A Spinor-Based Approach to Constrained Optimization in the Design of Correct Estimation Algorithms

A. Yu. Smetanin
Economical forecasting group
Aruji asset management LLC
Moscow, Russia
e-mail: alexander.smetanin@gmail.com

Abstract[1]

The problem of building correct algorithms of pattern recognition is considered. For some classes of estimation algorithms, criteria for correct algorithms are obtained. Conditions of correctness are formulated in terms of solving a constrained optimization problem. An optimization algorithm based on variable elimination using the spinor method is proposed. Under the conditions of correctness, the proposed method significantly reduces the computational complexity of synthesizing a correct algorithm.

1. Introduction

The notation and definitions from [1] are used. The pattern recognition problem $Z = \{I_0, \tilde{S}^q\}$ is considered, where I_0 is the training information, $I_0 = \{S_1, \ldots, S_m, \tilde{\alpha}(S_1), \ldots, \tilde{\alpha}(S_m)\}$, $S_i = (a_{i1}, \ldots, a_{in})$, $a_{ij} \in M_j$, $\tilde{\alpha}(S_i) \in E_2^l$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, and $\tilde{S}^q = (S^1, \ldots, S^q)$, $S^i = (b_{i1}, \ldots, b_{in})$, $i = 1, \ldots, q$, is the training sample. The information vectors $\tilde{\alpha}$ for $\tilde{S}^q$ are assumed to be known. The aim is to construct an estimation algorithm (EA) that works correctly on the test sample of the vectors of the estimation algorithm.In the general case, an EA is as follows (see [1]):

$$\Gamma_j(S) = x_1\Gamma_j^1(S) + x_0\Gamma_j^0(S), \quad \text{where} \quad x_0, x_1 \in \{0,1\},$$

$$\Gamma_j^1(S) = \frac{1}{Q_1} \sum_{S_i \in \tilde{K}_j} \sum_{\tilde{\omega} \in \{\tilde{\omega}_A\}} \gamma(S_i) p(\tilde{\omega}) B(\tilde{\omega}S_i, \tilde{\omega}S, \tilde{\varepsilon}),$$

$$\Gamma_j^0(S) = \frac{1}{Q_0} \sum_{S_i \in C\tilde{K}_j} \sum_{\tilde{\omega} \in \{\tilde{\omega}_A\}} \gamma(S_i) p(\tilde{\omega}) \overline{B}(\tilde{\omega}S_i, \tilde{\omega}S, \tilde{\varepsilon}); \text{ here,}$$

Γ_j is the estimate of the object S with respect to the jth class, Q_0 and Q_1 are constants, $\tilde{K}_j$ ($C\tilde{K}_j$) is the set of elements of the training sample that belong (respectively, do not belong) to the jth class, $\tilde{\omega}$ is the characteristic Boolean vector of the support set, $\gamma_i = \gamma(S_i)$ are the weights of the training objects, $\gamma_i \in [0, +\infty)$, $p(\tilde{\omega}) = \sum_{\tilde{\omega}} p_j$ are the weights of the features, $p_j \in [0, +\infty)$, B is the proximity function, and $\overline{B} = 1 - B$. We assume that the proximity function B satisfies the following conditions:

- $B(\tilde{\omega}S, \tilde{\omega}S', \varepsilon) = B(\tilde{\omega}S', \tilde{\omega}S, \varepsilon)$; $0 \le B(\tilde{\omega}S, \tilde{\omega}S', \varepsilon) \le 1$;
- $B(\tilde{\omega}S, \tilde{\omega}S', \varepsilon) = 1$, if $\tilde{\omega}S = \tilde{\omega}S'$.

