

Basic Properties of the Classification Hyperplanes Represented by Extremum Points in the Multidimensional Hough Space¹

Valeri Ilchev*, Zlatolilia S. Ilcheva**

* Institute of Information Technologies, 1113 Sofia

E-mail: ilch@iinf.bas.bg, vv@ilchev.net

** Institute of Computer and Communication Systems, 1113 Sofia

E-mail: zlat@iccs.isdip.bas.bg

Abstract: *In the paper, the indirect presentation of the convex covers of two arbitrary nonintercepted and compact sets in the multidimensional Hough space is examined. The basic purpose of this study is the analysis of the connection between the extremum sets of the images of these covers and the hyperplanes classifying these two compact sets.*

Two basic cases for the compact sets are examined: linearly separable and linearly unseparable sets. In both cases it is proved that the points of the extremum sets in the Hough space synonymously correspond by the reverse image, to the optimum classifying hyperplanes in the object space. In the second case it is proved that the extremum point corresponding to the classifying hyperplane which is the the best approximation hyperplane by Chebishev for the optimum hypersurface separating these two compact sets, in the sense of its definition.

Keywords: *Classification, theory of sets, Hough transform, Chebishev approximation, multidimensional variations.*

1. Introduction

The multidimensional Hough transform enables the compact presentation of different groups of hyperplains as sets of points in the n -dimensional space, defined by this transform (Hough space).

¹ This research is supported by the Institute of Information Technologies, BAS, Project No 010085, "Neuroinformation systems for classification and identification".

The analysis in the paper illustrates that if we have two non-intersecting sets S_1 and S_2 in the space E^n , which have to be classified, and if a set of hyperplanes is defined, which separate S_1 and S_2 (strictly or non-strictly) in some way, then the image of this set in the Hough-space will carry information for the mutual disposition of S_1 and S_2 and for the manner of their separation.

All hyperplanes in E^n can be divide into three groups which form three basic sets, namely:

- 1) hyperplain sets, which intercept the inside of the sets S_1 and (or) S_2 : $\text{Int}(S_1)$, $\text{Int}(S_2)$;
- 2) hyperplane sets which are supporting to S_1 and (or) S_2 ;
- 3) hyperplain sets which have not common points with S_1 and (or) S_2 .

By means of these sets we can present indirect S_1 and S_2 in the Hough space \mathcal{L}^n [1].

Let us consider such a way for presenting in the space \mathcal{L}^n of any of this two sets, for example, the set S_1 (Fig. 1a). For that purpose at first we will set the three basic hyperplane sets which will be subsets of the set of all hyperplanes, defined in the space E^n : $\mathbb{H} = \{H : H \subset E^n\}$. Depending on the location of the hyperplane sets in relation to S_1 , we will have the following conditions for their determination:

- 1) $\mathbb{H}_k = \{H_k : H_k \cap \text{Int}(S_1) \neq \emptyset\}$ – for example, the hyperplane H_k on Fig. 1a;
- 2) $\mathbb{H}_b = \{H_b : H_b \cap \text{Fr}(S_1) \neq \emptyset \wedge H_b \cap \text{Int}(S_1) = \emptyset\}$ – the hyperplanes H_a and H_b ;
- 3) $\mathbb{H}_e = \{H_e : H_e \cap \overline{S_1} \neq \emptyset\}$ – the hyperplanes H_{ab}^i and H_{ab}^j on Fig. 1a. In this way, the three subsets will be defined by the form and the location of S_1 in E^n , which means that they will contain information for the more common characteristics of the S_1 set. Further we will unify the second and the third subsets and will note them by \mathbb{H}_1 : $\mathbb{H}_1 = \mathbb{H}_b \cup \mathbb{H}_e$.

2. Basic properties of the multidimensional Hough transform

Initially, for the further analysis of this task, we will formulate the following theorem [2]:

Theorem 1. The set $\mathcal{H}_1 = T(\mathbb{H}_1)$ is convex and closed.

In this case by $T(\dots)$ is noted the Hough transform: $E^n \rightarrow \mathcal{L}^n$. According to this theorem, the images of the supporting to S_1 hyperplanes will be points from the bound of the set $\mathcal{H}_1 \subset \mathcal{L}^n$, for example $T(H_a) = h_a \in \text{Fr}(\mathcal{H}_1)$ and $T(H_b) = h_b \in \text{Fr}(\mathcal{H}_1)$, and the hyperplane images which do not intersect the set $\overline{S_1}$ will be points which belong to the inside of \mathcal{H}_1 set, for example: $T(H_{ab}^i) = h_{ab}^i \in \text{Int}(\mathcal{H}_1)$ and $T(H_{ab}^j) = h_{ab}^j \in \text{Int}(\mathcal{H}_1)$. Accordingly, the hyperplanes which intersect the set S_1 , i. e. which belong to the set \mathbb{H}_k , will be represented in the space \mathcal{L}^n as points outside \mathcal{H}_1 : $T(H_k) = h_k \notin \mathcal{H}_1$, where \mathcal{H}_1 is a closed set $\mathcal{H}_1 = \overline{\mathcal{H}_1}$ (Fig. 1b).

In the same manner we can represent in the space \mathcal{L}^n the three sets “interacting” with the set $S_2 \subset E^n$. For the set \mathbb{H}_2 we will have $\mathbb{H}_2 = \mathbb{H}_b^2 \cup \mathbb{H}_e^2$ and its image in \mathcal{L}^n will be $T(\mathbb{H}_2) = \mathcal{H}_2$, where $\mathcal{H}_2 = \overline{\mathcal{H}_2}$ is convex and closed set. In this case:

$\mathbb{H}_b^2 = \{H_b^2 : H_b^2 \cap \text{Fr}(S_1) \neq \emptyset \wedge H_b^2 \cap \text{Int}(S_1) = \emptyset\}$ and $\mathbb{H}_e^2 = \{H_e^2 : H_e^2 \cap \overline{S_1} \neq \emptyset\}$. With the help of transforms $T(\mathbb{H}_1)$ and $T(\mathbb{H}_2)$ we can represent (indirectly) in the space \mathcal{L}^n simultaneously the two sets S_1 and S_2 and to analyze the common properties of their images, which will be made in Section 3 of this paper. The receiving results will be used in Section 4 and Section 5 by image analysis of the classifying hyperplanes by linearly separable and linearly unseparable sets accordingly.

