. (19), the set of vectors $\{\theta\}$ forming these matrices in the algebraic closure of degree k is $\Theta^{K_i} \circ [\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{k-1}$.

Therefore, the problem is reduced to the question of correctness of the linear closure in the recognition problem $Z' = (I_0', \tilde{S}^q)$, which should be considered as a problem with nonoverlapping classes (due to Theorem 9). Note that, in the proof of the correctness of the linear closure on the basis of regularity conditions (9)–(11), only condition (10) was used. Moreover, any number of vectors in the representations with respect to different classes were allowed to coincide. Condition (10) is satisfied due to the regularity of the problem Z, and the representations of this problem are used to form the representations Θ'^{K_i} of the problem Z' by the formula $\Theta'^{K_i} = \Theta^{K_i} \circ [\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{k-1}$ $\forall i \in \{1, 2, \ldots, l\}$. Application of Theorem 4 to the problem Z' completes the proof.

Corollary 1. It follows from Theorem 10 that a necessary condition for the correctness of the algebraic closure of degree k is the equality $\operatorname{rank}([\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^k) = \operatorname{rank}([\Theta]^k) = q$, while $\operatorname{rank}([\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{k-1}) = \operatorname{rank}([\Theta]^{k-1}) = q$ is a sufficient condition. In particular, the algebraic closure of degree two is correct if $\operatorname{rank}([\Theta]) = \operatorname{rank}(\Theta) = q$.

Proof. We prove that the condition $\operatorname{rank}([\Theta]^{k-1}) = q$ is sufficient for the correctness of the algebraic closure of degree k. Indeed, by the definition of the 1-representation, the 1-representation $[\Theta^{K_i} \cup \Theta^{K_j}]$ contains a vector consisting of unities for any i and j; therefore, $\operatorname{rank}([\Theta^{K_i} \cup \Theta^{K_j}] \circ [\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{k-1}) \geq \operatorname{rank}([\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{k-1}) = q$, which implies the validity of the assertion.

Now, we prove the necessity of the condition $\operatorname{rank}([\Theta]^k) = q$ for the correctness of the algebraic closure of degree k. Note that $\forall i, j$, we have $\operatorname{rank}([\Theta]^k) = \operatorname{rank}([\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^k) \geq \operatorname{rank}([\Theta^{K_i} \cup \Theta^{K_j}] \circ [\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{k-1})$. Therefore, if the representation $\operatorname{rank}([\Theta]^k)$has no full rank, the algebraic closure of degree k is incorrect due to the correctness criterion.

Return to Example 1 considered in the preceding section; this example involves three training objects. It was shown that the linear closure of the EA model is incorrect in this example. However, it follows from Corollary 1 that the algebraic closure of degree two is correct.

4. ESTIMATE OF THE DEGREE OF THE CORRECT ALGORITHM IN THE EA MODEL

We left open the question of the degree of the polynomial that is sufficient for the construction of a correct algorithm for an arbitrary regular problem $Z = (I_0, \tilde{S}_q)$ with nonoverlapping classes. In the case of an arbitrary regular problem for l classes (without restrictions on the relationships between the classes), it was shown in [1] that the unimprovable degree of the polynomial is exactly $[\log_2 q] + [\log_2 l]$.

In this paper, we prove that, in the case of nonoverlapping classes, the degree of the correct polynomial does not exceed $I_0(l = 2)[\log_2 q] + I_0(l > 2)[\log_2 q]$ and, in addition, there exists a regular recognition problem in which the polynomial of a lower degree is insufficient for obtaining an arbitrary classification. Here, I_0 is the function defined by the equation

$$I_0(condition) = \begin{cases} 1, & \text{if the } condition \text{ is true,} \\ 0, & \text{if the } condition \text{ is false.} \end{cases}$$

Definition 10. The binary matrix $H_{q \times p}$ is said to be *mD-labelable* if it can be represented as a set of matrices $H^1_{q \times p_1}, \ldots, H^m_{q \times p_m}$: such that $H = [H^1 \ldots H^m]$, where each of the matrices H^i ($i \in \{1, 2, \ldots, m\}$) does not contain zero columns and each of its rows contains exactly one unity.

Note the following result obtained in [1], which plays an important role in the proof of Theorem 12.