We use the threshold decision rule C with the parameters c_1 and c_2 (see [2, 3]):

$C(c_1, c_2)$: $C(\|\Gamma_{ij}\|) = \|C(\Gamma_{ij})\|$, где

$$C(\Gamma_{ij}) = \begin{cases} 1, \Gamma_{ij} > c_2, \\ 0, \Gamma_{ij} < c_1, \\ \Delta, c_1 \le \Gamma_{ij} \le c_2 \end{cases}, \quad 0 < c_1 < c_2. \qquad ()$$

No explicit descriptions of the class of solvable problems for estimation algorithms are known. Solvability is only proved for some submodels (see [4, 5]). This is explained by the fact that the EA model has many degrees of freedom, and the problem of finding optimal values of the parameters is very difficult. In practice, locally optimal algorithms are usually sought (see [6, 7]). Moreover, the number of the parameters to be optimized is usually limited, which simplifies the optimization problem.

The rest of the article is as follows. In Chapter 2, the optimization problem is formulated. A method for solving the problem using spinors is described. In Chapter 3, a new algorithm for constructing correct EA is proposed. In Conclusion, the results and future research are briefly described.

2. Optimization criteria for EA

The aim of this study is to find out if it is possible to ensure that an algorithm works correctly on the objects of a validation sample by varying only the weights of features and the weights of objects. We also want to find the conditions under which such an algorithm can be

Proceedings of the 12th international workshop on computer science and information technologies CSIT'2010, Moscow – Saint-Petersburg, Russia, 2010

built. The problem of building a correct algorithm is equivalent to solving the system of inequalities

$$\begin{cases} \Gamma_1 > c_2, \ldots, \Gamma_T > c_2 \\ \Gamma_{T+1} < c_1, \ldots, \Gamma_Q < c_1 \end{cases}, \tag{2}$$

where $Q = ql$; $T < Q$; the index $i = 1, 2, \ldots, Q$ corresponds to an enumeration of the two-dimensional array $\{(u, v) \mid u = 1, 2, \ldots, l, v = 1, 2, \ldots, q\}$; and $\Gamma_i = \Gamma_u(S^v)$. For that purpose, the problem is decomposed into two parts: first, a solution of a certain transformed problem is found that contains among its solutions all the solutions of the original problem; then, the solutions of the original problem are among them.

It is well known that the method of algebraic closure of EA (see [1, 2]) makes it possible to build correct algorithms on given validation samples. Concrete algorithms are built as complex multialgorithmic constructs. For that reason, the problem of building correct algorithms in the framework of the original model remains important. In this paper, we consider some problems in which a correct algorithm can be constructed in the framework of the original family of EAs, and an algorithm is proposed that minimizes the number of errors on the validation sample.

Consistency criterion of system (2). We distinguish the case when there exists some $k \in \{1, 2, \ldots, m\}$ such that γ_k belongs to each Γ_i for $i = 1, 2, \ldots, T$ and does not belong to any Γ_i for $i = T + 1, T + 2, \ldots, Q$. In essence, this means that the problem is solvable with respect to the individual object S_k. In this case, one can construct the desired algorithm using, for example, the following weights: $\gamma_k > c_2$, $\gamma_j = 0$, for $j = 1, 2, \ldots, k\text{-}1, k\text{+}1, \ldots, m$, and $p_i = 1$, $i = 1, 2, \ldots, n$. A similar situation occurs when there exists $k \in \{1, 2, \ldots, n\}$ such that p_k belongs to each Γ_i for $i = 1, 2, \ldots, T$ and does not belong to any Γ_i for $i = T + 1, T + 2, \ldots, Q$.

Inconsistency criterion of system (2). Denote the set of all pairwise products of the weights $\pi_k = \gamma_i p_j$ ($k = 1, 2, \ldots, nm$; $i = 1, 2, \ldots, m$; and $j = 1, 2, \ldots, n$). Assume that there exist $t^* \in \{1, \ldots, T\}$ and $t_* \in \{T+1, \ldots, Q\}$, such that $\Gamma_{t^*} = \Gamma_{t^*}(\pi_{i^1}, \ldots, \pi_{i^I})$ and $\Gamma_{t_*} = \Gamma_{t_*}(\pi_{i_1}, \ldots, \pi_{i_J})$ for I, $J \in \{1, 2, \ldots, nm\}$. In this case, if $(i^1, \ldots, i^I) \subset (i_1, \ldots, i_J)$, then it is clear that, for any π_k ($k = 1, 2, \ldots, nm$; $i = 1, 2, \ldots, m$; and $j = 1, 2, \ldots, n$), $\Gamma_{t^*} \le \Gamma_{t_*}$, and system (2) is inconsistent.