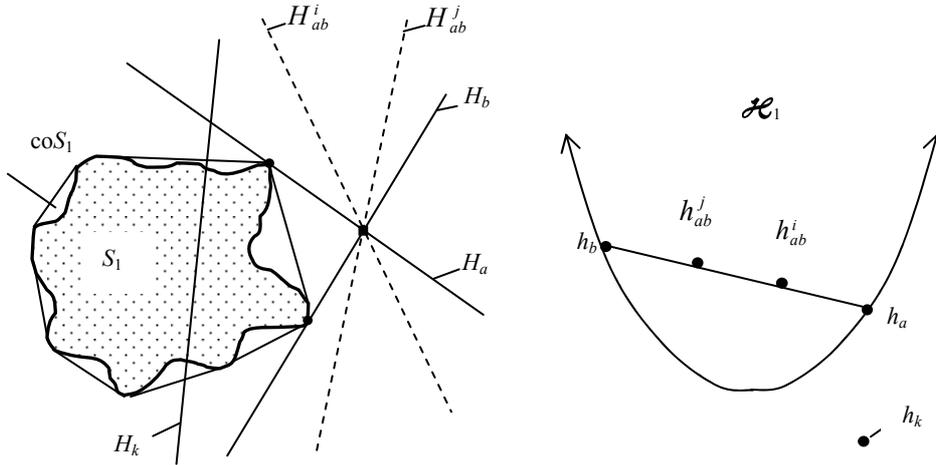


Fig. 1. The envelope $\text{co}S_1$ of the set S_1 and its supporting, not intersect and intersect the set S_1 hyperplanes (a). The images of these hyperplanes in the space $\mathcal{L}^n(b)$

This analysis will allow to determine the characteristic features of the classifying hyperplanes, whose images in the space \mathcal{L}^n belong to an extreme set of points, defining further in this paper.

3. Representing of two compact sets in the Hough space

Let two unintersected sets S_1 and S_2 be given, for which we suggest they are strongly linear separable: $S_{12} = \text{co}S_1 \cap \text{co}S_2 = \emptyset$ and let us consider the set of hyperplanes $\mathbb{H}_1 \cup \mathbb{H}_2$ (Fig. 2a), where $\text{co}S_1$ and $\text{co}S_2$ are the envelope of the sets S_1 and S_2 (the sets $\text{co}S_1$ and $\text{co}S_2$ are always convex).

Then according to Theorem 1, the hyperplanes H_a and H_b , which are simultaneously supporting to the sets $\text{co}S_1$ and $\text{co}S_2$ in the space E^n (Fig. 2a), will represent as the points h_a and h_b in \mathcal{L}^n , which will belong as well to the boundary $\text{Fr}(\mathcal{H}_1)$ of the set \mathcal{H}_1 , as to the boundary of the set \mathcal{H}_2 : $\text{Fr}(\mathcal{H}_2)$, i. e. $h_a \in \text{Fr}(\mathcal{H}_1) \cap \text{Fr}(\mathcal{H}_2)$ and $h_b \in \text{Fr}(\mathcal{H}_1) \cap \text{Fr}(\mathcal{H}_2)$ (Fig. 2b).

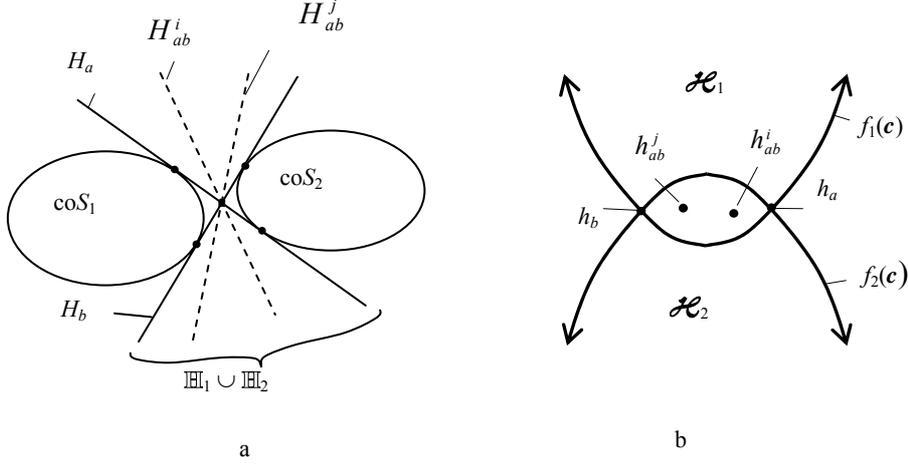


Fig. 2. Two unintersected sets $\text{co}S_1, \text{co}S_2$ and their set of hyperplanes $\mathbf{H}_1 \cup \mathbf{H}_2$ (a); the images of the hyperplanes of the set $\mathbf{H}_1 \cup \mathbf{H}_2$ in the space \mathcal{L}^n (b)

According to the same theorem, hyperplanes H_{ab}^i and H_{ab}^j , which strongly separate the two sets $\text{co}S_1$ and $\text{co}S_2$ and obviously fulfill the conditions $H_{ab}^i \cap \text{co}S_1 = \emptyset \wedge H_{ab}^i \cap \text{co}S_2 = \emptyset$ and $H_{ab}^j \cap \text{co}S_1 = \emptyset \wedge H_{ab}^j \cap \text{co}S_2 = \emptyset$, will be represented as inside points simultaneously of the two sets \mathcal{H}_1 and \mathcal{H}_2 in \mathcal{L}^n : $h_{ab}^i \in \text{Int}(\mathcal{H}_1) \wedge h_{ab}^i \in \text{Int}(\mathcal{H}_2)$ and $h_{ab}^j \in \text{Int}(\mathcal{H}_1) \wedge h_{ab}^j \in \text{Int}(\mathcal{H}_2)$ (Fig. 2b). From this immediately follows:

$$h_{ab}^i, h_{ab}^j \in \text{Int}(\mathcal{H}_1) \cap \text{Int}(\mathcal{H}_2) \Rightarrow \text{Int}(\mathcal{H}_1) \cap \text{Int}(\mathcal{H}_2) \neq \emptyset.$$

These theoretical conclusions allow us to formulate the following statement, which proof in details is given in [2]:

Statement 1. If in the space E^n the compact sets $\text{co}S_1$ and $\text{co}S_2$ are strongly linearly separable i. e. they do not intercept each other, then in the space \mathcal{L}^n the inside sets of the sets \mathcal{H}_1 and \mathcal{H}_2 will intercept each other:

$$\text{Int}(\mathcal{H}_1) \cap \text{Int}(\mathcal{H}_2) \neq \emptyset.$$

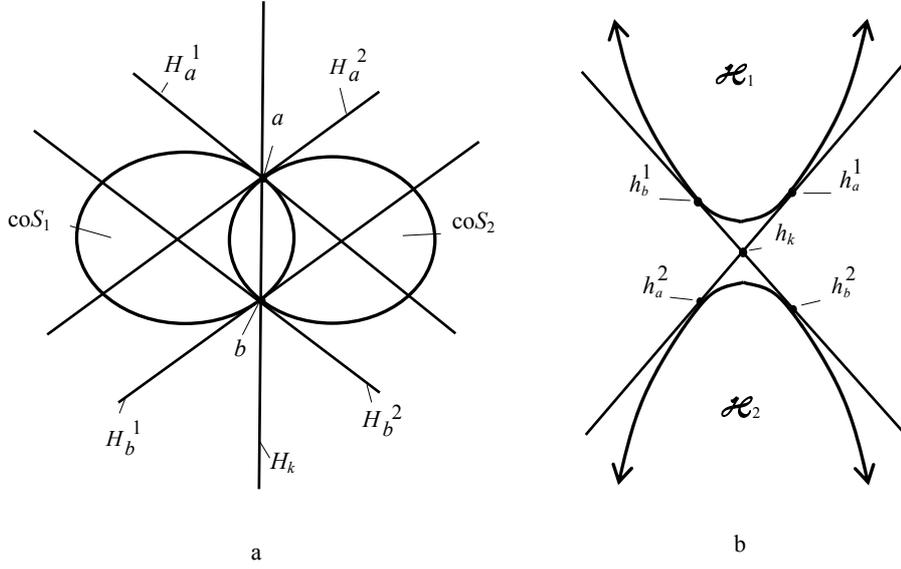


Fig. 3. Two intersected sets $\text{co}S_1, \text{co}S_2$ and their supporting hyperplanes H_a^1, H_a^2 and H_b^1, H_b^2 together the hyperplane H_k , which intersect the set S_{12} (a). The images of these hyperplanes in the space \mathcal{L}^n (b)

Let us examine the second case where we will have two intersected sets $S_{12} = \text{co}S_1 \cap \text{co}S_2 \neq \emptyset$ (Fig. 3a). If we take in the space E^n the points a and b , defined from the condition $a \in \text{Fr}(\text{co}S_1) \cap \text{Fr}(\text{co}S_2)$ and $b \in \text{Fr}(\text{co}S_1) \cap \text{Fr}(\text{co}S_2)$, then each of these points we can examine as a point, lying on the axis of the corresponding bunch of hyperplanes \mathbb{H}_a and \mathbb{H}_b .

The bunch \mathbb{H}_a will be limited from the supporting, respectively to $\text{co}S_1$ and $\text{co}S_2$, hyperplanes H_a^1 and H_a^2 , where $H_a^1, H_a^2 \subset \mathbb{H}_a$ and the bunch \mathbb{H}_b will be limited from the hyperplanes $H_b^1, H_b^2 \subset \mathbb{H}_b$, supporting as well (respectively) to $\text{co}S_1$ and $\text{co}S_2$. Then, according to the Theorem 1, the hyperplanes H_a^1 and H_a^2 will be represented in the space \mathcal{L}^n as bounding points respectively of the sets \mathcal{H}_1 and \mathcal{H}_2 : $T(H_a^1) = h_a^1 \in \text{Fr}(\mathcal{H}_1)$ and $T(H_a^2) = h_a^2 \in \text{Fr}(\mathcal{H}_2)$ – on Fig. 3b.

By analogy, the supporting hyperplanes H_b^1 and H_b^2 will be represented as bounding points, accordingly: $T(H_b^1) = h_b^1 \in \text{Fr}(\mathcal{H}_1)$ and $T(H_b^2) = h_b^2 \in \text{Fr}(\mathcal{H}_2)$ (Fig. 3b). Hyperplanes H_a and H_b , which are inside elements accordingly to the bunch \mathbb{H}_a : $H_a \subset \mathbb{H}_a$ and to the bunch \mathbb{H}_b : $H_b \subset \mathbb{H}_b$, will intercept the sets $\text{co}S_1$ and $\text{co}S_2$, i. e. they will satisfy the conditions: $H_a \cap (\text{co}S_1 \cap \text{co}S_2) \neq \emptyset$ and $H_b \cap (\text{co}S_1 \cap \text{co}S_2) \neq \emptyset$ (Fig. 3a). As follows from Theorem 1, these hyperplanes will be represented as points outside the sets \mathcal{H}_1 and \mathcal{H}_2 in the space \mathcal{L}^n , for example hyperplane H_k , where $H_k \cap (\text{co}S_1 \cap \text{co}S_2) \neq \emptyset$ will be represented as the point $h_k = T(H_k)$ in \mathcal{L}^n and will execute the condition: $h_k \in (\mathcal{L}^n \setminus \mathcal{H}_1) \cup (\mathcal{L}^n \setminus \mathcal{H}_2)$ (Fig. 3b). That means if we have a hyperplane bunch \mathbb{H}_z with a centre in the point z ,

belonging to the inside of the set $\text{co}S_1 \cap \text{co}S_2$, then for the image $T(H_z) = \mathcal{H}_z$, which will be a hyperplane in \mathcal{L}^n we will have $\mathcal{H}_z \cap \mathcal{H}_1 = \emptyset$ and $\mathcal{H}_z \cap \mathcal{H}_2 = \emptyset$, i. e. this hyperplane will divide strongly the sets \mathcal{H}_1 and \mathcal{H}_2 [3]. It is clear that if the closed, convex and infinite sets \mathcal{H}_1 and \mathcal{H}_2 can be strongly separated by the hyperplane (in this case the hyperplane \mathcal{H}_z), then for these sets will be executed the condition: $\mathcal{H}_1 \cap \mathcal{H}_2 = \emptyset$. From here we can formulate the following statement:

Statement 2. If in the space E^n the compact sets $\text{co}S_1$ and $\text{co}S_2$ are such that $\text{Int}(\text{co}S_1) \cap \text{Int}(\text{co}S_2) \neq \emptyset$, then the sets \mathcal{H}_1 and \mathcal{H}_2 (in the space \mathcal{L}^n) will execute the condition: $\mathcal{H}_1 \cap \mathcal{H}_2 = \emptyset$, i.e. they will be strongly separable.