Theorem 11. *For an arbitrary mD-labelable matrix H, there exists a regular recognition problem in which $\Theta = H$ (Θ is determined by Eq.* (8)).

In [1], the following lemma was proved.

Lemma 1. *Let all the rows in an mD-labelable matrix Θ be different. Then, any q-dimensional rational (real) vector can be represented by the linear combination*

$$\sum_{u \in X} c_u \tilde{\theta}_{u,1} \circ \ldots \circ \tilde{\theta}_{u, r(u)},$$

$\{\tilde{\theta}^{\mathrm{T}}_{u,z}\}^{r(u)}_{z=1,\, u \in X} \subseteq \mathrm{col}(\Theta)$, $1 \le r(u) \le [\log_2 q]$ for $u \in X$, where $|X| < \infty$. *The estimate for $r(u)$ is sharp.*

Consider the family of matrices $\{X_n\} : X_n = [x_1 \ldots x_n] \in \mathbb{Q}^{2^n \times n}$ $\forall n \in \mathbb{N}$ with the columns $x_1, \ldots, x_n$ in which every row is the binary representation of its index. The matrix

$$\Theta^*_n = [x_1 \bar{x}_1 \ldots x_n \bar{x}_n] \tag{22}$$

is *nD*-labelable; therefore, by Theorem 11, it is the estimation matrix for the labeling operators of a certain regular problem. It was proved in [1] that, in the problems in which the matrix Θ^*_n $\forall n \in \mathbb{N}$ is the matrix defined by Eq. (8), the estimate of Lemma 1 is sharp. The following result is a consequence of Lemma 1.

Lemma 2. *Let the binary matrix Θ with all the rows different be mD-labelable. Let $[\Theta]$ be the 1-closure of the representation consisting of the columns of Θ. Then, we have rank$([\Theta]^{[\log_2 q]}) = q$ and, in addition,* $\forall q \exists \Theta_q : \mathrm{rank}([\Theta_q]^{[\log_2 q] - 1}) < q$.

Theorem 12. *For an arbitrary regular problem with nonoverlapping classes, the algebraic closure $U^{I_0(l=2)[\log_2 q] + I_0(l>2)[\log_2 q]}$ of the EA model is correct, and this estimate is sharp.*

Proof. *Case 1.* Let the number of classes be two. The algebraic closure of degree $[\log_2 q]$ is correct by Theorem 10 if and only if

$$\mathrm{rank}([\Theta^{K_1} \cup \Theta^{K_2}] \circ [\Theta^{K_1} \cup \Theta^{K_2}]^{[\log_2 q] - 1}) = \mathrm{rank}([\Theta^{K_1} \cup \Theta^{K_2}]^{[\log_2 q]}) = q.$$

This equality holds true due to Lemma 2. The estimate is sharp due to the same lemma and Theorem 11.

Case 2. Let the number of classes be $l > 2$ and $\exists n \in \mathbb{N}$ such that $2^n + 1 \le q < 2^{n+1}$.

Sufficiency. In Case 2, we have $[\log_2 q] = [\log_2 q] + 1 = n + 1$. It follows from Lemma 2 that $\mathrm{rank}([\Theta]^{[\log_2 q]}) = n$ and, by Corollary 1, this equality is sufficient for the correctness of the closure of degree $n + 1$.

Necessity. Consider the case of three classes. Let $\Theta_n^* = (\theta_{ij}^n)_{i=1,j=1}^{i=2^n,j=2n}$ be the nD-labelable matrix defined by Eq. (22) for which the degree $r(u) = [\log_2 q] - 1 = n$ is insufficient for the construction of a basis in the space of columns (see Lemma 1). Consider the binary matrix $A = (a_{ij})_{i,j=1}^{i,j=2^n+1,2n+3}$, where

$$a_{ij} = \begin{cases} \theta_{ij}^n, & i \le 2^n, j \le 2n \\ 0, & i = 2^n + 1, j \le 2n, j \text{ is even}, \\ 1, & i = 2^n + 1, j \le 2n, j \text{ is odd}, \\ \delta_{i,2^n+1}, & j = 2n + 1 \\ 1 - \delta_{i,2^n+1}, & j = 2n + 2 \\ 1, & j = 2n + 3. \end{cases}$$