The main idea of the proposed approach is as follows. Some inequalities in (2), for example $\Gamma_i < c_1$, $i = T+1, \ldots, Q$, remain intact, and the other inequalities are replaced by the nondecreasing functional $f(\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n)$, such that all the solutions to system (2) are solutions to the constrained optimization problem

$$\begin{cases} f(\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n) \Rightarrow max \\ \Gamma_{T+1} < c_1, \ldots, \Gamma_Q < c_1 \end{cases}. \tag{3}$$

The functional $f(\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n)$ can be chosen such that the solution of problem (2) is reduced to problem (3). The following example confirms this idea. In the case of the functional

$$f(\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n) = \sum_{i=1}^{T} \mathrm{sgn}(\Gamma_i - c_2),$$

if $\max f = T$, then the corresponding solution $W^{(0)} = (\gamma_1^{(0)}, \ldots, \gamma_m^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)})$ obviously is a solution to problem (2). If $\max f < T$, then system (2) is inconsistent.

If system (2) is inconsistent, then we want to find the maximal consistent subsystem. This can be done using the algorithms described in [6, 8].When the simplest functional

$$f(\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n) = \sum_{i=1}^{m} \alpha_i \gamma_i + \sum_{i=1}^{n} \beta_i p_i$$

where $\alpha_i = 1$ for γ_i belonging to at least one $\Gamma_1, \ldots, \Gamma_T$ and $\alpha_i = 0$ for γ_i that do not belong to any $\Gamma_1, \ldots, \Gamma_T$ and, similarly, $\beta_j = 1$ for p_j belonging to at least one $\Gamma_1, \ldots, \Gamma_T$ and $\beta_j = 0$ for p_j that do not belong to any $\Gamma_1, \ldots, \Gamma_T$, problem (3) takes the form

$$\begin{cases} \sum_{i=1}^{m} \alpha_i \gamma_i + \sum_{i=1}^{n} \beta_i p_i \Rightarrow \max \\ \Gamma_{T+1} < c_1, \ldots, \Gamma_Q < c_1 \end{cases}. \tag{4}$$

Theorem 1 (see [9]). Let $f(\gamma, p)$, $\gamma = (\gamma_1, \ldots, \gamma_m)$, $p = (p_1, \ldots, p_n)$, and $\Gamma_i(\gamma, p)$, $i = T+1, \ldots, Q$, be convex functions; $x = (x_1, \ldots, x_{nm})$, $x_k = \gamma_i p_j$, $k = 1, \ldots, nm$, $x \in \Xi$, Ξ be a convex set; and the Slater condition be satisfied (that is, there exists an x^0 such that $\Gamma_i(x^0) < 0$, for $i = T+1, \ldots, Q$. Then, a feasible point x^* is a global solution to problem (3) if and only if there exist $\lambda_i^* \ge 0$ such that $\lambda_i^* \Gamma_i(x^*) = 0$ for $i = T + 1, \ldots, Q$ and $L(x, \lambda^*) \ge L(x^*, \lambda^*)$ for any $x \in \Xi$, where $\lambda^* = (\lambda_{T+1}^*, \ldots, \lambda_Q^*)$ and $L(x, \lambda)$ is the Lagrange function.