4. Representation of the classification hyperplanes for linear separable sets in the Hough space

Let for the compact sets $\text{co}S_1$ and $\text{co}S_2$ hold the condition $\text{co}S_1 \cap \text{co}S_2 = \emptyset$. Then according to Statement 1, in the space \mathcal{L}^n will be valid the condition $\mathcal{H}_{12} = \text{Int}(\mathcal{H}_1) \cap \text{Int}(\mathcal{H}_2) \neq \emptyset$. This means that every inside point $h_i \in \text{Int}(\mathcal{H}_{12})$ will be an image in the space \mathcal{L}^n of strongly separating the sets S_1 and S_2 hyperplane $H_i = T^{-1}(h_i)$, defined in the object space E^n , where: $H_i \cap \bar{S}_1 = \emptyset \wedge H_i \cap \bar{S}_2 = \emptyset$ and $\bar{S}_1 \subset [H_i]$, $\bar{S}_2 \subset [H_i^-]$ or $\bar{S}_1 \subset [H_i^-]$, $\bar{S}_2 \subset [H_i]$. The sets $S_1 = \bar{S}_1$ and $S_2 = \bar{S}_2$ are closed and $[H_i]$, $[H_i^-]$ are accordingly the positive and the negative half-spaces of the hyperplane H_i . Since the set \mathcal{H}_{12} is closed (because it is an intersection of two closed sets \mathcal{H}_1 and \mathcal{H}_2 , according to the Theorem 1), then $\forall h_i \in \text{Int}(\mathcal{H}_{12})$ will be limited simultaneously from below – by $\text{Fr}(\mathcal{H}_1)$ and from above – by $\text{Fr}(\mathcal{H}_2)$ (Fig. 2b).

Let us specify the functions $f_1(\mathbf{c})$ and $f_2(\mathbf{c})$, describing analytically the set boundaries of \mathcal{H}_1 and \mathcal{H}_2 and let us determine the area $\Omega \subset \mathcal{C}^{n-1}$ in the subspace of the arguments $\{\mathbf{c}\} = \mathcal{C}^{n-1}$ by means of the condition $\mathbf{c}_\Omega = \{\mathbf{c} \in \Omega: f_1(\mathbf{c}) \leq f_2(\mathbf{c})\}$, where $\mathcal{C}^{n-1} \times L^1 = \mathcal{L}^n$. Then in the area Ω for the function:

$$(1) \quad f_r(\mathbf{c}) = f_1(\mathbf{c}) - f_2(\mathbf{c}) \text{ we will have } f_r(\mathbf{c}) \leq 0.$$

Since the set \mathcal{H}_{12} is convex, then for any pair of points $h_i(\mathbf{c}_i, L_i) \in \text{Int}(\mathcal{H}_{12})$ and $h_j(\mathbf{c}_j, L_j) \in \text{Int}(\mathcal{H}_{12})$, where $\mathbf{c}_i = \mathbf{c}_j$, $\mathbf{c}_i, \mathbf{c}_j \in \Omega$, will be valid the inequality $|L_i - L_j| \leq f_r(\mathbf{c}_i)$, because for the points $h_1^i[\mathbf{c}_i, f_1(\mathbf{c}_i)]$ and $h_2^i[\mathbf{c}_i, f_2(\mathbf{c}_i)]$ we have: $h_1^i \in \text{Fr}(\mathcal{H}_1)$ and $h_2^i \in \text{Fr}(\mathcal{H}_2)$.

Let us analyze in more details the function $f_r(\mathbf{c})$ from condition (1). For its extremum we will have: $\min_{\mathbf{c}} f_r(\mathbf{c}) < 0$, from where for the extremum point (or

points) \mathbf{c}_e will be valid the condition: $|f_1(\mathbf{c}_e) - f_2(\mathbf{c}_e)| = \max_{\mathbf{c}} |f_1(\mathbf{c}) - f_2(\mathbf{c})|$, for $\mathbf{c} \in \Omega$.

The reverse image of the extremum of $f_r(\mathbf{c})$ will have specific properties in comparison with the reverse image of any other point of this function, i. e. at $\mathbf{c} \notin \{\mathbf{c}_e\}$. This fact is illustrated by means of theorem:

Theorem 2. If $\mathbf{c}_e \in \mathcal{C}^{n-1}$ is extremum point of the function $f_r(\mathbf{c})$, for which $f_r(\mathbf{c}_e) = \min_{\mathbf{c}} f_r(\mathbf{c})$ in the space \mathcal{L}^n , then in the space E^n , the vector \mathbf{c}_e will coincide

in direction with the normal vector of the pair parallel hyperplanes H_e^1 and H_e^2 , separating unstrictly the sets S_1 , S_2 and executing the conditions: $|y_e^1 - y_e^2| = \max_{\mathbf{c}} |y_p^1 - y_p^2|$, where: $y_e^1 = \mathbf{c}_e^T \mathbf{x} - L_{e1}$, $y_e^2 = \mathbf{c}_e^T \mathbf{x} - L_{e2}$ and $y_p^1 = \mathbf{c}_p^T \mathbf{x} - L_{p1}$,

$y_p^2 = \mathbf{c}_p^T \mathbf{x} - L_{p2}$, $\mathbf{x} = (x_1, x_2, \dots, x_{n-1}) \in \mathbf{X}^{n-1}$.

P r o o f. Let us examine anyone pair of parallel hyperplanes H_p^1 and H_p^2 , separating unstrictly the two sets S_1, S_2 and defining by the equations $y_p^1 = \mathbf{c}_p^T \mathbf{x} - L_{p1}$, $y_p^2 = \mathbf{c}_p^T \mathbf{x} - L_{p2}$. The distance between the two hyperplanes (along the axis Y) will be

$$(2) \quad |y_p^1 - y_p^2| = |L_{p1} - L_{p2}| = |\mathbf{c}_p^T \mathbf{x}_0 - y_{01} - \mathbf{c}_p^T \mathbf{x}_0 + y_{02}| = |y_{01} - y_{02}|,$$