Observe that this matrix is $(n + 2)$-labelable and all its rows are distinct; therefore, due to Theorem 11, there exists a regular recognition problem in which A plays the role of the matrix Θ defined by Eq. (8). Let the objects $S^1, \ldots, S^n$ determining the first $2n$ columns of A belong to the class K_1; the object S^{n+1} determining the columns $2n + 1$ and $2n + 2$ belong to the class K_2; and the training object S^{n+2} determining the remaining submatrix belongs to the class K_3. Therefore, the representation Θ^{K_2} with respect to K_2 contains only two vectors $\mathbf{e}_{2^n+1}$ and $\bar{\mathbf{e}}_{2^n+1}$. The representation $\Theta^{K_3} = \{(1, \ldots, 1)\}$.

By Theorem 10, for the correctness of the algebraic closure of degree $[\log q]$, the following equality must hold:

$$\operatorname{rank}([\Theta^{K_2} \cup \Theta^{K_3}] \circ [\Theta^{K_1} \cup \Theta^{K_2} \cup \Theta^{K_3}]^{[\log q]-1}) = q.$$

It follows from Lemma 2 that the representation $[\Theta_n^*]^{[\log q]-1}$ is of non-full rank; therefore, there exists a nonzero vector $\tilde{\mathbf{l}} = (\tilde{l}_1, \ldots, \tilde{l}_q)$ that is orthogonal to all the vectors of this representation. This implies that the vector $\mathbf{l} = (\tilde{l}_1, \ldots, \tilde{l}_q, 0)$ is orthogonal to all the vectors of the representation $[\Theta^{K_1} \cup \Theta^{K_2} \cup \Theta^{K_3}]^{[\log q]-1}$.

Consider an arbitrary vector $\theta \in [\Theta^{K_1} \cup \Theta^{K_2} \cup \Theta^{K_3}]^{[\log q]-1}$ that has zero as its last component. It is clear that the product of the vectors $\theta \circ \theta_1$, where $\theta_1 \in \Theta^{K_2} \cup \Theta^{K_3}$, is either a zero vector or a vector that is identical to θ. Therefore, $(\theta \circ \theta_1, \mathbf{l}) = 0$. Moreover, $(\mathbf{e}_{2^n+1} \circ \theta_1, \mathbf{l}) = 0$ for $\theta_1 \in \Theta^{K_2} \cup \Theta^{K_3}$; this implies that $\mathbf{l}$ is orthogonal to any vector in $[\Theta^{K_2} \cup \Theta^{K_3}] \circ [\Theta^{K_1} \cup \Theta^{K_2} \cup \Theta^{K_3}]^{[\log q]-1}$, which completes the proof for Case 2. Note that this construct can be easily generalized for the case of an arbitrary number of nonoverlapping classes for which there exists a regular problem. Indeed, it is sufficient that the matrix A be a submatrix of the matrix $\Theta \in \mathbb{Q}^{q \times k}$, where $2^n < q < 2^{n+1}$, the classes K_2 and K_3 consist of a single object each, and the description of this object be induced by the description of the objects S^{n+1} and S^{n+2}, respectively.

Case 3. Let the number of clusters $l > 2$ and $\exists n \in \mathbb{N} \cup \{0\}$ such that $q = 2^n$.

In this case, we have $[\log_2 q] = [\log_2 2^n] = n$.

Necessity. The necessity of the condition under examination follows from Lemma 2 and the necessary condition for the correctness of the algebraic closure given in Corollary 1.

Sufficiency. Note that, if the number of training objects is 2^n ($n \in \mathbb{N}$) and the algebraic closure of degree $n - 1$ is incorrect, the rank of the subspace that is orthogonal to any vector in $[\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{[\log q]-1}$ is one. Indeed, removing one test object, we obtain that, for the sample of length $2^n - 1$, a basis in the closure of degree $[\log_2 2^n - 1] = n - 1$ can be obtained (by Lemma 2). Therefore, $\operatorname{rank}[\Theta^{K_1} \cup \ldots \cup \Theta^{K_l}]^{[\log q]-1} \ge q$. Note that, for any nonzero vector inducing an orthogonal subspace, none of its components is zero (indeed, if a component is zero, the corresponding test object can be arbitrarily classified).