The Lagrange function is defined as

$$L(\gamma_1, \ldots, \gamma_m, p_1, \ldots, p_n) =$$

$$= \sum_{i=1}^{m} \alpha_i \gamma_i + \sum_{i=1}^{n} \beta_i p_i + \lambda_1 (\Gamma_{T+1} - c_1) + \ldots + \lambda_t (\Gamma_Q - c_1),$$

$$t = Q - T.$$

The corresponding system consists of $n + m + t$ equations with $n + m + t$ unknowns; it has the form

$$
\begin{cases}
0 = \dfrac{\partial L}{\partial \lambda_1} = \Gamma_{T+1} - c_1, \\
\ldots \\
0 = \dfrac{\partial L}{\partial \gamma_i} = \alpha_i + \lambda_1 \dfrac{\partial \Gamma_{T+1}}{\partial \gamma_i} + \ldots + \lambda_t \dfrac{\partial \Gamma_Q}{\partial \gamma_i}, \\
\ldots \\
0 = \dfrac{\partial L}{\partial p_n} = \beta_i + \lambda_1 \dfrac{\partial \Gamma_{T+1}}{\partial p_n} + \ldots + \lambda_t \dfrac{\partial \Gamma_Q}{\partial p_n}
\end{cases} \tag{5}
$$

For the simple estimates used in recognition problems solved using EAs,

$$
\Gamma_u(S^v) = \frac{1}{N_u} \sum_{\omega} \sum_{S_r \in W_u^1} \gamma_r (p_{i_1} + \ldots + p_{i_k}) B_{v_j}^{\varepsilon}
$$

and system (5) is bilinear:

$$
\begin{cases}
0 = \dfrac{1}{N_{u_1}} \sum\limits_{\omega} \sum\limits_{S_r \in W_{u_1}^1} \gamma_r (p_{i_1} + \ldots + p_{i_k}) B_{v_1}^{\varepsilon} - c_1 \\
\ldots \\
0 = \dfrac{1}{N_{u_t}} \sum\limits_{\omega} \sum\limits_{S_r \in W_{u_t}^1} \gamma_r (p_{i_1} + \ldots + p_{i_k}) B_{v_t}^{\varepsilon} - c_1 \\
0 = \beta_1 + \lambda_1 \dfrac{1}{N_{u_1}} \sum\limits_{\omega} \sum\limits_{S_r \in W_{u_1}^1} \xi_1^{r1} \gamma_r B_{v_1}^{\varepsilon} + \ldots + \\
\quad + \lambda_t \dfrac{1}{N_{u_t}} \sum\limits_{\omega} \sum\limits_{S_r \in W_{u_t}^1} \xi_1^{rt} \gamma_r B_{v_t}^{\varepsilon} \\
\ldots \\
0 = \beta_n + \lambda_1 \dfrac{1}{N_{u_1}} \sum\limits_{\omega} \sum\limits_{S_r \in W_{u_1}^1} \xi_n^{r1} \gamma_r B_{v_1}^{\varepsilon} + \ldots + \\
\quad + \lambda_t \dfrac{1}{N_{u_t}} \sum\limits_{\omega} \sum\limits_{S_r \in W_{u_t}^1} \xi_n^{rt} \gamma_r B_{v_t}^{\varepsilon} \\
0 = \alpha_1 + \lambda_1 \dfrac{1}{N_{u_1}} \sum\limits_{\omega} \xi_{n+1}^{r1} (p_{i_1} + \ldots + p_{i_k}) B_{v_1}^{\varepsilon} + \ldots + \\
\quad + \lambda_t \dfrac{1}{N_{u_t}} \sum\limits_{\omega} \xi_{n+1}^{rt} (p_{i_1} + \ldots + p_{i_k}) B_{v_t}^{\varepsilon} \\
\ldots \\
0 = \alpha_m + \lambda_1 \dfrac{1}{N_{u_1}} \sum\limits_{\omega} \xi_{n+m}^{r1} (p_{i_1} + \ldots + p_{i_k}) B_{v_1}^{\varepsilon} + \ldots + \\
\quad + \lambda_t \dfrac{1}{N_{u_t}} \sum\limits_{\omega} \xi_{n+m}^{rt} (p_{i_1} + \ldots + p_{i_k}) B_{v_t}^{\varepsilon}
\end{cases}
$$

Here, $\xi_k^{ij} \in \{0, 1\}$ depending on the corresponding values of ω and S_r. It is clear that it is difficult to solve such a system analytically in the general case. Various methods for solving systems of nonlinear algebraic equations are available (see [10]). From the viewpoint of obtaining linear relations between the variables in the case when the number of equations is not very large, the spinor method of variable elimination is efficient [11].