where for the points $\mathbf{x}_0^1(x_0, y_{01})$ and $\mathbf{x}_0^2(x_0, y_{02})$ in the object space E^n we will have: $\mathbf{x}_0^1 \in H_p^1$ and $\mathbf{x}_0^2 \in H_p^2$. Since by condition the hyperplanes H_p^1 and H_p^2 divide unstrictly the sets S_1 and S_2 accordingly, then H_p^1 will be a supporting hyperplane to S_1 , and H_p^2 will be a supporting hyperplane to S_2 . Then, according to the Theorem 1, for the images of these hyperplanes in the space \mathcal{L}^n , we will have: $T(H_p^1) = h_p^1(\mathbf{c}_p, L_{p1}) \in \text{Fr}(\mathcal{H}_1)$ and $T(H_p^2) = h_p^2(\mathbf{c}_p, L_{p2}) \in \text{Fr}(\mathcal{H}_2)$, where the boundaries $\text{Fr}(\mathcal{H}_1)$ and $\text{Fr}(\mathcal{H}_2)$ are described analytically by the functions $f_1(\mathbf{c})$ and $f_2(\mathbf{c})$. That means that each of the points h_p^1 and h_p^2 can be also specified in the following way: $h_p^1 = [\mathbf{c}_p, f_1(\mathbf{c}_p)]$ and $h_p^2 = [\mathbf{c}_p, f_2(\mathbf{c}_p)]$, from where the equalities: $f_1(\mathbf{c}_p) = L_{p1}$ and $f_2(\mathbf{c}_p) = L_{p2}$ follow.

For each hyperplane H_p^1 and H_p^2 , unstrongly separating the sets S_1 and S_2 , we will have two cases: 1) $H_p^1 \cap \overline{S}_2 = \emptyset$ and $H_p^2 \cap \overline{S}_1 = \emptyset$, where $\overline{S}_1 = S_1$ and $\overline{S}_2 = S_2$ are closed sets, and 2) H_p^1 and (or) H_p^2 are supporting to the $\text{co}S_1$ and $\text{co}S_2$ simultaneously.

Let us examine the first case, for which according to the Theorem 1 we will have $T(H_p^1) = h_p^1 \in \text{Int}(\mathcal{H}_2)$, where $\mathcal{H}_2 = \overline{\mathcal{H}_2}$. Along with this, obviously

$h_p^1 \in \mathcal{H}_{12}$ (because $h_p^1 \in \text{Fr}(\mathcal{H}_1)$), where $\mathcal{H}_{12} = \overline{\mathcal{H}_{12}}$ is closed set. By analogy, for $T(H_p^2) = h_p^2 \in \text{Int}(\mathcal{H}_1)$, we will receive $h_p^2 \in \mathcal{H}_{12}$.

For the second case we will have

$$h_p^1 \in \text{Fr}(\mathcal{H}_1) \wedge h_p^1 \in \text{Fr}(\mathcal{H}_2) \Rightarrow h_p^1 \in \text{Fr}(\mathcal{H}_1) \cap \text{Fr}(\mathcal{H}_2)$$

if H_p^1 executes the condition of case 2) and respectively $h_p^2 \in \text{Fr}(\mathcal{H}_1) \cap \text{Fr}(\mathcal{H}_2)$ if H_p^2 executes the same condition.

Let us analyze in more detail the first case. Since \mathcal{H}_2 is a convex set, and the function $f_2(\mathbf{c})$, which analytically describes the bounding of this set is concave [2, Property 1.3], then \mathcal{H}_2 will be hypograph of the function $f_2(\mathbf{c})$: $\text{hyp}f_2(\mathbf{c})$ and will be determined by the condition $\{h_2(\mathbf{c}, L): \mathbf{c} \in \mathcal{C}^{n-1}, L \in L^1, L \leq f_2(\mathbf{c})\} = \mathcal{H}_2$. Then as soon as the point $h_p^1 = [\mathbf{c}_p, f_1(\mathbf{c}_p)]$ is such that $h_p^1 \in \text{Int}(\mathcal{H}_2)$, we will receive the conditions $L_{p1} < f_2(\mathbf{c}_p) \Rightarrow f_1(\mathbf{c}_p) < f_2(\mathbf{c}_p)$, in view of the equality $L_{p1} = f_1(\mathbf{c}_p)$.

For the second case from the condition $h_p^1 \in \text{Fr}(\mathcal{H}_1) \cap \text{Fr}(\mathcal{H}_2)$ follows $f_1(\mathbf{c}_p) = f_2(\mathbf{c}_p)$. And the same equality will be valid for h_p^2 : $h_p^2 \in \text{Fr}(\mathcal{H}_1) \cap \text{Fr}(\mathcal{H}_2)$, from where immediately follows: $f_r(\mathbf{c}_p) = 0$. Let us examine all the points $h_p^1, h_p^2 \in \text{Fr}(\mathcal{H}_1) \cap \text{Fr}(\mathcal{H}_2)$. They will belong to the intersection of the function $f_r(\mathbf{c}_p)$ with the subspace \mathcal{C}^{n-1} , by that for the convex function $f_r(\mathbf{c})$, which is difference between convex and concave function, we will have $f_r(\mathbf{c}_g) > 0$ for $\mathbf{c}_g \notin \overline{\mathbf{c}_i \mathbf{c}_j}$, where the end points will execute the condition $\mathbf{c}_i, \mathbf{c}_j \in \{\mathbf{c}_\Omega^0: f_r(\mathbf{c}_\Omega^0) = 0\}$. It is clearly that for the last set we will have $\{\mathbf{c}_\Omega^0: f_r(\mathbf{c}_\Omega^0) = 0\} = \text{Fr}(\Omega)$, that means all the points h_p^1 and h_p^2 will be projected on the boundary of the set Ω : $\text{Fr}(\Omega)$. Since for all points $h_p^1, h_p^2 \in \mathcal{H}_{12}$ is executed the condition $f_r(\mathbf{c}_p) \leq 0$, then it follows the set \mathcal{H}_{12} is projected completely on the set $\Omega \subset \mathcal{C}^{n-1}$, where $\Omega = \overline{\Omega}$ is a closed set.

By condition the set \mathcal{H}_{12} contains the images of all the hyperplanes which separate strictly or unstrictly S_1 and S_2 , because of that we will analyze only this set. Since for the function $f_r(\mathbf{c})$ we have $f_r(\mathbf{c}) \leq 0, \forall \mathbf{c} \in \Omega$, where $h_1[\mathbf{c}, f_1(\mathbf{c})] \in \text{Fr}(\mathcal{H}_1) \subset \mathcal{H}_{12}$ and $h_2[\mathbf{c}, f_2(\mathbf{c})] \in \text{Fr}(\mathcal{H}_2) \subset \mathcal{H}_{12}$, then $\min_{\mathbf{c}} f_r(\mathbf{c}) < 0$, for $\mathbf{c} \in \Omega$.