Thus, if the algebraic closure of degree n is incorrect for the regular recognition problem under examination, the set of vectors that are orthogonal to every vector in the representation $[\Theta^{K_1} \cup ... \cup \Theta^{K_l}]^{n-1}$ forms a one-dimensional subspace and $\operatorname{rank}([\Theta^{K_1} \cup ... \cup \Theta^{K_l}]^{n-1}) = q - 1$.

If the closure of degree n is correct, then we have, for any $\forall i, j$: $\operatorname{rank}([\Theta^{K_i} \cup \Theta^{K_j}] \circ [\Theta^{K_1} \cup ... \cup \Theta^{K_l}]^{n-1}) = n$. Consider arbitrary classes K_{i_0} and K_{j_0}. The problem is regular; therefore, $\Theta^{K_{i_0}} \neq \Theta^{K_{j_0}}$. Consequently, $\Theta^{K_{i_0}} \cup \Theta^{K_{j_0}}$ contains a vector that is distinct from $\tilde{\mathbf{1}} = (1, ..., 1)$ and $0\ \tilde{\mathbf{0}} = (0, ..., 0)$. Denote it by $\tilde{\alpha}$. Without loss of generality, set $|\operatorname{Ind}(\tilde{\alpha})| \leq 2^{n-1}$ (otherwise, its negation can be used as $\tilde{\alpha}$). Rearrange the components of the vector $\tilde{\alpha}$ in such a way that the first $|\operatorname{Ind}(\tilde{\alpha})| \leq 2^{n-1}$ components contain unities and the other components are zero. Rearrange the rows in the matrix Θ in the same fashion. (The length of $\tilde{\alpha}$ is the same as the number of rows in the matrix Θ.) Below, Θ denotes the matrix with the rearranged rows. It is clear that the linear closure of the set of vectors of the representation $[\Theta^{K_1} \cup ... \cup \Theta^{K_l}]^{n-1}$ contains a vector $\tilde{\mathbf{e}} = (1, 0, ..., 0, x)$, where x is a number determined by the vector $\hat{\mathbf{l}}$. It remains to note that $(\tilde{\alpha} \circ \tilde{\mathbf{e}}, \tilde{\mathbf{l}}) = (\tilde{\mathbf{e}}_1, \tilde{\mathbf{l}}) \neq 0$; therefore, the representation $[\Theta^{K_{i_0}} \cup \Theta^{K_{j_0}}] \circ [\Theta^{K_1} \cup ... \cup \Theta^{K_l}]^{[\log q]-1}\ \forall i_0, j_0 \in \{1, ..., l\}$ has a greater rank than the representation $[\Theta^{K_1} \cup ... \cup \Theta^{K_l}]^{[\log q]-1}$; that is, it is complete, which completes the proof.

5. CONCLUSIONS

The correctness criteria for the linear closure and the algebraic closure of an arbitrary degree of the family of EA algorithms with arbitrary fixed weights of the feasible objects, weights of the features and normalizations (defined in [1]) are obtained. These criteria make it possible to efficiently find the minimal sufficient degree of the algorithmic polynomial that can correctly classify the given test sample. In addition, it is shown that the degree of the closure that makes it possible to construct a correct classifier for an arbitrary regular problem with nonoverlapping classes can be considerably reduced compared with the general case (which is described, for example, in [1]).

REFERENCES

1. A. G. D'yakonov, *Algebras over Estimation Algorithms* (Mosk. Gos. Univ., Moscow, 2006) [in Russian].
2. Yu. I. Zhuravlev, "An Algebraic Approach to Recognition and Classification Problems," in *Problems of Cybernetics* issue 33 (Nauka, Moscow, 1978; Hafner, 1986), pp. 5–68.
3. V. L. Matrosov, *Correct Algebras of Bounded Capacity over the Set of Estimation Algorithms*, Doctoral Dissertation in Mathematics and Physics (Computing Center, Russian Academy of Sciences, Moscow, 1986).
4. K. V. Rudakov, *On the Algebraic Theory of Universal and Local Constraints for Classification Problems*, Doctoral Dissertation in Mathematics and Physics (Computing Center, Russian Academy of Sciences, Moscow, 1992).
5. A. G. D'yakonov, "Algebra over Estimation Algorithms: Normalization and Division," Zh. Vychisl. Mat. Mat. Fiz. **47** 1099–1109 (2007) [Comput. Math. Math. Phys. **47**, 1050–1060 (2007)].