The idea underlying this method is to design a procedure for splitting the original nonlinear equations into linear ones in the unknown factors B_i and Φ, where Φ is the spinor that is common for all the equations of the system and B_i is the matrix corresponding to the spinor for the ith equation. The procedure is designed using special algebras of the type of the Clifford algebras. To eliminate the variables, it is sufficient to know the explicit form of B_i. In the forward elimination, the solution to the original equations is found in terms of the basis elements of these algebras. The unknowns are eigenvalues of this solution.

For convenience, we rename the variables in (2.4) as follows: $\gamma_1 = x_1$, …, $\gamma_m = x_m$, $p_1 = x_{m+1}$, …, $p_n = x_{m+n}$, $\lambda_1 = x_{m+n+1}$, …, $\lambda_t = x_{m+n+t}$, $M = m+n+t$,. The corresponding coefficients of the variables are renamed so that a_{ijk} is the coefficient of the term $x_j x_k$ in the ith equation, where $i, j, k \in \{1, 2, \ldots, M\}$. Then, the system can be rewritten as follows:

$$
\begin{aligned}
{}^2R_i \equiv x_1(a_{i11}x_1 + a_{i12}x_2 + \ldots + a_{i1M}x_M + b_{i1}) + \\
+ x_2(a_{i22}x_2 + a_{i23}x_3 + \ldots + a_{i2M}x_M + b_{i2}) + \\
+ \ldots + x_M(a_{iMM}x_M + b_{iM}) + c_i = 0, \; i = 1, \ldots, M.
\end{aligned} \tag{6}
$$

The canonical form of a system of nonlinear algebraic equations is

$$
\begin{aligned}
B_i\Phi \equiv [z_1^{1i}x_1 + z_2^{1i}(a_{i11}x_1 + \ldots + a_{i1M}x_M + b_{i1}) + \\
+ z_1^{2i}x_2 + z_2^{2i}(a_{i22}x_2 + \ldots + a_{i2M}x_M + b_{i2}) + \\
+ \ldots + z_1^{Mi}x_M + z_2^{Mi}(a_{iMM}x_M + b_{iM}) + e_i\sqrt{c_i}\,]\Phi = 0.
\end{aligned} \tag{7}
$$

The elements e_i are generators of the algebra of unipotent alternions (see [12]) defined by the generating relation

$$
e_{i_1}e_{i_2} + e_{i_2}e_{i_1} = 2\varepsilon_{i_1}\delta_{i_1 i_2}e \quad (i = \overline{1, 2p}), \tag{8}
$$

where ε_{i_1} is equal to +1 in s equations and is equal to −1 in the other $2M - s$ equations, e is the unit of the algebra, and $\delta_{i_1 i_2}$ is the Kronecker delta. In what follows, the generators of the algebra of unipotent alternions whose square is equal to e are denoted by α_k and those whose square is equal to $-e$ are denoted by β_k.

The finite number of elements z_σ^{ij} (σ = 1, 2) can be enumerated using a single index t_k to form the elements $z_\sigma^{t_k}$, where t_k ranges through a finite set of integer values; for example, $z_1^{11} = z_1^1$, $z_2^{11} = z_2^1, \ldots$. The elements $z_\sigma^{t_k}$ (σ = 1, 2) are generators of the algebra of nilpotent alternions defined by the generating relation

$$
z_{\sigma_1}^{t_1} z_{\sigma_2}^{t_2} + z_{\sigma_2}^{t_2} z_{\sigma_1}^{t_1} = \varepsilon_{t_1}\delta^{t_1 t_2}\tilde{\delta}_{\sigma_1\sigma_2}e, \tag{9}
$$

where $\delta^{t_1 t_2}$ is the Kronecker delta and $\tilde{\delta}_{\sigma_1\sigma_2} = 1 - \delta_{\sigma_1\sigma_2}$.