If we have in mind the equation $f_r(\mathbf{c}) = f_1(\mathbf{c}) - f_2(\mathbf{c})$, we will receive $f_r(\mathbf{c}_e) = \min_{\mathbf{c}} f_r(\mathbf{c}) = \min_L (L_{e1} - L_{e2}) < 0$ for $\mathbf{c}_e \in \Omega$, from where for $|L_{e1} - L_{e2}|$ in the area Ω , for the points $h_e^1(\mathbf{c}_e, L_{e1})$ and $h_e^2(\mathbf{c}_e, L_{e2})$ we will have $|L_{e1} - L_{e2}| = \max_L |L_1 - L_2|, \mathbf{c}_e \in \Omega$

Ω . The last equation defines in the space E^n the condition: $|y_e^1 - y_e^2| = \max_y |y_p^1 - y_p^2|$ for the dividing S_1 and S_2 mutually parallel hyperplanes H_p^1 and H_p^2 , in accordance with the equation (2): $|y_e^1 - y_e^2| = |L_{e1} - L_{e2}|$ By this the Theorem 2 is proved completely ■

By means of this theorem we can determine the correspondence between the extremum point c_e of the function $f_r(c)$ in the space \mathcal{L}^n and the orientation of the best separating the sets S_1 and S_2 , mutually parallel hyperplanes in the object space E^n .

5. Representation of the classification hyperplanes for non-linear separable sets in the Hough space

If the compact sets S_1 and S_2 defined in the object space E^n are such that $S_1 \cap S_2 = \emptyset$ but for their covers $\text{co}S_1$ and $\text{co}S_2$ the condition $\text{Int}(\text{co}S_1) \cap \text{Int}(\text{co}S_2) \neq \emptyset$ is examined, then, according to the Statement 2, for the sets \mathcal{H}_1 and \mathcal{H}_2 (in the space \mathcal{L}^n) the condition

$$\mathcal{H}_1 \cap \mathcal{H}_2 = \emptyset \text{ will be valid, where } \mathcal{H}_1 = T(\mathbb{H}_1) \text{ and } \mathcal{H}_2 = T(\mathbb{H}_2).$$

For the functions $f_1(c)$ and $f_2(c)$ we will have the inequalities $f_1(c) > L_i > f_2(c) \Rightarrow f_1(c) > f_2(c)$, where $h_i(c, L_i) \in \mathcal{H}_i$, and \mathcal{H}_i is any hyperplane strongly divided the sets \mathcal{H}_1 and \mathcal{H}_2 . From these inequalities follows the condition $f_r(c) = f_1(c) - f_2(c) > 0$, where $f_r(c)$ is a convex function in the space \mathcal{L}^n . That means for the extremum of this function (in this case-minimum) we will have: $f_r(c_e) = \min_c f_r(c) > 0$. The reverse

image of the minimum point $h_e^r [c_e, f_r(c_e)]$ of the function $f_r(c)$ will represent the hyperplane H_r , which will intercept the set S_{12} in the space E^n , where $S_{12} = \text{co}S_1 \cap \text{co}S_2 \neq \emptyset$.

Before to analyze the properties of this characteristic hyperplane, which classifies (although badly) the two sets S_1 and S_2 in the space E^n , we will define the concept for the best dividing S_1 and S_2 hypersurface among the family \mathcal{P} of all dividing hypersurfaces.

Definition 1. We shall call $P_o, P_o \subset \mathbb{P}$, a hypersurface which divides the sets S_1 and S_2 in E^n , in the best way if for its function $P_o(x)$ is examined the condition:

$$\mathbf{V}[P_o(x), \omega] \leq \mathbf{V}[P_i(x), \omega], \forall \omega \subset \Omega = \mathbf{X}_{12}, x \in \omega, \text{ where:}$$

$\mathbf{X}_{12} = \Omega = \text{Pr}_X(S_{12})$ – the project of the set S_{12} in to the subspace $\mathbf{X}^{n-1} \subset E^n$; $\{\omega\}$ – subsets of Ω ,

$\mathbf{V}(P_i, \omega)$ is the $(n-1)$ -th variation of the function $P_i(\mathbf{x})$, defining by the similar way as in [4]; $\mathbf{V}(P_i, \omega) = (1/K) \int_{\mathbf{B}} \mathbf{V}_1(P_i, \mathbf{I}) d\mu_{\mathbf{B}}$; $\mathbf{I} = \omega \cap \beta_1$; β_1 – in this case straight line,

$$\int_{\mathbf{B}} \mathbf{V}_1(\dots) d\mu_{\mathbf{B}} - \text{Lebsegue integral [5],}$$

$\mathbf{B} = \{\beta_1: \beta_1 \subset \mathbf{X}^{n-1}\}$ – the space of all lines in the subspace $\mathbf{X}^{n-1} \subset E^n$,

$\mathbf{V}_1(P_i, \mathbf{I})$ – the first variation of the function $P_i(\mathbf{x})$ for $\mathbf{x} \in \mathbf{I}$, which as is known from [6], for the one-dimensional area \mathbf{I} is equal to the variation of one-dimensional cross-section of this function (Banach theorem [6]); $\mu_{\mathbf{B}}$ – Lebsegue measure (in this case linear) of the set ω in the space \mathbf{B} ,

$$K = \int_{\mathbf{B}} \mathbf{V}_0(\mathbf{I}) d\mu_{\mathbf{B}} - \text{the first variation of the set } \omega.$$

$\mathbf{V}_0(\mathbf{I})$ – zero-variation of the set $\omega \cap \beta_1$ – (component number in \mathbf{I}).

We will formulate three important properties of this hypersurface, which are examined and proved in detail in [7], and which we will use further in the proof of the Theorem 3.

Property 1. $F_2(\mathbf{x}) \leq P_0(\mathbf{x}) \leq F_1(\mathbf{x})$, $\mathbf{x} \in \mathbf{X}_{12}$; where: $F_1(\mathbf{x})$ and $F_2(\mathbf{x})$ are the functions describing the hypersurface $\text{Fr}(\text{coS}_1)$ and $\text{Fr}(\text{coS}_2)$ accordingly, in the area \mathbf{X}_{12} .

Property 2. If we define the sets $s_1 = \{s_{i1}: s_{i1} \subset \text{Fr}(S_1) \cap \text{Fr}(\text{coS}_1) \cap \text{coS}_2\}$ and $s_2 = \{s_{j2}: s_{j2} \subset \text{Fr}(S_2) \cap \text{Fr}(\text{coS}_2) \cap \text{coS}_1\}$, where $\bigcup_{j1} s_{i1} = \text{Fr}(S_1) \cap \text{Fr}(\text{coS}_1) \cap \text{coS}_2$

and $\bigcup_{j2} s_{j2} = \text{Fr}(S_2) \cap \text{Fr}(\text{coS}_2) \cap \text{coS}_1$, then for these sets will be valid the

conditions: $s_1 \subset \mathbf{P}_0$ and $s_2 \subset \mathbf{P}_0$.

Property 3. Every supporting to the set $S_{12} = \text{coS}_1 \cap \text{coS}_2$ hyperplane H , will be simultaneously supporting to the separating, the two sets S_1 and S_2 , hypersurface \mathbf{P}_0 , where $\text{Pr}_X(\mathbf{P}_0) \subset \mathbf{X}_{12}$.

From Property 3 immediately follows a corollary important from a theoretical point of view [7]:

Corollary 3.1. Every supporting to the set S_{12} hyperplane H , will be tangential to the function $P_0(\mathbf{x})$, which defines analytically the hypersurface \mathbf{P}_0 in the area \mathbf{X}_{12} .

Theorem 3. If $\mathbf{c}_m \subset \mathcal{C}^{n-1}$ is such argument of the function $f_r(\mathbf{c})$ that $f_r(\mathbf{c}_m) = \min_{\mathcal{C}} f_r(\mathbf{c})$ in \mathcal{L}^n , than in the space E^n the vector \mathbf{c}_m will coincide by direction with the normal vector of the hyperplane H_0 of the best approximation by Chebishev for the function $P_0(\mathbf{x})$ in E^n , i. e.

$$\max |P_0(\mathbf{x}) - H_0(\mathbf{x})| = \min_H [\max_x |P_0(\mathbf{x}) - H(\mathbf{x})|], \text{ for } \mathbf{x} \in \mathbf{X}_{12},$$

where $H_0(\mathbf{x})$ and $H(\mathbf{x})$ are denoted the equations of the appropriate hyperplanes.

P r o o f. Let us first examine the set $\text{Fr}(\text{coS}_1) \cap \text{coS}_2$ and respectively – the set s_1 . From Property 2 and Property 3 it is clear that the set s_1 contains all

extremum points of the set $\text{Fr}(\text{co}S_1) \cap \text{co}S_2$. Then each supporting to this set hyperplane H_{1t} will contain obligatory at least one point $s_1 \in s_1$, i. e. $\forall H_{1t} \cap s_1 \neq \emptyset$ and will be supporting also to P_o (Property 3). Along with this, from the execution of the condition $s_1 \subset P_o(\mathbf{x})$, according to Corollary 3.1, every such hyperplane will be supporting to the function $P_o(\mathbf{x})$, describing the hypersurface P_o . For the function $P_o(\mathbf{x})$, according to Property 1 we have the inequality $P_o(\mathbf{x}) \leq F_1(\mathbf{x})$. For the function $F_1(\mathbf{x})$, describing this part from the boundary of the set $\text{co}S_1$, which is projected on the area \mathbf{X}_{12} , i. e. the hypersurface $\text{Fr}(\text{co}S_1) \cap \text{co}S_2$, we will have the inequality: $F_1(\mathbf{x}) \leq \mathbf{F}_1(\mathbf{x}) \Rightarrow P_o(\mathbf{x}) \leq \mathbf{F}_1(\mathbf{x})$, according to the basic properties of the concave cover of the sets.

Then from the conditions $H_{1t} \cap s_1 \neq \emptyset, \forall H_{1t}$ and $s_1 \subset P_o(\mathbf{x})$, we can conclude that $\forall H_{1t}$ which is supporting to $\text{Fr}(\text{co}S_1) \cap \text{co}S_2$, will be simultaneously supporting to the hypersurface P_o . That means the images $T(H_{1t})$ of the supporting hyperplanes H_{1t} in the space \mathcal{L}^n , which are such that: $\{T(H_{1t})\} \subset f_1(\mathbf{c})$, will reflect the basic form of $\mathbf{F}_1(\mathbf{x})$, and to some extent the form of $P_o(\mathbf{x})$, defined in E^n . The same reasonings will be valid for the set s_2 with the difference that we will have in E^n the inequalities $\mathbf{F}_2(\mathbf{x}) \leq F_2(\mathbf{x}) \leq P_o(\mathbf{x}) \Rightarrow \mathbf{F}_2(\mathbf{x}) \leq P_o(\mathbf{x})$, where the function $\mathbf{F}_2(\mathbf{x})$ describes the set (the surface) $\text{Fr}(\text{co}S_2) \cap \text{co}S_1$ accordingly. By analogy in the space \mathcal{L}^n for the images $T(H_{2t})$ of the supporting hyperplanes H_{2t} will be valid the condition: $\{T(H_{2t})\} \subset f_2(\mathbf{c})$.

Let us examine in detail the functions $f_1(\mathbf{c})$ and $f_2(\mathbf{c})$, and their difference – the function $f_r(\mathbf{c})$: $f_r(\mathbf{c}) = f_1(\mathbf{c}) - f_2(\mathbf{c})$. In the extremum point \mathbf{c}_m , for which $f_r(\mathbf{c}_m) = \min_{\mathbf{c}} f_r(\mathbf{c})$, we can define the coordinates of the corresponding points for every of this functions in the following way: $h_{1m}[\mathbf{c}_m, f_1(\mathbf{c}_m)]$, $h_{2m}[\mathbf{c}_m, f_2(\mathbf{c}_m)]$ and $h_{rm}[\mathbf{c}_m, f_r(\mathbf{c}_m)]$, from where it is clear that $L_{rm} = L_{1m} - L_{2m} = \min$.