Algebras generated by relations (8) and (9) can be combined into the unified algebra of unipotent and nilpotent alternions (or, shortly, the algebra of unions) in which the elements satisfying those permutation relations are linearly independent. Due to the linear independence, we introduce a column Φ that has the order equal to the order of the matrices to guarantee that a solution exists. Such a column Φ, which is common for all the equations in the system ($B_i\Phi = 0$), exists and is defined up to a factor by specifying an irreducible matrix representation of B_i (see [12]). The possibility to pass from (6) to (7) is justified by the following result.

Theorem 2 (see [11]). *Every solution to Eqs. (6) is a solution to Eqs. (7) and conversely.*

System (2.6) can be written in matrix form as $\Lambda X = Q$, where $\Lambda = \|\Lambda_{ij}\|$, $X = \|x_j\Phi\|$, $Q = \|Q_i\|$, $Q_i = -(z_2^{1i}b_{i1} + z_2^{2i}b_{i2} + \ldots + z_2^{Mi}b_{iM} + e_i\sqrt{c_i})\Phi$, and $\Lambda_{ij} = z_2^{1i}a_{i1j} + z_2^{2i}a_{i2j} + \ldots + z_2^{ji}a_{ijj} + z_1^{ji}$. The elimination of unknowns is performed by inverting the alternion block matrix Λ and by multiplying the system by Λ^{-1} on the left. Then, the system takes the form $X = \Lambda^{-1}Q$, and each equation takes the form $x_i\Phi = C_i\Phi$, where C_i is the alternion component of Λ^{-1}Q.

Theorem 3 (see [11]). *The solution $x_i\Phi = C_i\Phi$ is a vector solution to system (6).*

The generators of the algebra of unipotent alternions (8) satisfy the relations $\alpha_k^2 = E, \beta_k^2 = -E$. Therefore, $\alpha_k^{-1} = \alpha_k$, $\beta_k^{-1} = -\beta_k$; in particular, $(\Lambda_{ij})^{-1} = \Lambda_{ij} / a_{ljj}$.

Thus, we have reduced the solution of bilinear system (5) to the eigenvalue problem for x_i of the form $(C_i - x_iE)\Phi = 0$, i.e., to the system of equations

$$(C_i - x_iE)\Phi = ((\Lambda_{ij} / a_{ljj})Q_i - x_iE)\Phi =$$

$$= (\prod_{j=1}^{M}\{-((z_2^{1i}a_{i1j} + z_2^{2i}a_{i2j} + \ldots + z_2^{ji}a_{ijj} + z_1^{ji}) / a_{ljj})$$

$$(z_2^{1i}b_{i1} + z_2^{2i}b_{i2} + \ldots + z_2^{Mi}b_{iM} + e_i\sqrt{c_i})\} - x_iE)\Phi = 0,$$

where $i = 1, \ldots, m+n+t$, $M = m+n+t$, $x_1 = \gamma_1, \ldots, x_m = \gamma_m$, $x_{m+1} = p_1, \ldots, x_{m+n} = p_n$, $x_{m+n+1} = \lambda_1, \ldots, x_{m+n+t} = x_M = \lambda_t$, the coefficients $a_{ijk} = \{ \sum_{g=1}^{m}\frac{\theta_g B_{v_g}^{\varepsilon}}{N_{u_g}}$, for $j = 1, \ldots, m$, $k = m + 1, \ldots, m + n$; $\sum_{g=1}^{t}\frac{\theta_g \xi_g^{r_j} B_{v_j}^{\varepsilon}}{N_{u_j}}$, for $j = m + n + 1, \ldots, m + n + t$, $k = 1, \ldots, m$; $\sum_{g=1}^{t}\frac{\theta_g \xi_{n+g}^{r_j} B_{v_j}^{\varepsilon}}{N_{u_j}}$, for $j = m + n + 1, \ldots, m + n + t$, $k = m + 1, \ldots, m + n$; $\theta_g = \theta_g(j, k) \in \{1, 2\}\}$ and Φ is the spinor that is common for all the equations in the system.

If these equations are multiplied on the left by $C_i : x_iC_i\Phi = (C_i)^2\Phi$, then $x_i^2\Phi = (C_i)^2\Phi$, and we obtain the following result.

Theorem 4 (see [11]). *The problem $(C_i - x_iE)\Phi = 0$, $(i = 1, \ldots, p)$ is an eigenvalue problem. To obtain the characteristic equation, no matrix representation of e_i and $z_\sigma^{t_k}$ is required.*

The resulting characteristic equation can be solved numerically to obtain 2^M values of x_i. For each of them, the corresponding Φ can be found, and the spinor Φ determines the tuples of solutions.

3. Optimization algorithm

The general scheme of building an algorithm that minimizes the number of errors for a certain sample is as follows. The violated inequalities for Γ_i in (2) ($i \in \{1, \ldots, T\}$) are found. According to a given criterion Cr_1, one of them is selected. An attempt to satisfy it is made by modifying some variables such that the valid inequalities remain true. The variables are chosen in accordance with a given criterion Cr_2. If such a modification can be done, the next violated inequality is chosen, and the procedure is repeated. Below, we consider an example of the implementation of the proposed scheme.

1. An arbitrary solution $W^{(0)} = (\gamma_1^{(0)}, \ldots, \gamma_m^{(0)}, p_1^{(0)}, \ldots, p_n^{(0)})$ to problem (3) is substituted into (2). It is clear that the inequalities in (2) with the indices $i = T + 1, \ldots, Q$ are satisfied.
2. The inequalities in (2) are arranged in descending order of their left-hand sides upon the substitution of $W_{(0)}$. Among the inequalities with the indices $i = 1, \ldots, T$, select those that are violated, i.e., the inequalities in which the sum on the left-hand side is less than c_2:

$$\Gamma_u(S^v) = \sum_{i=1}^{n} p_i(\gamma_{j_1} + \ldots + \gamma_{j_s}) = p(\gamma_{j_1} + \ldots + \gamma_{j_s}) =$$

$$= a_{uv} = c_2 - \delta_{uv},\ \delta_{uv} > 0.$$

3. Among these inequalities, find the one for which δ_{uv} is the largest. Assume that the left-hand side of this inequality contains the weights $\gamma_{j_1}, \ldots, \gamma_{j_s}$, with the indices $j_1, \ldots, j_s$. Without loss of generality, we assume that the indexation of the inequalities remains the same as in (2) when they are ordered. Thus, the inequality with the index T has the following form:

№	γ_1	γ_2	…	γ_k	…	γ_{m-}	γ_m	
1								$> c_2$
…								…
S								$> c_2$
S+1								$= c_2 - \delta_{S+2}$
…								…
T	0	1	…	1	…	1	0	1	$= c_2 - \delta_T$
T+1				0					$= c_1 - \delta_{T+1}$

…				0					…
P				1					$= c_1 - \delta_P$
…									…
Q									$= c_1 - \delta_Q$

Fig. 1. The form of the inequality with index T

4. Among the inequalities with the indices $i = T + 1$, …, Q involving the weights $\gamma_{j_1}, K, \gamma_{j_s}$ the inequality with the minimum index P_i is found for each γ_{j_i}. Then, the inequality with the index $P = \max\limits_{i=1,s}\{P_i\}$ and the weight γ_{j_i} corresponding to this inequality are chosen. In this inequality, the difference between the left-hand side and c_1 is the largest.
5. The weight γ_{j_i} increases by the maximum possible quantity ξ_P, where $\xi_P < \delta_P \Big/ \sum\limits_{i=1}^{n} p_i$. If $\delta_P > \delta_T$, then go to the inequality with the index $T - 1$, and the procedure is repeated from step 4. Otherwise, the inequality with the index $P' = \max\limits_{i=1,s}(\{P_i\} \setminus P)$ and the corresponding weight γ_{j_i} is increased by $\xi_{P'}$, where $\xi_{P'} < \delta_{P'} \Big/ \sum\limits_{i=1}^{n} p_i$.
6. The procedure is continued until the left-hand sides of all the inequalities with the indices from 1 through T exceed c_2 or until the reserve in the inequalities with the indices from $T + 1$ to Q is exhausted.