Then in the space E^n , the difference between the equations (in obvious form) of the hyperplanes $H_{1m} = T^{-1}(h_{1m})$ and $H_{2m} = T^{-1}(h_{2m})$ will be

$$(3) \quad Y_r = Y_1 - Y_2 = (\mathbf{x}^T \mathbf{c}_m + L_{1m}) - (\mathbf{x}^T \mathbf{c}_m + L_{2m}) = L_{1m} - L_{2m} = \min,$$

where $\mathbf{x} = (x_1, x_2, \dots, x_{n-1})$, $\mathbf{c}_m = (c_{1m}, c_{2m}, \dots, c_{n-1, m})$. The linear functions, which define the hyperplanes H , contain in their equations the additive constant L , which is parameter. In these cases, as it is known [8], the couple parallel hyperplanes H_{1m} and H_{2m} , which are supported accordingly from below and from above to the function $P_o(\mathbf{x})$ and execute the condition (3), will be the best one-sided approximation by Chebishev, accordingly from above (H_{1m}) and from below (H_{2m}) for the function $P_o(\mathbf{x})$. Then the equation of the hyperplane H_o of the best approximation by Chebishev for the function $P_o(\mathbf{x})$ in E^n will differ from the equations of hyperplanes H_{1m} and H_{2m} , only with the additive constant L , i. e. for the normal vector \mathbf{c}_o , defined the orientation of this hyperplane in the space E^n , we will have: $\mathbf{c}_o = \mathbf{c}_m$. By this the Theorem 3 is proved ■

From Theorem 3 follows immediately the following statement:

Statement 3. If the hyperplane $H_o \subset E^n$, defined with the equation $H_o(\mathbf{x}) = \mathbf{c}_o^T \mathbf{x} - L_o = 0$, is the hyperplane of the best approximation by Chebishev for the function $P_o(\mathbf{x})$, then the position of H_o in E^n will be defined synonymously from the

coordinates of the point $h_m(\mathbf{c}_m, L_m)$ in the space \mathcal{L}^n : $\mathbf{c}_o = \mathbf{c}_m$, $L_o = L_m$, where: $L_m = [f_1(\mathbf{c}_m) - f_2(\mathbf{c}_m)]/2$, \mathbf{c}_m is extremum point of the function $f_r(\mathbf{c})$ and $f_r(\mathbf{c}_m) = \min_{\mathbf{c}} f_r(\mathbf{c})$.

P r o o f. For the couple parallel hyperplanes H_{1j} and H_{2j} , supported from below and from above to the function $P_o(\mathbf{x})$, we will have $|P_o(\mathbf{x}) - H_{1j}(\mathbf{x})| = |P_o(\mathbf{x}) - H_{2j}(\mathbf{x})| = \sup_j |P_o(\mathbf{x}) - H_j(\mathbf{x})|$, for $\forall H_j \parallel H_{1j} \parallel H_{2j}$, where $H_j \cap P_o \neq \emptyset$ and with $H(\mathbf{x})$ are denoted the equations of the appropriate hyperplanes. As in the concrete case we have the equations $|P_o(\mathbf{x}) - H_{1j}(\mathbf{x})| = |P_o(\mathbf{x}) - H_{2j}(\mathbf{x})| = |H_{1j}(\mathbf{x}) - H_{2j}(\mathbf{x})| = L_{1j} - L_{2j}$, then the minimum difference $|H_{1j}(\mathbf{x}) - H_{2j}(\mathbf{x})|$ by j , according to the Theorem 2, will be gain by $j = m$: $|H_{1m}(\mathbf{x}) - H_{2m}(\mathbf{x})| = L_{1m} - L_{2m} = \min$. For the hyperplane H_o of the best (two-sided) approximation by Chebishev we will have

$$|P_o(\mathbf{x}) - H_o(\mathbf{x})| = \min_x [\max_x |P_o(\mathbf{x}) - H_j(\mathbf{x})|] = (L_{1m} - L_{2m})/2 = L_m,$$

where L_m is a constant term in the equation which defines the function $H_o(\mathbf{x})$; $H_o \cap P_o \neq \emptyset$. From here if we have in mind the equation $f_1(\mathbf{c}_m) - f_2(\mathbf{c}_m) = L_{1m} - L_{2m}$ (used in the proof of the Theorem 2), for the extremum point \mathbf{c}_m of the function $f_r(\mathbf{c})$, finally we receive

$$L_o = L_m = (L_{1m} - L_{2m})/2 = [f_1(\mathbf{c}_m) - f_2(\mathbf{c}_m)]/2,$$

by which Statement 3 is proved ■

6. Conclusion

The results obtained in this paper could be used for theoretical investigations as well as for the creation of practical classification methods. An essential result, directly connected with the real classification problems, is the proof of the fact that the sign of the function $f_r(\mathbf{c})$, defined in the Hough space, is an indicator for the linear separability of the two compact sets, defined in the object space. It means that for them the function $f_r(\mathbf{c})$ is a characteristic function. Besides, if these sets are linearly separable, strictly or non strictly, i. e. if $f_r(\mathbf{c}) \leq 0$, then the extremum point \mathbf{c}_m of the function $f_r(\mathbf{c})$ will, in practice, unambiguously define the optimal separating hyperplane for the two sets.

If the sets are linearly inseparable, then the properties of the optimal hyperplane which is the best Chebishev approximation of the optimal separating hypersurface, could be used mainly for theoretical research. The results of these investigations can be applied in the creation of neural networks for classification as well as in some methods for multidimensional spline approximation.

References

1. Ilchev, V., Z. Ilcheva. Some Basic Invariant Properties of the Multidimensional Hough Transform. – Cybernetics and Information Technologies, Vol. 7, 2007, No 2, 45-56.

2. Ilchev, V. Transformation of Lineary Separable Sets in Hough Spase. – Compt. Rend. Acad. Bulg. Sci., Vol. **56**, 2003, No 1, 31-36.
3. Ilchev, V., Z. Ilcheva. Properties of the Transform of the Separating Hyperplanes in the Hough's Space. Working Paper, Institute of Information Technologies, IIT/WP-155b, December, 2002, Bulgarian Academy of Sciences (in Bulgarian).
4. Vitushkin, A. About Multidimensional Variations. Moscow, Technical-Theoretical Literature Press, 1955 (in Russian).
5. Kolmogorov, A., S. Fomin. Elements of Theory of Functions and Functional Analyses. Moscow, Nauka, 1968 (in Russian).
6. Natanson, I. Theory of Functions of Real Variables. Moscow, Tehnical-Theoretical Literature Press, 1957 (in Russian).
7. Ilchev, V., Z. Ilcheva. Estimation of the Classification Hypersurfaces by the Variations of Their Functions. – In: Proc. of the Institute of Computer and Communication Systems, Vol. **1**, 2007, Sofia, Bulgarian Academy of Sciences, 133-138 (in Bulgarian).
8. Collatz, L., W. Krabs. Approximation Theory. Chebishev's Approximations. Stuttgart, Teubner, 1973 (Russian Edition: Moscow, Nauka, 1978).