This algorithm, which minimizes the number of errors in the recognition problem under examination, belongs to the class of greedy algorithms. Greedy algorithms are computationally efficient. In many combinatorial optimization problems, they produce solutions that are close to the optimal ones. If the set of inequalities (2) is a matroid (see [13]), then, by the Rado–Edmonds theorem [14], the greedy algorithm produces an optimal solution.

4. Conclusion

- An algorithm for solving optimization problems with inequalities as constraints is proposed. The algorithm is based on the spinor algebra.
- The algorithm is used for constructing correct EA algorithms of pattern recognition.
- For some classes of problems, criteria of feasibility are obtained, which are formulated in terms of solving a constrained optimization problem.
- Under the obtained conditions, the proposed method significantly reduces the computational complexity of synthesizing a correct algorithm.

Future work is related to the investigation of classes of pattern recognition problems that can be efficiently solved using the proposed approach.

References

1. Zhuravlev Yu. I. "Correct Algorithms over Sets of Incorrect (Heuristic) Algorithms: Part II". *Kibernetika* 1977; 6:21-27.
2. Zhuravlev Yu. I. "Correct Algorithms over Sets of Incorrect (Heuristic) Algorithms: Part III". *Kibernetika* 1978; 2:35-43.
3. Zhuravlev Yu. I., Isaev I. V. "Constructing Recognition Algorithms That are Correct for a Given Validation Sample". *Zh. Vychisl. Mat. Mat. Fiz.* 1979; 19:726–738.
4. Rudakov K. V. "On Certain Classes of Recognition Algorithms (General Results)". Vychisl. Tsentr Akad. Nauk SSSR, Moscow, 1980.
5. Rudakov K. V. "On Certain Classes of Recognition Algorithms (Parametric Models)". Vychisl. Tsentr Akad. Nauk SSSR, Moscow, 1981.
6. Ryazanov V. V. "Optimization of Estimation Algorithms Using Parameters Describing the Representative Property of Reference Rows". *Zh. Vychisl. Mat. Mat. Fiz.* 1976; 16:1559–1570 (1976).
7. Ryazanov V. V. *Optimal Group Decisions in Recognition and Classification*. Doctoral Dissertation in Mathematics and Physics, VTs RAN, Moscow, 1994.
8. Katerinochkina N. N. "Methods for Finding a Maximal Consistent Subsystem of a System of Linear Inequalities". Communications in Applied Mathematics, Vychisl. Tsentr, Ross. Akad. Nauk, Moscow, 1997.
9. Polyak B. T. "Introduction to Optimization". Nauka, Moscow, 1983.
10. Pshenichnikov S. B., Shtern A. G. "Methods for Solving Systems of Nonlinear Algebraic Equations". Mathematical Methods and Computer-Aided Systems in Geology: A Survey, Moscow, 1995.
11. Pshenichnikov S. B. "Spinor Elimination and Formulae for Relations between Unknowns in Systems of Nonlinear Algebraic Equations". *Latv. Mat. Ezhegodnik* 1986; 30:150–161.
12. Zaitsev G. A. "Algebraic Problems in Mathematical and Theoretical Physics" Nauka, Moscow, 1974.
13. Rybnikov K. A. "Introduction to Combinatorial Analysis". Mosk. Gos. Univ., Moscow, 1985.
14. Edmonds J. "Matroids and the Greedy Algorithm". *Math. Program.* 1971; 1:127 – 136.